Grammatical theory

From transformational grammar to constraint-based approaches

Open Review Part 2

Stefan Müller

Textbooks in Language Sciences 1

language science press

Part II

General discussion

13 The innateness of linguistic knowledge

If we try and compare the theories presented in this book as well as their developments, we notice that there are a number of similarities. In all of the frameworks, there are variants of theories that use feature-value pairs to describe linguistic objects. The syntactic structures assumed are sometimes similar. Nevertheless, there are some differences that have often led to fierce debates between proponents of the various schools. The theories differ with regard to whether they assume transformations, empty heads, phrasal or lexical analyses, binary branching or flat structures.

Every theory has to not only describe natural language, but also explain it. It is possible to formulate infinitely many grammars that can create structures for a given language (see Exercise 1 on page 80). These grammars are *observationally adequate*. A grammar achieves *descriptive adequacy* if it corresponds to observations and the intuitions of native speakers.[1] A linguistic theory is descriptively adequate if it can be used to formulate a descriptively adequate grammar for every natural language. However, grammars achieving descriptive adequacy do not always necessarily reach *explanatory adequacy*. Grammars that achieve explanatory adequacy are those that are compatible with acquisition data, that is, grammars that could plausibly be acquired by human speakers (Chomsky 1965: 24–25).

Chomsky (1965: 25) assumes that children already have domain-specific knowledge about what grammars could, in principle, look like and then extract information about what a given grammar actually looks like from the linguistic input. The most prominent variant of acquisition theory in Mainstream Generative Grammar (MGG) is the Principles & Parameters theory, which claims that parametrized principles restrict the grammatical structures possible and children just have to set parameters during language acquisition (see Section 3.1.2).

Over the years, the innateness hypothesis, also known as nativism, has undergone a number of modifications. In particular, the assumptions about exactly what forms part of this innate linguistic knowledge, so-called Universal Grammar (UG), have often undergone changes.

Nativism is often rejected by proponents of Construction Grammar, Cognitive Grammar as well as many other researchers working in other theories. Other explanations are

[1] This term is not particularly useful as subjective factors play a role. Not everybody finds grammatical theories intuitively correct where it is assumed that every observed order in the languages of the world has to be derived from a common Specifier-Head-Complement configuration, and also only by movement to the left (see Section 4.6.1 for the discussion of such proposals).

offered for the facts normally used to argue for the innateness of grammatical categories, syntactic structures or relations between linguistic objects in syntactic structures. Another point of criticism is that the actual complexity of analyses is blurred by the fact that many of the stipulations are simply made to be part of UG. The following is a caricature of a certain kind of argumentation in GB analyses:

1. I have developed an analysis for the phenomenon P in the language S.
2. The analysis is elegant/conceptually simple/my own[2].
3. There is no possibility to learn the relevant structures or principles.
4. Therefore, the assumptions A_1 through A_n that are made in this analysis must be part of the innate knowledge of speakers.

By attributing arbitrary assumptions to UG, it is possible to keep the rest of the analysis very simple. The following section will briefly review some of the arguments for language-specific innate knowledge. We will see that none of these arguments is universally accepted. In the following chapters, I will discuss fundamental questions about the architecture of grammar, the distinction between competence and performance and how to model performance phenomena, the theory of language acquisition as well as other controversial questions, e. g. whether it is desirable to postulate null elements in linguistic representations and whether language should be explained primarily based on the properties of words or rather phrasal patterns.

Before we turn to all these hotly debated topics, I want to discuss the one that is most fiercely debated, namely the question of innate linguistic knowledge. In the literature, one finds the following arguments for such innate knowledge:

- the existence of syntactic universals,
- the speed of acquisition,
- the fact that there is a 'critical period' for language acquisition,
- the fact that all children learn a language, but primates do not,
- the fact that children spontaneously regularize pidgin languages,
- the localization of language processing in particular parts of the brain,
- the alleged dissociation of language and general cognition:
  - Williams Syndrome,
  - the KE family with FoxP2 mutation and
- the Poverty of the Stimulus Argument.

Pinker (1994) offers a nice overview of these arguments. Tomasello (1995) provides a critical review of this book. The individual points will be discussed in what follows.

[2] Also, see http://www.youtube.com/watch?v=cAYDiPizDIs. 20.03.2010.

13.1 Syntactic universals

The existence of syntactic universals has been taken as an argument for the innateness of linguistic knowledge (e. g. Chomsky 1998: 33; Pinker 1994: 237–238). There are varying claims in the literature with regard to what is universal and language-specific. The most prominent candidates for universals are:[3]

- the Head Directionality Parameter
- $\overline{\mathrm{X}}$ structures
- grammatical functions such as subject or object
- binding principles
- properties of long-distance dependencies
- grammatical morphemes for tense, mood and aspect
- parts of speech
- recursion or self-embedding

These supposed universals will each be discussed briefly in what follows. One should emphasize that there is in no way a consensus that these are universal and that the observed properties actually require the postulation of innate linguistic knowledge.

13.1.1 Head Directionality Parameter

The Head Directionality Parameter was already introduced in Section 3.1.2. The examples in (7) on page 88, repeated below as (1), show that the structures in Japanese are the mirror image of the English structures:

(1) a. be showing pictures of himself

b. zibun -no syasin-o mise-te iru
himself of picture showing be

[3] Frans Plank has an archive of universals in Konstanz (Plank & Filimonova 2000): http://typo.uni-konstanz.de/archive/intro/. On 06.03.2010, it contained 2029 entries. The entries are annotated with regard to their quality, and it turns out that many of the universals are statistical universals, that is, they hold for the overwhelming majority of languages, but there are some exceptions. Some of the universals are marked as almost absolute, that is, very few exceptions are known. 1153 were marked as absolute or absolute with a question mark. 1021 of these are marked as absolute without a question mark. Many of the universals captured are implicational universals, that is, they have the form: If a language has the property X, then it also has the property Y. The universals listed in the archive are, in part, very specific and refer to the diachronic development of particular grammatical properties. For example, the fourth entry states that: *If the exponent of vocative is a prefix, then this prefix has arisen from 1st person possessor or a 2nd person subject.*

In order to capture these facts, a parameter was proposed that is responsible for the position of the head relative to its arguments (e. g. Chomsky 1986b: 146; 1988: 70).

By assuming a Head Directionality Parameter, Radford (1990: 60–61; 1997: 19–22), Pinker (1994: 234, 238), Baker (2003: 350) and other authors claim, either explicitly or implicitly, that there is a correlation between the direction of government of verbs and that of adpositions, that is, languages with verb-final order have postpositions and languages with VO order have prepositions. This claim can be illustrated with the language pair English/Japanese and the examples in (1): The *no* occurs after the pronoun in the prepositional phrase, the noun *syasin-o* 'picture' follows the PP belonging to it, the main verb follows its object and the auxiliary *iru* occurs after the main verb *mise-te*.

A single counter example is enough to disprove a universal claim and in fact, it is possible to find a language that has verb-final order but nevertheless has prepositions. Persian is such a language. An example is given in (2):

(2) man ketâb-â-ro be Sepide dâd-am.
I book-PL-RÂ to Sepide gave-1SG
'I gave the books to Sepide.'

In Section 3.1.7, it was shown that German cannot be easily described with this parameter: German is a verb-final language but has both prepositions and postpositions

Furthermore, Gibson & Wexler (1994: 422) point out that a single parameter for the position of heads would not be enough since complementizers in both English and German/Dutch occur before their complements, however, English is a VO language, whereas German and Dutch count as OV languages.

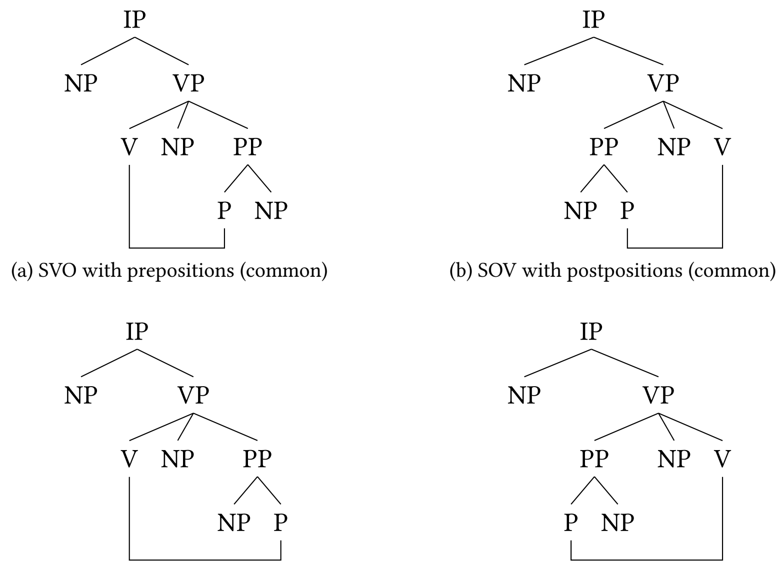
If one wishes to determine the direction of government based on syntactic categories (Gibson & Wexler 1994: 422, Chomsky 2005: 15), then one has to assume that the syntactic categories in question belong to the inventory of Universal Grammar (see Section 13.1.7, for more on this). Difficulties with prepositions and postpositions also arise for this kind of assumption as these are normally assigned to the same category (P). If we were to introduce special categories for both prepositions and postpositions, then a four-way division of parts of speech like the one on page 96 would no longer be possible. One would instead require a binary feature in addition and one would thereby automatically predict eight categories although only five (the four commonly assumed plus an extra one) are actually needed.

One can see that the relation between direction of government that Pinker formulated as a universal claim is in fact a correct tendency, that is, there are many languages where there is a correlation between the use of prepositions or positions and the position the verb.[4]

[4] Pinker (1994: 234) uses the word *usually* in his formulation. He thereby implies that there are exceptions and that it is actually a tendency rather than a universally applicable rule. However, in the pages that follow, he argues that the Head Directionality Parameter forms part of innate linguistic knowledge. Travis (1984: 55) discusses data from Mandarin Chinese that do not correspond to the correlations she assumes. She then proposes treating the Head Directionality Parameter as a kind of Default Parameter that can be overridden by other constraints in the language.

In many languages, adpositions have evolved from verbs. In Chinese grammar, it is commonplace to refer to a particular class of words as coverbs. These are words that can be used both as prepositions and as verbs. If we view languages historically, then we can find explanations for these tendencies that do not have to make reference to innate linguistic knowledge (see Evans & Levinson 2009a: 445).

Furthermore, it is possible to explain these correlations with reference to processing preferences: In languages with the same direction of government, the distance between the verb and the pre-/postposition is less (Figure 13.1a–b) than in languages with differing directions of government (Figure 13.1c–d). From the point of view of processing,



(a) SVO with prepositions (common)
(b) SOV with postpositions (common)
(c) SVO with postpositions (rare)
(d) SOV with prepositions (rare)

Figure 13.1: Distance between verb and preposition for various head orders

languages with the same direction of government should be preferred since they allow the hearer to better identify the parts of the verb phrase (Newmeyer (2004a: 219–221) cites Hawkins (2004: 32) with a relevant general processing preference). This tendency can thus be explained as the grammaticalization of a performance preference (see Chapter 15 for the distinction between competence and performance) and recourse to innate language-specific knowledge is not necessary.

13.1.2 $\overline{\text{X}}$ structures

It is often assumed that all languages have syntactic structures that correspond to the $\overline{\text{X}}$ schema (see Section 2.5) (Pinker 1994: 238; Meisel 1995: 11, 14; Pinker & Jackendoff 2005:

216). There are, however, languages such as Dyirbal (Australia) where it does not seem to make sense to assume hierarchical structure for sentences. Thus, Bresnan (2001: 110) assumes that Tagalog, Hungarian, Malayalam, Warlpiri, Jiwarli, Wambaya, Jakaltek and other corresponding languages do not have a VP node, but rather a rule taking the form of (3):

(3) $S \rightarrow C^*$

Here, C^* stands for an arbitrary number of constituents and there is no head in the structure. Other examples for structures without heads will be discussed in Section 23.10.

$\overline{X}$ structure was introduced to restrict the form of possible rules. The assumption was that these restrictions reduce the class of grammars one can formulate and thus – according to the assumption – make the grammars easier to acquire. But as Kornai & Pullum (1990) have shown, the assumption of $\overline{X}$ structures does not lead to a restriction with regard to the number of possible grammars if one allows for empty heads. In GB, a number of null heads were used and in the Minimalist Program, there has been a significant increase of these. For example, the rule in (3) can be reformulated as follows:

(4) $V' \rightarrow V^0\ C^*$

Here, V^0 is an empty head. Since specifiers are optional, V' can be projected to VP and we arrive at a structure corresponding to the $\overline{X}$ template.

Apart from the problem with languages with very free constituent order, there are further problems with adjunction structures: Chomsky's analysis of adjective structure in $\overline{X}$ Theory (Chomsky 1970: 210; see also Section 2.5 of this book) is not straightforwardly applicable to German since, unlike English, adjective phrases in German are head-final and degree modifiers must directly precede the adjective:

(5) a. der auf seinen Sohn sehr stolze Mann
the of his son very proud man
'The man very proud of his son'

b. * der sehr auf seinen Sohn stolze Mann
the very of his son proud man

c. * der auf seinen Sohn stolze sehr Mann
the of his son proud very man

Following the $\overline{X}$ template, *auf seinen Sohn* has to be combined with *stolze* and only then can the resulting $\overline{A}$ projection be combined with its specifier (see Figure 2.8 on page 77 for the structure of adjective phrases in English). It is therefore only possible to derive orders such as (5b) or (5c). Neither of these is possible in German. It is only possible to rescue the $\overline{X}$ schema if one assumes that German is exactly like English and, for some reason, the complements of adjectives must be moved to the left. If we allow these kind of repair approaches, then of course any language can be described using the $\overline{X}$ template. The result would be that one would have to postulate a vast number of movement rules for many languages and this would be extremely complex and difficult to motivate from a psycholinguistic perspective. See Chapter 15 for grammars compatible with performance.

A further problem for $\overline{\text{X}}$ Theory in its strictest form as presented in Section 2.5 is posed by so-called hydra clauses (Perlmutter & Ross 1970; Link 1984; Kiss 2005):

(6) a. [[der Kater] und [die Katze]], die einander lieben
the tomcat and the cat that each.other love
'the tomcat and the (female) cat that love each other'

b. [[The boy] and [the girl]] who dated each other are friends of mine.

Since the relative clauses in (6) refer to a group of referents, they can only make reference to the result of the coordination. The entire coordination is an NP, however, and adjuncts should actually be attached at the $\overline{\text{X}}$ level. The reverse case of relative clauses in German and English is posed by adjectives in Persian: Samvelian (2007) argues for an analysis where adjectives are combined with nouns directly and only the combination of nouns and adjectives is then combined with a PP argument.

The discussion of German and English shows that the introduction of specifiers and adjuncts cannot be restricted to particular projection levels, and the preceding discussion of non-configurational languages has shown that the assumption of intermediate levels does not make sense for every language.

It should also be noted that Chomsky himself assumed in 1970 that languages can deviate from the $\overline{\text{X}}$ template (1970: 210).

If one is willing to encode all information about combination in the lexicon, then one could get by with very abstract combinatorial rules that would hold universally. Examples for this kind of combinatorial rules are the multiplication rules of Categorial Grammar (see Chapter 8) as well as Merge in the Minimalist Program (see Section 4). The rules in question simply state that two linguistic objects are combined. These kinds of combination of course exist in every language. With completely lexicalized grammars, however, it is only possible to describe languages if one allows for null heads and makes certain ad hoc assumptions. This will be discussed in Section 23.10.

13.1.3 Grammatical functions such as subject and object

Bresnan & Kaplan (1982: xxv), Pinker (1994: 236–237), Baker (2003: 349) and others assume that all languages have subjects and objects. In order to determine what exactly this claim means, we have to explore the terms themselves. For most European languages, it is easy to say what a subject and an object is (see Section 1.7), however, it has been argued that it is not possible for all languages or that it does not make sense to use these terms at all (Croft 2001: Chapter 4; Evans & Levinson 2009a: Section 4).

In theories such as LFG – the one in which Pinker worked – grammatical functions play a primary role. The fact that it is still controversial whether one should view sentences as subjects, objects or as specially defined sentential arguments (XCOMP) (Dalrymple & Lødrup 2000; Berman 2003b; 2007; Alsina, Mohanan & Mohanan 2005; Forst 2006) serves to show that there is at least some leeway for argumentation when it comes to assigning grammatical functions to arguments. It is therefore likely that one can find an assignment of grammatical functions to the arguments of a functor in all languages.

Unlike LFG, grammatical functions are irrelevant in GB (see Williams 1984; Sternefeld 1985a) and Categorial Grammar. In GB, grammatical functions can only be assigned indirectly to the positions in the tree that they correspond to. Thus, in the approach discussed in Chapter 3, the subject is the phrase in the specifier of IP.

In later versions of Chomskyan linguistics, there are functional nodes that seem to correspond to grammatical functions (AgrS, AgrO, AgrIO, see page 146). However, Chomsky (1995: Section 4.10.1) remarks that these functional categories are only assumed for theory internal reasons and should be removed from UG. See Haider (1997a) and Sternefeld (2006: 509–510) for a description of German that does without functional projections that cannot be motivated in the language in question.

The position taken by HPSG is somewhere in the middle: A special valency feature is used for subjects (in grammars of German, there is a head feature that contains a representation of the subject for non-finite verb forms). However, the value of the SUBJ feature is derived from general theoretical considerations: In German, the least oblique element with structural case is the subject (Müller 2002a: 153; Müller 2007b: 311).

In GB theory (Extended Projection Principle, EPP, Chomsky (1982: 10)) and also in LFG (Subject Condition), there are principles ensuring that every sentence must have a subject. It is usually assumed that these principles hold universally.[5]

As previously mentioned, there are no grammatical functions in GB, but there are structural positions that correspond to grammatical functions. The position corresponding to the subject is the specifier of IP. The EPP states that there must be an element in SpecIP. If we assume universality of this principle, then every language must have an element in this position. As we have already seen, there is a counterexample to this universal claim: German. German has an impersonal passive (7a) and there are also subjectless verbs (7b,c) and adjectives (7d–f).[6]

(7) a. dass noch gearbeitet wird
that still worked is
'that people are still working'

b. Ihm graut vor der Prüfung.
him.DAT dreads before the exam
'He dreads the exam.'

c. Mich friert.
me.ACC freezes
'I am freezing.'

[5] However, Chomsky (1981a: 27) allows for languages not to have a subject. He assumes that this is handled by a parameter. Bresnan (2001: 311) formulates the Subject Condition, but mentions in a footnote that it might be necessary to parametrize this condition so that it only holds for certain languages.

[6] For further discussion of subjectless verbs in German, see Haider (1993: Sections 6.2.1, 6.5), Fanselow (2000b), Nerbonne (1986b: 912) and Müller (2007b: Section 3.2).

d. weil schulfrei ist
because school.free is
'because there is no school today'

e. weil ihm schlecht ist
because him.DAT ill is
'because he is not feeling well'

f. Für dich ist immer offen.[7]
for you is always open
'We are always open for you.'

Most of the predicates that can be used without subjects can also be used with an expletive subject. An example is given in(8):

(8) dass es ihm vor der Prüfung graut
that it him before the exam dreads
'He dreads the exam.'

However, there are verbs such as *liegen* 'lie' in example (9a) from Reis (1982: 185) that cannot occur with an *es* ('it').

(9) a. Mir liegt an diesem Plan.
me.DAT lies on this plan
'This plan matters a lot to me.'

b. * Mir liegt es an diesem Plan.
me.DAT lies it on this plan

Nevertheless, the applicability of the EPP and the Subject Condition is sometimes also assumed for German. Grewendorf (1993: 1311) assumes that there is a null expletive that fills the subject position of subjectless constructions.

Berman (1999; 2003a: Chapter 4), working in LFG, assumes that verbal morphology can fulfill the subject role in German and therefore even in sentences where no subject is overtly present, the position for the subject is filled in the f-structure. A constraint stating that all f-structures without a PRED value must be third person singular is responsible for the verbal inflection in subjectless constructions (Berman 1999).

As we saw on page 164, some researchers working in the Minimalist Program even assume that there is an object in every sentence (Stabler quoted in Veenstra (1998: 61, 124)). Objects of mono-valent verbs are assumed to be null elements.

If we allow these kinds of tools, then it is of course easy to maintain the existence of many universals: We claim that a language X has the property Y and then assume that the structural items are invisible and have no meaning. These analyses can only be justified theory-internally with the goal of uniformity (see Culicover & Jackendoff 2005: Section 2.1.2)[8].

[7] Haider (1986a: 18).
[8] For arguments on language acquisition, see Section 16.

13.1.4 Binding principles

The principles governing the binding of pronouns are also assumed to be part of UG (Chomsky 1998: 33; Crain, Thornton & Khlentzos 2009: 146; Rizzi 2009a: 468). Binding Theory in GB theory has three principles: Principle A states that reflexives such as *sich* or *himself* refer to an element (antecedent) inside of a certain local domain (binding domain). Simplyfying a bit one could say that a reflexive has to refer to a co-argument.

(10) Klaus$_i$ sagt, dass Peter$_j$ sich$_{*i/j}$ rasiert hat.
Klaus says that Peter himself shaved has

Principle B holds for personal pronouns and states that these cannot refer to elements inside of their binding domain.

(11) Klaus$_i$ sagt, dass Peter$_j$ ihn$_{i/*j}$ rasiert hat.
Klaus says that Peter him shaved has

Principle C determines what referential expressions can refer to. According to Principle C, an expression A_1 cannot refer to another expression A_2 if A_2 c-commands A_1. c-command is defined with reference to the structure of the utterance. There are various definitions of c-command; a simple version states that A c-commands B if there is a path in the constituent structure that goes upwards from A to the next branching node and then only downwards to B.

For the example in (12a), this means that *Max* and *er* 'he' cannot refer to the same individual since *er* c-commands *Max*.

(12) a. Er sagt, dass Max Brause getrunken hat.
he says that Max soda drunk has
'He said that Max drank soda.'

b. Max sagt, dass er Brause getrunken hat.
Max said that he soda drunk has
'Max said that he drank soda.'

c. Als er hereinkam, trank Max Brause.
as he came.in drank Max soda
'As he came in, Max was drinking soda.'

This is possible in (12b), however, as there is no such c-command relation. For *er* 'he', it must only be the case that it does not refer to another argument of the verb *getrunken* 'drunk' and this is indeed the case in (12b). Similarly, there is no c-command relation between *er* 'he' and *Max* in (12c) since the pronoun *er* is inside a complex structure. *er* 'he' and *Max* can therefore refer to the the same or different individuals in (12b) and (12c).

Crain, Thornton & Khlentzos (2009: 147) point out that (12b,c) and the corresponding English examples are ambiguous, whereas (12a) is not, due to Principle C. This means that one reading is not available. In order to acquire the correct binding principles, the learner would need information about which meanings expressions do not have. The authors note that children already master Principle C at age three and they conclude from this that Principle C is a plausible candidate for innate linguistic knowledge. (This is a classic kind of argumentation. For Poverty of the Stimulus arguments, see Section 13.8 and for more on negative evidence, see Section 13.8.4).

Evans & Levinson (2009b: 483) note that Principle C is a strong cross-linguistic tendency that yet has some exceptions. As an example, they mention both reciprocal expressions in Abaza, where affixes that correspond to *each other* occur in subject position rather than object position as well as Guugu Yimidhirr, where pronouns in a superordinate clause can be coreferent with full NPs in a subordinate clause.

Furthermore, Fanselow (1992b: 351) refers to the examples in (13) that show that Principle C is a poor candidate for a syntactic principle.

(13) a. Mord ist ein Verbrechen.
murder is a crime

b. Ein gutes Gespräch hilft Probleme überwinden.
a good conversation helps problems overcome
'A good conversation helps to overcome problems.'

In (13a), it is expressed that it is a crime when somebody kills someone else, and (13b) refers to conversations with another person rather than talking to oneself. In these sentences, the nominalizations *Mord* 'murder' and *Gespräch* 'conversation' are used without any arguments of the original verbs. So there aren't any arguments that stand in a syntactic command relation to one another. Nevertheless the arguments of the nominalized verbs cannot be coreferential. Therefore it seems that there is a principle at work that says that the argument slots of a predicate must be interpreted as non-coreferential as long as the identity of the arguments is not explicitly expressed by linguistic means.

In sum, one can say that there are still a number of unsolved problems with Binding Theory. The HPSG variants of Principles A–C in English cannot even be applied to German (Müller 1999a: Chapter 20). Working in LFG, Dalrymple (1993) proposes a variant of Binding Theory where the binding properties of pronominal expressions are determined by the lexicon. In this way, the language-specific properties of pronouns can be accounted for.

13.1.5 Properties of long-distance dependencies

It seems to be the case that the long-distance dependencies I discussed in the preceding chapters are subject to some kind of restrictions. For example, nothing can be extracted out of sentences that are part of a noun phrase in English. Ross (1967: 70) calls the relevant constraint the *Complex NP Constraint*. In later work, the attempt was made to group this, and other constraints such as the *Right Roof Constraint* also formulated by Ross (1967: Section 5.1.2), into a single, more general constraint, namely the Subjacency Principle (Chomsky: 1973: 271; 1986a: 40; Baltin 1981; 2006). Subjacency was assumed to hold universally. The Subjacency Constraint states that movement operations can cross at most one bounding node, whereby what exactly counts as a bounding node depends on the language in question (Baltin: 1981: 262; 2006; Rizzi 1982: 57; Chomsky 1986a: 38–40).[9]

Currently, there are varying opinions in the GB/Minimalism tradition with regard to the question of whether subjacency should be considered as part of innate linguistic knowledge. Hauser, Chomsky & Fitch (2002) assume that subjacency does not form part of language-specific abilities, at least not in the strictest sense, but rather is a linguistically relevant constraint in the more broad sense that the constraints in question can be derived from more general cognitive ones (see p. 441). In other contemporary works, subjacency still plays a role as a UG principle (Newmeyer 2005: 15, 74–75; 2004a: 184; Baltin 2006[10]; Baker 2009; Freidin 2009; Rizzi 2009a;b), so that the Subjacency Principle will be discussed here in some further detail.

It is possible to distinguish between two types of movement: Movement to the left (normally called extraction) and movement to the right (normally referred to as extraposition. Both movement types constitute long-distance dependencies. In the following section, I will discuss some of the restrictions on extraposition. Extraction will be discussed in Section 13.1.5.2 following it.

13.1.5.1 Extraposition

Baltin (1981) and Chomsky (1986a: 40) claim that the extraposed relative clauses in (14) have to be interpreted with reference to the embedding NP, that is, the sentences are not equivalent to those where the relative clause would occur in the position marked with t, but rather they correspond to examples where it would occur in the position of the t′.

(14) a. [$_{NP}$ Many books [$_{PP}$ with [stories t]] t′] were sold [that I wanted to read].
b. [$_{NP}$ Many proofs [$_{PP}$ of [the theorem t]] t′] appeared
[that I wanted to think about].

Here, it is assumed that NP, PP, VP and AP are bounding nodes for rightward movement (at least in English) and the interpretation under question here is thereby ruled out by the Subjacency Principle (Baltin 1981: 262).

If we construct a German example parallel to (14a) and replace the embedding noun so that it is ruled out or dispreferred as a referent, then we arrive at (15):

[9] Newmeyer (2004b: 539–540) points out a conceptual problem following from the language-specific determination of bounding nodes: It is argued that subjacency is a language-specific principle since it so abstract that it is impossible for speakers to learn it. However, if parametrization requires that a speaker chooses from a set of categories in the linguistic input, then the corresponding constraints must be derivable from the input at least to the degree that it is possible to determine the categories involved. This raises the question as to whether the original claim of the impossibility of acquisition is actually justified. See Section 13.8 on the *Poverty of the Stimulus* and Section 16.1 on parameter-based theories of language acquisition.

[10] However, see Baltin (2004: 552).

(15) weil viele Schallplatten mit Geschichten verkauft wurden, die ich noch
because many records with stories sold were that I still
lesen wollte
read wanted
'because many records with stories were sold that I wanted to read'

This sentence can be uttered in a situation where somebody in a record store sees particular records and remembers that he had wanted to read those fairy tales on those records. Since one does not read records, adjunction to the superordinate noun is implausible and thus adjunction to *Geschichten* is preferred. By carefully choosing the nouns, it is possible to construct examples such as (16) that show that extraposition can take place across multiple NP nodes:[11]

(16) a. Karl hat mir [ein Bild [einer Frau _$_i$]] gegeben, [die schon lange
Karl has me a picture a woman given that PRT long
tot ist]$_i$.
dead is
'Karl gave me a picture of a woman that has been dead some time.'

b. Karl hat mir [eine Fälschung [des Bildes [einer Frau _$_i$]]] gegeben,
Karl has me a forgery of.the picture of.a woman given
[die schon lange tot ist]$_i$.
that PRT long dead is
'Karl gave me a forgery of the picture of a woman that has been dead for some time.'

c. Karl hat mir [eine Kopie [einer Fälschung [des Bildes [einer Frau
Karl has me a copy of.a forgery of.the picture of.a woman
_$_i$]]]] gegeben, [die schon lange tot ist]$_i$.
given that PRT long dead is
'Karl gave me a copy of a forgery of the picture of a woman that has been dead for some time.'

This kind of embedding could continue further if one were to not eventually run out of nouns that allow for semantically-plausible embedding. NP is viewed as a bounding node in German (Grewendorf: 1988: 81; 2002: 17–18; Haider 2001: 285). These examples show that it is possible for rightward extraposed relative clauses to cross any number of bounding nodes.

Koster (1978: 52–54) discusses some possible explanations for the data in (16), where it is assumed that relative clauses move to the NP/PP border and are then moved on further from there (this movement requires so-called escape hatches or escape routes). He argues that these approaches will also for work for the very sentences that should be ruled out by subjacency, that is, for examples such as (14). This means that either data such as (14)

[11] See Müller (1999a: 211) and Müller: 2004b; 2007d: Section 3. For parallel examples from Dutch, see Koster (1978: 52).

can be explained by subjacency and the sentences in (16) are counterexamples, or there are escape hatches and the examples in (14) are irrelevant, deviant sentences that cannot be explained by subjacency.

In the examples in (16), a relative clause was extraposed in each case. These relative clauses are treated as adjuncts and there are analyses that assume that extraposed adjuncts are not moved but rather base generated in their position and coreference/coindexation is achieved by special mechanisms (Kiss 2005). For proponents of these kinds of analyses, the examples in (16) would be irrelevant to the subjacency discussion as the Subjacency Principle only constrains movement. However, extraposition across phrase boundaries is not limited to relative clauses; sentential complements can also be extraposed:

(17) a. Ich habe [von [der Vermutung $\_{i}$]] gehört, [dass es Zahlen gibt, die
I have from the conjecture heard that there numbers are that
die folgenden Bedingungen erfüllen]$_i$.
the following requirements fulfill

'I have heard of the conjecture that there are numbers that fulfill these requirements.'

b. Ich habe [von [einem Beweis [der Vermutung $\_{i}$]]] gehört, [dass es
I have from a proof of.the conjecture heard that there
Zahlen gibt, die die folgenden Bedingungen erfüllen]$_i$.
numbers are that the following requirements fulfill

'I have heard of the proof of the conjecture that there are numbers that fulfill these requirements.'

c. Ich habe [von [dem Versuch [eines Beweises [der Vermutung $\_{i}$]]]]
I have from the attempt of.a proof of.the conjecture
gehört, [dass es Zahlen gibt, die die folgenden Bedingungen
heard that there numbers are that the following requirements
erfüllen]$_i$.
fulfill

'I have heard of the attempt to prove the conjecture that there are numbers that fulfill these requirements.'

Since there are nouns that embed *zu*-infinitives or prepositional phrases and since these sentences are interpreted as if they were extraposed, it must be ensured that the syntactic category of the postposed element corresponds to the category required by the noun. This means that there has to be some kind of relation between the governing noun and the extraposed element. For this reason, the examples in (17) have to be analyzed as instances of extraposition and provide counter evidence to the claims discussed above.

If one wishes to discuss the possibility of recursive embedding, then one is forced to refer to constructed examples as the likelihood of stumbling across groups of sentences such as those in (16) and (17) is very remote. It is, however, possible to find some individual cases of deep embedding: (18) gives examples of relative clause extraposition and complement extraposition taken from the Tiger corpus[12] (Müller 2007d: 78–79; Meurers & Müller 2009: Section 2.1).

(18) a. Der 43jährige will nach eigener Darstellung damit [$_{NP}$ den Weg [$_{PP}$
the 43.year.old wants after own depiction there.with the way
für [$_{NP}$ eine Diskussion [$_{PP}$ über [$_{NP}$ den künftigen Kurs [$_{NP}$ der
for a discussion about the future course of.the
stärksten Oppositionsgruppierung]]]]]] freimachen, [die aber mit
strongest opposition-grouping free.make that however with
10,4 Prozent der Stimmen bei der Wahl im Oktober weit hinter den
10.4 percent of.the votes at the election in October far behind the
Erwartungen zurückgeblieben war]. (s27639)
expectations stayed.back was

'In his own words, the 43-year old wanted to clear the way for a discussion about the future course of the strongest opposition group that had, however, performed well below expectations gaining only 10.4 percent of the votes.'

b. […] die Erfindung der Guillotine könnte [$_{NP}$ die Folge [$_{NP}$ eines
the invention of.the guillotine could the result of.a
verzweifelten Versuches des gleichnamigen Doktors] gewesen sein, [seine
desperate tries the same.name doctor have been his
Patienten ein für allemal von Kopfschmerzen infolge schlechter
patients once for all.time of headaches following bad
Kissen zu befreien]. (s16977)
pillows to free

'The invention of the guillotine could have been the result of a desperate attempt of the eponymous doctor to rid his patients once and for all of headaches from bad pillows.'

It is also possible to construct sentences for English that violate the Subjacency Condition. Uszkoreit (1990: 2333) provides the following example:

(19) [$_{NP}$ Only letters [$_{PP}$ from [$_{NP}$ those people $\_{i}$]]] remained unanswered [that had received our earlier reply]$_i$.

Jan Strunk has found examples for extraposition of both restrictive and non-restrictive relative clauses across multiple phrase boundaries:

(20) a. For example, we understand that Ariva buses have won [$_{NP}$ a number [$_{PP}$ of [$_{NP}$ contracts [$_{PP}$ for [$_{NP}$ routes in London $\_{i}$]]]]] recently, [which will not be run by low floor accessible buses]$_i$.[13]

b. I picked up [$_{NP}$ a copy of [$_{NP}$ a book $\_{i}$]] today, by a law professor, about law, [that is not assigned or in any way required to read]$_i$.[14]

[12] See Brants et al. (2004) for more information on the Tiger corpus.
[13] http://www.publications.parliament.uk/pa/cm199899/cmselect/cmenvtra/32ii/32115.htm, 24.02.2007.
[14] http://greyhame.org/archives/date/2005/09/, 27.09.2008.

c. We drafted [$_{\mathrm{NP}}$ a list of [$_{\mathrm{NP}}$ basic demands $\_{i}$]] that night [that had to be unconditionally met or we would stop making and delivering pizza and go on strike]$_i$.[15]

The preceding discussion has shown that subjacency constraints on rightward movement do not hold for English or German and thus cannot be viewed as universal. One could simply claim that NP and PP are not bounding nodes in English or German. Then, these extraposition data would no longer be problematic for theories assuming subjacency. However, subjacency constraints are also assumed for leftward movement. We will see this in more detail in the following section.

13.1.5.2 Extraction

Under certain conditions, leftward movement is not possible from certain constituents (Ross 1967). These constituents are referred to as islands for extraction. Ross (1967: Section 4.1) formulated the *Complex NP Constraint* (CNPC) that states that extraction is not possible from complex noun phrases. An example of extraction from a relative clause inside a noun phrase is the following:

(21) * Who$_i$ did he just read [$_{\mathrm{NP}}$ the report [$_{\mathrm{S}}$ that was about $\_{i}$]?

Although (21) would be a semantically plausible question, the sentence is still ungrammatical. This is explained by the fact that the question pronoun has been extracted across the sentence boundary of a relative clause and then across the NP boundary and has therefore crossed two bounding nodes. It is assumed that the CNPC holds for all languages. This is not the cases, however, as the corresponding structures are possible in Danish (Erteschik-Shir & Lappin 1979: 55), Norwegian, Swedish, Japanese, Korean, Tamil and Akan (see Hawkins (1999: 245, 262) and references therein). Since the restrictions on the CNPC are integrated into the Subjacency Principle, then it follows that the Subjacency Principle cannot be universally applicable unless one claims that NP is not a bounding node in the problematic languages. It seems indeed to be the case that the majority of languages do not allow extraction from complex noun phrases. Hawkins explains this on the basis of the processing difficulties associated with the structures in question (Section 4.1). He explains the difference between languages that allow this kind of extraction and languages that do not with reference to the differing processing load for structures that stem from the interaction of extraction with other grammatical properties such as verb position and other conventionalized grammatical structures in the respective languages (Section 4.2).

Unlike extraction from complex noun phrases, extraction across a single sentence boundary (22) is not ruled out by the Subjacency Principle.

(22) Who$_i$ did she think that he saw $\_{i}$?

Movement across multiple sentence boundaries, as discussed in previous chapters, is explained by so-called cyclic movement in transformational theories: A question pronoun

[15] http://portland.indymedia.org/en/2005/07/321809.shtml, 27.09.2008.

is moved to a specifier position and can then be moved further to the next highest specifier. Each of these movement steps is subject to the Subjacency Principle. The Subjacency Principle rules out long-distance movement in one fell swoop.

The Subjacency Principle cannot explain why extraction from sentences embedded under verbs that specify the kind of utterance (23a) or factive verbs (23b) is deviant (Erteschik-Shir & Lappin 1979: 68–69).

(23) a. ?? Who$_i$ did she mumble that he saw $\_{}_i$?
b. ?? Who$_i$ did she realize that he saw $\_{}_i$?

The structure of these sentences seems to be the same as (22). In entirely syntactic approaches, it was also attempted to explain these differences as subjacency violations or as a violation of Ross' constraints. It has therefore been assumed (Stowell 1981: 401–402) that the sentences in (23) have a structure different from those in (22). Stowell treats these sentential arguments of verbs characterizing ways of speaking as adjuncts. Since adjunct clauses are islands for extraction by assumption, this would explain why (23a) is marked. The adjunct analysis is compatible with the fact that these sentential arguments can be omitted:

(24) a. She shouted that he left.
b. She shouted.

Ambridge & Goldberg (2008: 352) have pointed out that treating such clauses as adjuncts is not justified as they are only possible with a very restricted class of verbs, namely verbs of saying and thinking. This property is a property of arguments and not of adjuncts. Adjuncts such as place modifiers are possible with a wide number of verb classes. Furthermore, the meaning changes if the sentential argument is omitted as in (24b): Whereas (24a) requires that some information is communicated, this does not have to be the case with (24b). It is also possible to replace the sentential argument with an NP as in (25), which one would certainly not want to treat as an adjunct.

(25) She shouted the remark/the question/something I could not understand.

The possibility of classifying sentential arguments as adjuncts cannot be extended to factive verbs as these do not allow the sentential argument to be optional (Ambridge & Goldberg 2008: 352):

(26) a. She realized that he left.
b. ?? She realized.

Kiparsky & Kiparsky (1970) suggest an analysis of factive verbs that assumes a complex noun phrase with a nominal head. An optional *fact* Deletion-Transformation removes the head noun and the determiner of the NP in sentences such as (27a) to derive sentences such as (27b) (page 159).

(27) a. She realized [$_{\mathrm{NP}}$ the fact [$_{\mathrm{S}}$ that he left]].

b. She realized [$_{NP}$ [$_{S}$ that he left]].

The impossibility of extraction can then be explained by that fact that movement would have to cross two phrases boundaries (on the island status of this construction, see Kiparsky & Kiparsky 1970: Section 4). This analysis predicts that extraction from complement clauses of factive verbs should be just as bad as extraction from overt NP arguments since the structure for both is the same. According to Ambridge & Goldberg (2008: 353), this is, however, not the case:

(28) a. * Who did she realize the fact that he saw $\_{i}$?
b. ?? Who did she realize that he saw $\_{i}$?

Together with Erteschik-Shir (1981), Erteschik-Shir & Lappin (1979), Takami (1988) and Van Valin (1998), Goldberg (2006: Section 7.2) assumes that there must be a gap in a part of the utterance that can potentially form the focus of an utterance (see Cook (2001), De Kuthy (2002) and Fanselow (2003c) for German). This means that this part cannot be presupposed.[16] If we consider what this means for the data from the subjacency discussion, then one notices that in each case extraction has taken place out of presupposed material:

(29) a. Complex NP
She didn't see the report that was about him. → The report was about him.
b. Complement of a verbs of thinking or saying
She didn't whisper that he left. → He left.
c. Factive verb
She didn't realize that he left. → He left.

Goldberg assumes that constituents that belong to backgrounded information are islands (*Backgrounded constructions are islands* (BCI)). Ambridge & Goldberg (2008) have tested this semantic/pragmatic analysis experimentally and compared it to a purely syntactic approach. They could confirm the fact that information structural properties play a significant role for the extractability of elements. Along with Erteschik-Shir (1973: Section 3.H), Ambridge & Goldberg (2008: 375) assume that languages differ with regard to how much constituents have to belong to background knowledge in order to rule out extraction. In any case we should not rule out extraction from adjuncts for all languages as there are languages such as Danish where it is possible to extract from relative clauses.[17]

[16] Information is presupposed if it is true regardless of whether the utterance is negated or not. Thus, it follows from both (i.a) and (i.b) that there is a king of France.

(i) a. The King of France is bald.
b. The King of France is not bald.

[17] Crain, Khlentzos & Thornton (2010: 2669) claim that it is not possible to extract from relative clauses and the existence of such languages would call into question the very concept of UG. (*If a child acquiring any language could learn to extract linguistic expressions from a relative clause, then this would seriously cast doubt on one of the basic tenets of UG.*) They thereby contradict Evans and Levinson as well as Tomasello, who claim that UG approaches are not falsifiable. If the argumentation of Crain, Khlentzos and Thornton were correct, then (30) would falsify UG and that would be the end of the discussion.

Erteschik-Shir (1973: 61) provides the following examples and more:

(30) a. Det$_i$ er der mange [der kan lide $\_{i}$].
that are there many that can like
'Many people like that.' (lit.: 'That, there are many who like.')

b. Det hus$_i$ kender jeg en mand [som har købt $\_{i}$].
this house know I a man that has bought
'I know a man that has bought this house.' (lit.: 'This house, I know a man that has bought.')

Rizzi's parametrization of the subjacency restriction has been abandoned in many works and the relevant effects have been ascribed to differences in other areas of grammar (Adams 1984; Chung & McCloskey 1983; Grimshaw 1986; Kluender 1992).

We have seen that there are reasons other than syntactic properties of structure as to why leftward movement might be blocked. In addition to information structural properties, processing considerations also play a role (Grosu 1973; Ellefson & Christiansen 2000; Gibson 1998; Kluender & Kutas 1993; Hawkins 1999; Sag et al. 2007). The length of constituents involved, the distance between filler and gap, definiteness, complexity of syntactic structure and interference effects between similar discourse referents in the space between the filler and gap are all important factors for the acceptability of utterances. Since languages differ with regard to their syntactic structure, varying effects of performance, such as the ones found for extraposition and extraction, are to be expected.

In sum, we can say that subjacency constraints do not hold for extraposition in either German or English and furthermore that one can better explain constraints on extraction with reference to information structure and processing phenomena than with the Subjacency Principle. Assuming subjacency as a syntactic constraint in a universal competence grammar is therefore unnecessary to explain the facts.

13.1.6 Grammatical morphemes for tense, mood and aspect

Pinker (1994: 238) is correct in claiming that there are morphemes for tense, mood, aspect, case and negation in many languages. However, there is a great deal of variation with regard to which of these grammatical properties are used in a language and how they are expressed.

For examples of differences in the tense system see Dahl & Velupillai (2013b;a)). Mandarin Chinese is a clear case: It has next to no morphology. The fact that the same morphemes occur in one form or another in almost every language can be attributed to the fact that certain things need to be expressed repeatedly and then things which are constantly repeated become grammaticalized.

13.1.7 Parts of speech

In Section 4.6, so-called cartographic approaches were mentioned, some of which assume over thirty functional categories see Table 4.1 on page 146 for Cinque's functional heads) and assume that these categories form part of UG together with corresponding fixed syntactic structures. Cinque & Rizzi (2010: 55, 57) even assume over 400 functional categories that are claimed to play a role in the grammars of all languages.[18] Also, when formulating principles that are assumed to be universal, specific parts of speech such as Infl (inflection) and Comp (complementizer) are referred to (Baltin: 1981: 262; 2006; Rizzi 1982; Chomsky 1986a: 38; Hornstein 2013: 397).

Chomsky (1988: 68; 1991; 1995: 131), Pinker (1994: 284, 286), Briscoe (2000: 270) and Wunderlich (2004: 621) make comparatively fewer assumptions about the innate inventory of parts of speech: Chomsky assumes that all lexical categories (verbs, nouns, adjectives and adpositions) belong to UG and languages have these at their disposal. Pinker, Briscoe and Wunderlich assume that all languages have nouns and verbs.

This also begs the question as to whether these syntactic categories can be found in other languages in the form that we know them from languages such as German and English.

Braine (1987: 72) argues that parts of speech such as verb and noun should be viewed as derived from fundamental concepts of argument and predicate (also see Wunderlich (2008: 257)). This means that there is an independent explanation for the presence of these categories that is not based on language-specific knowledge.

Evans & Levinson (2009a: Section 2.2.4) discuss the typological literature and give examples of languages which lack adverbs and adjectives. The authors cite Straits Salish as a language in which there may be no difference between verbs and nouns (also, see Evans & Levinson 2009b: 481). They remark that it does make sense to assume the additional word classes ideophone, positional, coverb, classifier on top of the four or five normally used for the analysis of non Indo-European languages.[19] This situation is not a problem for UG-based theories if one assumes that languages can choose from an inventory of possibilities (a toolkit) but do not have to exhaust it (Jackendoff 2002: 263; Newmeyer 2005: 11; Fitch, Hauser & Chomsky 2005: 204; Chomsky 2007: 6–7; Cinque & Rizzi 2010: 55, 58, 65). If we condone this view, then there is a certain arbitrariness. It is possible to assume any parts of speech that one requires for the analysis of at least one language, attribute them to UG and then claim that most (or maybe even all) languages not do not make use of the entire set of parts of speech. This is what is suggested by Villavicencio (2002: 157), working in the framework of Categorial Grammar, for the categories S, NP, N, PP and PRT. This kind of assumption is not falsifiable (see Evans & Levinson 2009a: 436; Tomasello 2009: 471 for a discussion of similar cases and a more general discussion).

Whereas Evans and Levinson assume that one needs additional categories, Haspelmath (2009: 458) and Croft (2009: 453) go so far as to deny the existence of cross-linguistic parts of speech. I consider this is an extreme and believe that a better research strategy is to try and find commonalities between languages.[20] One should, however,

[18] The question of whether these categories form part of UG is left open.

[19] For the opposite view, see Jackendoff & Pinker (2009: 465).

[20] Compare Chomsky (1999: 2):

> In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.

expect to find languages that do not fit into our Indo-European-biased conceptions of grammar.

13.1.8 Recursion and infinitude

In an article in *Science*, Hauser, Chomsky & Fitch (2002) put forward the hypothesis that the only domain-specific universal is recursion. [21] This assumption is controversial and there have been both formal and empirical objections to it.

13.1.8.1 Formal problems

The claim that our linguistic capabilities are infinite is widespread and can already be found in Humboldt's work:[22]

> Das Verfahren der Sprache ist aber nicht bloß ein solches, wodurch eine einzelne Erscheinung zustande kommt; es muss derselben zugleich die Möglichkeit eröffnen, eine unbestimmbare Menge solcher Erscheinungen und unter allen, ihr von dem Gedanken gestellten Bedingungen hervorzubringen. Denn sie steht ganz eigentlich einem unendlichen und wahrhaft grenzenlosen Gebiete, dem Inbegriff alles Denkbaren gegenüber. Sie muss daher von endlichen Mitteln einen unendlichen Gebrauch machen, und vermag dies durch die Identität der gedanken- und spracheerzeugenden Kraft. (Humboldt 1988: 108)

If we just look at the data, we can see that there is an upper bound for the length of utterances. This has to do with the fact that extremely long instances cannot be processed and that speakers have to sleep or will eventually die at some point. If we set a generous maximal sentence length at 100,000 morphemes and then assume a morpheme inventory of X then one can form less than $X^{100,000}$ utterances. We arrive at the number $X^{100,000}$ if we use each of the morphemes at each of the 100,000 positions. Since not all of these sequences will be well-formed, then there are actually less than $X^{100,000}$ possible utterances. This number is incredibly large, but still finite. The same is true of thought: We do not have infinitely many possible thoughts (if *infinitely* is used in the

[21] In a discussion article in *Cognition*, Fitch, Hauser & Chomsky (2005) clarify that their claim that recursion is the only language-specific and human-specific property is a hypothesis and it could be the case that are not any language-specific/species-specific properties at all. Then, a particular combination of abilities and properties would be specific to humans (p. 182–201). An alternative they consider is that innate language-specific knowledge has a complexity corresponding to what was assumed in earlier versions of Mainstream Generative Grammar (p. 182).
Chomsky (2007: 7) notes that Merge could be a non language-specific operation but still attributes it to UG.

[22] The algorithm of language is not simply one where an individual instantiation is created; at the same time it must allow for an indefinite set of such instantiations and must above all allow the expression of the conditions imposed by thought. Language faces an infinite and truly unrestricted domain, the epitome of everything one can think of. Therefore, it must make infinite use of finite means and this is possible through the identity of the power that is responsible for the production of thought and language.

mathematical sense of the word), despite claims by Humboldt and Chomsky (2008: 137) to the contrary.

In the literature, one sometimes finds the claim that it is possible to produce infinitely long sentences (see for instance Nowak, Komarova & Niyogi (2001: 117) and Kim & Sells (2008: 3) and Dan Everett in O'Neill & Wood (2012) at 25:19). This is most certainly not the case. It is also not the case that the rewrite grammars we encountered in Chapter 2 allow for the creation of infinite sentences as the set of symbols of the right-hand side of the rule have to be finite by definition. While it is possible to derive an infinite number of sentences, the sentences themselves cannot be infinite, since it is always one symbol that is replaced by finitely many other symbols and hence no infinite symbol sequence may result.

Chomsky (1965: Section I.1) follows de Saussure (1916b) and draws a distinction between competence and performance: Competence is the knowledge about what kind of linguistic structures are well-formed and performance is the application of this knowledge (see Section 12.6.3 and Chapter 15). Our restricted brain capacity as well as other constraints are responsible for the fact that we cannot deal with an arbitrary amount of embedding and that we cannot produce utterances longer than 100,000 morphemes. The separation between competence and performance makes sense and allows us to formulate rules for the analysis of sentences such as (31):

(31) a. Richard is sleeping.
b. Karl suspects that Richard is sleeping.
c. Otto claims that Karl suspects that Richard is sleeping.
d. Julius believes that Otto claims that Karl suspects that Richard is sleeping.
e. Max knows that Julius believes that Otto claims that Karl suspects that Richard is sleeping.

The rule takes the following form: Combine a noun phrase with a verb of a certain class and a clause. By applying this rule successively, it is possible to form strings of arbitrary length. Pullum & Scholz (2010) point out that one has to keep two things apart: the question of whether language is a recursive system and whether it is just the case that the best models that we can devise for a particular language happen to be recursive. For more on this point and on processing in the brain, see Luuk & Luuk (2011). When constructing strings of words using the system above, it cannot be shown that (a particular) language is infinite, even if this is often claimed to be the case (Pinker 1994: 86; Hauser, Chomsky & Fitch 2002: 1571; Müller 2007b: 1; Hornstein, Nunes & Grohmann 2005: 7; Kim & Sells (2008: 3)).

The "proof" of this infinitude of language is led as an indirect proof parallel to the proof that shows that there is no largest natural number (Pinker 1994: 86). In the domain of natural numbers, this works as follows: Assume x is the largest natural number. Then form x + 1 and, since this is by definition a natural number. We have now found a natural number that is greater than x. We have therefore shown that the assumption that x is the highest number leads to a contradiction and thus that there cannot be such a thing as the largest natural number.

When transferring this proof into the domain of natural language, the question arises as to whether one would still want to class a string of 1,000,000,000 words as part of the language we want to describe. If we do not want this, then this proof will not work.

If we view language as a biological construct, then one has to accept the fact that it is finite. Otherwise, one is forced to assume that it is infinite, but that an infinitely large part of the biologically real object is not biologically real (Postal 2009: 111). Luuk & Luuk (2011) refer to languages as physically uncountable but finite sets of strings. They point out that a distinction must be made between the ability to imagine extending a sentence indefinitely and the ability to take a sentence from a non-countable set of strings and really extend it. We possess the first ability but not the second.

One possibility to provide arguments for the infinitude of languages is to claim that only generative grammars, which create sets of well-formed utterances, are suited to modeling language and that we need recursive rules to capture the data, which is why mental representations have a recursive procedure that generates infinite numbers of expressions (Chomsky, 1956: 115; 2002: 86–87), which then implies that languages consist of infinitely many expressions. There are two mistakes in this argument that have been pointed out by Pullum & Scholz (2010): Even if one assumes generative grammars, it can still be the case that a context-sensitive grammar can still only generate a finite set even with recursive rules. Pullum & Scholz (2010: 120–121) give an interesting example from András Kornai.

The more important mistake is that it is not necessary to assume that grammars generate sets. There are three explicitly formalized alternatives of which only the third is mentioned here, namely the model-theoretic and therefore constraint-based approaches (see Chapter 14). Johnson & Postal's Arc Pair Grammar (1980), LFG in the formalization of Kaplan (1995), GPSG in the reformalization of Rogers (1997) and HPSG with the assumptions of King (1999), Pollard (1999) and Richter (2007) are examples of model-theoretic approaches.

In constraint-based theories, one would say for an example like (31) that certain attitude verbs select a nominative NP and a *that* clause and that these can only occur in certain local configuration where a particular relation holds between the elements involved. One of these relations is subject-verb agreement. In this way, one can represent expressions such as (31) and does not have to say anything about how many sentences can be embedded. This means that constraint-based theories are compatible with every answer to the question of whether there is a finite or infinite amount of structures. Using correspondingly formulated competence grammars, it is possible to develop performance models that explain why certain strings are unacceptable (see Chapter 15).

13.1.8.2 Empirical problems

It is sometimes claimed that all natural languages are recursive and that sentences of an arbitrary length are possible in languages (Hornstein, Nunes & Grohmann 2005: 7 for an overview, and see Pullum & Scholz (2010: Section 2) for further references). When one speaks of recursion, what is often meant are structures with self-embedding as we saw in the analysis of (31) (Fitch 2010). However, it is possible that there are languages that do not allow self-embedding. Everett (2005) claims that Pirahã is such a language (however, see Nevins, Pesetsky & Rodrigues (2009) and Everett (2009)). A further example of a language without recursion, which is sometimes cited with reference to Hale (1976), is Warlpiri. Hale's rules for the combination of a sentence with a relative clause are recursive, however (page 85). This recursion is made explicit on page 98.[23] Pullum & Scholz (2010: 131) discuss Hixkaryana, an Amazonian language from the Caribbean language family that is not related to Pirahã. This language does have embedding, but the embedded material has a different from to that of the matrix clause. It could be the case that these embeddings cannot be carried out indefinitely. In Hixkaryána, there is also no possibility to coordinate phrases or clauses (Derbyshire (1979: 45) cited by Pullum & Scholz (2010: 131)), which is why the possibility of forming recursive sentence embedding does not exist in this language. Other languages without self-embedding seem to be Akkadian, Dyirbal and Proto-Uralic.

There is of course a trivial sense in which all languages are recursive: They follow a rule that says that a particular number of symbols can be combined to form another symbol.[24]

(32) $X \rightarrow X \ldots X$

In this sense, all natural languages are recursive and the combination of simple symbols to more complex ones is a basic property of language (Hockett 1960: 6). The fact that the debate about Pirahã is so fierce could go to show that this is not the kind of recursion that is meant. Also, see Fitch (2010).

It is also assumed that the combinatorial rules of Categorial Grammar hold universally. It is possible to use these rules to combine a functor with its arguments (X/Y * Y = X). These rules are almost as abstract as the rules in (32). The difference is that one of the elements has to be the functor. There are also corresponding restrictions in the Minimalist Program such as selectional features (see Section 4.6.4) and restrictions on the assignment of semantic roles. Whether or not a Categorial Grammar licenses recursive structures does not depend on the very general combinatorial schemata, but rather on the lexical entries. Using the lexical entries in (33), it is only possible to analyze two sentences and certainly not to build recursive structures.

(33) a. the: np/n
b. woman: n

[23] However, he does note on page 78 that relative clauses are separated from the sentence containing the head noun by a pause. Relative clauses in Warlpiri are always peripheral, that is, they occur to the left or right of a sentence with the noun they refer to. Similar constructions can be found in German:

(i) Es war einmal ein Mann. Der hatte sieben Söhne.
there was once a man he had seven sons
'There once was a man. He had seven sons.'

It could be the case that we are dealing with linking of sentences at text level and not recursion at sentence level.

[24] Chomsky (2005: 11) assumes that Merge combines n objects. A special instance of this is binary Merge.

c. cat: n
d. sees: (s\np)/np

If we expand the lexicon to include modifiers of the category n/n or conjunctions of the category (X\X)/X, then we arrive at a recursive grammar.

Fitch, Hauser & Chomsky (2005: 203) note that the existence of languages that do not license recursive structures is not a problem for UG-based theories as not all the possibilities in UG have to be utilized by an individual language. With this view we have actually the same situation that you can posit any number of properties belonging to UG and then decide on a language by language basis whether they play a role or not. The extreme case of this analysis would be that grammars of all languages become part of UG (perhaps with different symbols such as $NP_{Spanish}$, NP_{German}). This variant of a UG-based theory of the human capacity for language would be truly unfalsifiable (Evans & Levinson (2009a: 436, 443); Tomasello (2009: 471)).

13.1.8.3 Recursion in other areas of cognition

There are also phenomena in domains outside of language that can be described with recursive rules: Hauser, Chomsky & Fitch (2002: 1571) mention navigation, family relations and counting systems.[25] One could perhaps argue that the relevant abilities are acquired late and that higher mathematics is a matter of individual accomplishments that do not have anything to do with the cognitive capacities of the majority, but even children at the age of 3 years and 9 months are already able to produce recursive structures: In 2008, there were newspaper reports about an indigenous Brazilian tribe that was photographed from a plane. I showed this picture to my son and told him that Native Americans shot at the plane with a bow and arrow. He then asked me what kind of plane it was. I told him that one cannot see that because the people who took the photograph were sitting in the plane. He then answered that one would then need another plane if one wanted to take a photo that contained both the plane and the Native Americans. He was pleased with his idea and said "And then another one. And then another one. One after the other". He was therefore very much able to imagine the consequence of embeddings.

Culicover & Jackendoff (2005: 113–114) discuss visual perception and music as recursive systems that are independent of language. Jackendoff (2011) extends the discussion of visual perception and music and adds the domains of planning (with the example of making coffee) and wordless comic strips. Chomsky (2007: 7–8) claims that examples from visual perception are irrelevant but then admits that the ability to build up recursive structures could belong to general cognitive abilities (p. 8). He still attributes this

[25] Pinker & Jackendoff (2005: 230) note, however, that navigation differs from the kind of recursive system described by Chomsky and that recursion is not part of counting systems in all cultures. They assume that those cultures have developed infinite counting systems that were made possible by their linguistic capability. This is also assumed by Fitch, Hauser & Chomsky (2005: 203). The latter authors claim that all forms of recursion in other domains depend on language. For more on this point, see Chomsky (2007: 7–8). Luuk & Luuk (2011) note that natural numbers are defined recursively, but the mathematical definition does not necessarily play a role for the kinds of arithmetic operations carried out by humans.

ability to UG. He views UG as a subset of the Faculty of Language, that is, as a subset of non domain-specific abilities (Faculty of Language in the Broad Sense = FLB) and the domain-specific abilities (Faculty of Language in the Narrow Sense = FLN) required for language.

13.1.9 Summary

In sum, we can say that there are no linguistic universals for which there is a consensus that one has to assume domain-specific innate knowledge to explain them. At the 2008 meeting of the *Deutsche Gesellschaft für Sprachwissenschaft*, Wolfgang Klein promised € 100 to anyone who could name him a non-trivial property that all languages share (see Klein 2009). This begs the question of what is meant by 'trivial'. It seems clear that all languages share predicate-argument structures and dependency relations in some sense (Hudson 2010; Longobardi & Roberts 2010: 2701) and, all languages have complex expressions whose meaning can be determined compositionally (Manfred Krifka was promised 20 € for coming up with compositionality). As has been noted at various points, universality by no means implies innateness (Bates 1984: 189; Newmeyer 2005: 205): Newmeyer gives the example that words for sun and moon probably exist in all languages. This has to do with the fact that these celestial bodies play an important role in everyone's lives and thus one needs words to refer to them. It cannot be concluded from this that the corresponding concepts have to be innate. Similarly, a word that is used to express a relation between two objects (e. g. *catch*) has to be connected to the words describing both of these objects (*I*, *elephant*) in a transparent way. This does not necessarily entail, however, that this property of language is innate.

Even if we can find structural properties shared by all languages, this is still not proof of innate linguistic knowledge, as these similarities could be traced back to other factors. It is argued that all languages must be made in such a way as to be acquirable with the paucity of resource available to small children (Hurford 2002: Section 10.7.2; Behrens 2009: 433). It follows from this that our brain in the relevant phases of its development is a constraining factor. Languages have to fit into our brains and since our brains are similar, languages are also similar in certain respects (see Kluender 1992: 251).

13.2 Speed of language acquisition

It is often argued that children learn language extraordinarily quickly and this can only be explained by assuming that they already possess knowledge about language that does not have to be acquired (e. g. Chomsky 1976b: 144; Hornstein 2013: 395). In order for this argument to hold up to scrutiny, it must be demonstrated that other areas of knowledge with a comparable degree of complexity require longer to acquire (Sampson 1989: 214–218). This has not yet been shown. Language acquisition spans several years and it is not possible to simply state that language is acquired following *brief exposure*. Chomsky compares languages to physics and points out that is considerably more difficult for us to acquire knowledge about physics. Sampson (1989: 215) notes, however, that the knowledge about physics one acquires at school or university is not a basis for comparison and one should instead consider the acquisition of everyday knowledge about the physical world around us. For example, the kind of knowledge we need when we want to pour liquids into a container, skip with a skipping rope or the knowledge we have about the ballistic properties of objects. The complexity in comparing these domains of knowledge in order to be able to make claims about language acquisition may turn out to be far from trivial. For an in-depth discussion of this aspect, see Sampson (1989: 214–218). Müller & Riemer (1998: 1) point out that children at the age of six can understand 23,700 words and use over 5000. It follows from this that, in the space of four and a half years, they learn on average 14 new words every day. This is indeed an impressive feat, but cannot be used as an argument for innate linguistic knowledge as all theories of acquisition assume that words have to be learned from data rather than being predetermined by a genetically-determined Universal Grammar. In any case the assumption of genetic encoding would be highly implausible for newly created words such as *fax*, *iPod* or *e-mail*.

Furthermore, the claim that first language acquisition is effortless and rapid when compared to second language acquisition is a myth as has been shown by estimations by Klein (1986: 9): If we assume that children hear linguistic utterances for five hours a day (as a conservative estimate), then in the first five years of their lives, they have 9100 hours of linguistic training. But at the age of five, they have still not acquired all complex constructions. In comparison, second-language learners, assuming the necessary motivation, can learn the grammar of a language rather well in a six-week crash course with twelve hours a day (500 hours in total).

13.3 Critical period for acquisition

Among ducks, there is a critical phase in which their behavior is influenced significantly. Normally, baby ducks follow their mother. If, however, a human is present rather than the mother during a particular time span, the ducks will follow the human. After a certain period, this influence on their behavior can no longer be identified (Lorenz 1970). This kind of critical period can also be identified in other areas of cognition, for example the acquisition of visual abilities among primates. Certain abilities are acquired in a given time frame, whereby the presence of the relevant input is important for determining the start of this time frame.

Lenneberg (1964) claims that language acquisition is only possible up to the age of twelve and concludes from the fact that children can learn language much better than adults that this is also due to a critical period and that language acquisition has similar properties to the behavior of ducks and that hence, the predisposition for language acquisition must be innate (Lenneberg 1967: Chapter 4).

The assumptions about the length of the critical period for language vary considerably. It is possible to find suggestions for 5, 6, 12 and even 15 years (Hakuta et al. 2003: 31). An alternative assumption to the critical period would be to assume that the ability to acquire languages decreases continuously over time. Johnson & Newport (1989) tried to determine a critical period for second-language acquisition and they claim that a second language is learned significantly worse from the age of 15. Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett (1996) have however pointed out that there is also another curve for Johnson and Newport's data that fits the individual data better. The alternative curve shows no abrupt change but rather a steady decrease in the ability to learn language and therefore offers no proof of an effect created by a critical period.

Hakuta, Bialystok & Wiley (2003) evaluate data from a questionnaire of 2,016,317 Spanish speakers and 324,444 speakers of Mandarin Chinese that immigrated to the United States. They investigated which correlations there were between age, the point at immigration, the general level of education of the speakers and the level of English they acquired. They could not identify a critical point in time after which language acquisition was severely restricted. Instead, there is a steady decline in the ability to learn as age increases. This can also be observed in other domains: for example, learning to drive at an older age is much harder.

Summing up, it seems to be relatively clear that a critical period cannot be proven to exist for second-language acquisition. Sometimes, it is assumed anyway that second-language acquisition is not driven by an innate UG, but is in fact a learning process that accesses knowledge already acquired during the critical period (Lenneberg 1967: 176). One would therefore have to show that there is a critical period for first-language acquisition. This is, however, not straightforward as, for ethical reasons, one cannot experimentally manipulate the point at which the input is available. We cannot, say, take 20 children and let them grow up without linguistic input to the age of 3, 4, 5, 6, … or 15 and then compare the results. This kind of research is dependent on thankfully very rare cases of neglect. For example, Curtiss (1977) studied a girl called Genie. At the time, Genie was 13 years old and had grown up in isolation. She is a so-called wolf child. As Curtiss showed, she was no longer able to learn certain linguistic rules. For an objective comparison, one would need other test subjects that had not grown up in complete isolation and in inhumane conditions. The only possibility of gaining relevant experimental data is to study deaf subjects that did not receive any input from a sign language up to a certain age. Johnson & Newport (1989: 63) carried out relevant experiments with learners of American Sign Language. It was also shown here that there is a linear decline in the ability to learn, however nothing like a sudden drop after a certain age or even a complete loss of the ability to acquire language.

13.4 Lack of acquisition among non-human primates

The fact that non-human primates cannot learn natural language is viewed as evidence for the genetic determination of our linguistic ability. All scientists agree on the fact that there are genetically-determined differences between humans and primates and that these are relevant for linguistic ability. Friederici (2009) offers an overview of the literature that claims that in chimpanzees and macaques (and small children), the connections between parts of the brain are not as developed as in adult humans. The connected regions of the brain are together responsible for the processing of lexical-semantic knowledge and could constitute an important prerequisite for the development of language

(p. 179).

The question is, however, whether we differ from other primates in having special cognitive capabilities that are specific to language or whether our capability to acquire languages is due to domain-general differences in cognition. Fanselow (1992b: Section 2) speaks of a human-specific formal competence that does not necessarily have to be specific to language, however. Similarly, Chomsky (2007: 7–8) has considered whether Merge (the only structure-building operation, in his opinion), does not belong to language-specific innate abilities, but rather to general human-specific competence (see, however, Section 13.1.8, in particular footnote 25).

One can ascertain that non-human primates do not understand particular pointing gestures. Humans like to imitate things. Other primates also imitate, however, not for social reasons (Tomasello 2006b: 9–10). According to Tomasello et al. (2005: 676), only humans have the ability and motivation to carry out coordinated activities with common goals and socially-coordinated action plans. Primates do understand intentional actions, however, only humans act with a common goal in mind (*shared intentionality*). Only humans use and understand hand gestures (Tomasello et al. 2005: 685, 724, 726). Language is collaborative to a high degree: Symbols are used to refer to objects and sometimes also to the speaker or hearer. In order to be able to use this kind of communication system, one has to be able to put oneself in the shoes of the interlocutor and develop common expectations and goals (Tomasello et al. 2005: 683). Non-human primates could thus lack the social/cognitive prerequisites for language, that is, the difference between humans and other primates does not have to be explained by innate linguistic knowledge (Tomasello 2003: Section 8.1.2; Tomasello et al. 2005).

13.5 Creole and sign languages

When speakers that do not share a common language wish to communicate with each other, they develop so-called pidgin languages. These are languages that use parts of the vocabularies of the languages involved but have a very rudimentary grammar. It can be noted that children of pidgin speakers regularize these languages. The next generation of speakers creates a new language with an independent grammar. These languages are referred to as *creole languages*. One hypothesis is that the form of languages that develop from creolization is restricted by an innate UG (Bickerton 1984b). It is assumed that the parameter setting of creole languages correspond to the default values of parameters (Bickerton: 1984a: 217; 1984b: 178), that is, parameters already have values at birth and these correspond to the values that creole languages have. The defaults would have to be modified when learning other languages.[26] Bickerton claims that creole languages contain elements that language learners could not have acquired from the input, that is from the pidgin languages. His argumentation is a variant of the classic Poverty of the Stimulus Argument that will be discussed in more detail in Section 13.8.

[26] For problems that can arise from the assumption of defaults values, see Meisel (1995: 17). Bickerton (1997: 56, fn. 13) distances himself from the claim that creole languages have the default values of parameters.

Bickerton's claims have been criticized as it cannot be verified whether children had the input in the individual languages of the adults (Samarin 1984: 207; Seuren 1984: 209). All that can be said considering this lack of evidence is that there are a number of demographic facts that suggest that this was the case for at least some creole languages. This means that children did not only have the strings from the pidgin languages as an input but also sentences from the individual languages. Many creolists assume that adults contribute specific grammatical forms to the emerging language. For example, in the case of Hawaiian Creole English one can observe that there are influences from the mother tongues of the speakers involved: Japanese speakers use SOV order as well as SVO and Philippinos use VOS order as well as SVO order. In total, there is quite a lot of variation in the language that can be traced back to the various native languages of the individual speakers.

It is also possible to explain the effects observed for creolization without the assumption of innate language-specific knowledge: The fact that children regularize language can be attributed to a phenomenon independent of language. In experiments, participants were shown two light bulbs and the test subjects had to predict which of the light bulbs would be turned on next. If one of the bulbs was switched on 70% of the time, the participants also picked this one 70% of the time (although they would have actually had a higher success rate if they had only chosen the bulb turned on with 70% probability). This behavior is known as *Probability Matching*. If we add another light bulb to this scenario and then turn this lamp on in 70% of cases and the other two each 15% of the time, then participants choose the more frequently lit one 80–90% of the time, that is, they regularize in the direction of the most frequent occurrence (Gardner 1957; Weir 1964).

Children regularize more than adults (Hudson & Newport 1999; Hudson Kam & Newport 2005), a fact that can be traced back to their limited brain capacity („less is more"-hypothesis, Newport 1990; Elman 1993).

Like creolization, a similar situation can be found in certain social contexts with the acquisition of sign language: Singleton & Newport (2004) have shown that a child (Simon) that learned American Sign Language (ASL) makes considerably less mistakes than his parents. The parents first learned ASL at the age of 15 or 16 and performed particular obligatory movements only 70% of the time. Simon made these movements 90% of the time. He regularized the input from his parents, whereby the consistent use of form-meaning pairs plays an important role, that is, he does not simply use Probability Matching, but learns selectively. Singleton & Newport (2004: 401) suspect that these kind of regularizations also play a role for the emergence of creole and sign languages. However, the relevant statistical data that one would need to confirm this hypothesis are not available.

13.6 Localization in special parts of the brain

By measuring brain activity during speech production/processing and also by investigating patients with brain damage, one can identify special parts of the brain (Broca's area and Wernicke's area) that play an important role for language production and processing (see Friederici (2009) for a current overview). Chomsky talks about there being a center of language and even calls this (metaphorically) an *Organ* (Chomsky 1977: 164; Chomsky 2005: 1; Chomsky 2008: 133). On the basis of this localization, it has been assumed that there is an innate basis for our linguistic knowledge (also see Pinker 1994: 297–314).

However, it is the case that if these parts are damaged, other areas of the brain can take over the relevant functions. If this damage occurs in early childhood, language can also be learned without these special areas of the brain (for sources, see Dąbrowska 2004: Section 4.1).

Apart from that, it can also be observed that a particular area of the brain is activated when reading. If the conclusion about the localization of processing in a particular part of the brain leading to the innateness of linguistic knowledge were valid, then the activation of certain areas of the brain during reading should also lead us to conclude that the ability to read is innate (Elman et al. 1996; Bishop 2002: 57). This is, however, not assumed (also see Fitch, Hauser & Chomsky 2005: 196).

It should also be noted that language processing affects several areas of the brain and not just Broca's and Wernicke's areas (Fisher & Marcus 2005: 11; Friederici 2009). On the other hand, Broca's and Wernicke's areas are also active during non-linguistic tasks such as imitation, motoric coordination and processing of music (Maess et al. 2001). For an overview and further sources, see Fisher & Marcus (2005).

Musso et al. (2003) investigated brain activity during second-language acquisition. They gave German native speakers data from Italian and Japanese and noticed that there was activation in Broca's area. They then compared this to artificial languages that used Italian and Japanese words but did not correspond to the principles of the Universal Grammar assumed by the authors. An example of the processes assumed in their artificial language is the formation of questions by reversing of word order as shown in (34).

(34) a. This is an statement.
     b. Statement a is this?

The authors then observed that different areas of the brain were activated when learning this artificial language. This is an interesting result, but does not show that we have innate linguistic knowledge. It only shows that the areas that are active when processing our native languages are also active when we learn other languages and that artistry with words such as reversing the order of words in a sentence affects other areas of the brain.

A detailed discussion of localization of languages in particular parts of the brain can be found in Dąbrowska (2004: Chapter 4).

13.7 Differences between language and general cognition

Researchers who believe that there is no such thing as innate linguistic knowledge assume that language can be acquired with general cognitive means. If it can be shown that humans with severely impaired cognition can still acquire normal linguistic abilities or that there are people of normal intelligence whose linguistic ability is restricted, then one can show that language and general cognition are independent.

13.7.1 Williams Syndrome

There are people with a relatively low IQ, who can nevertheless produce grammatical utterances. Among these are people with Williams Syndrome (see Bellugi, Lichtenberger, Jones, Lai & George (2000) for a discussion of the abilities of people with Williams Syndrome). Yamada (1981) takes the existence of such cases as evidence for a separate module of grammar, independent of the remaining intelligence.

IQ is determined by dividing a score in an intelligence test (the mental age) by chronological age. The teenagers that were studied all had a mental age corresponding to that of a four to six year-old child. Yet children at this age already boast impressive linguistic ability that comes close to that of adults in many respects. Gosch, Städing & Pankau (1994: 295) have shown that children with Williams Syndrome do show a linguistic deficit and that their language ability corresponds to what would be expected from their mental age. For problems of sufferers of Williams Syndrome in the area of morphosyntax, see Karmiloff-Smith et al. (1997). The discussion about Williams Syndrome is summarized nicely in Karmiloff-Smith (1998).

13.7.2 KE family with FoxP2 mutation

There is a British family – the so-called KE family – that has problems with language. The members of this family with these linguistic problems have a genetic defect. Fisher et al. (1998) and Lai et al. (2001) discovered that this is due to a mutation of the FoxP2 gene (FoxP2 stands for *Forkhead-Box P2*). Gopnik & Cargo (1991) conclude from the fact that deficits in the realm of morphology are inherited with genetic defects that there must be a gene that is responsible for a particular module of grammar (morphology). Vargha-Khadem et al. (1995: 930) have demonstrated, however, that the KE family did not just have problems with morphosyntax: The affected family members have intellectual and linguistic problems together with motoric problems with facial muscles. Due to the considerably restricted motion in their facial muscles, it would make sense to assume that their linguistic difficulties also stem from motory problems (Tomasello 2003: 285). The linguistic problems in the KE family are not just limited to production problems, however, but also comprehension problems (Bishop 2002: 58).

Nevertheless, one cannot associated linguistic deficiencies directly with FoxP2 as there are a number of other abilities that are affected by the FoxP2 mutation: As well as hindering pronunciation, morphology and syntax, it also has an effect on non-verbal IQ and motory problems with the facial muscles, dealing with non-linguistic tasks, too (Vargha-Khadem et al. 1995).

Furthermore, FoxP2 also occurs in animals. For example, the human gene differs from the analogous gene of a mouse in only three amino acid positions, and those of chimpanzees, gorillas and rhesus apes by only two positions (Enard et al. 2002).

In addition, FoxP2 also occurs in other body tissues: It is also responsible for the development of the lungs, the heart, the intestine and various regions of the brain (Marcus & Fisher 2003). Marcus & Fisher (2003: 260–261) point out that FoxP2 is probably not directly responsible for the development of organs or areas of organs but rather regulates a cascade of different genes. FoxP2 can therefore not be referred to as the language gene, it is just a gene that interacts with other genes in complex ways. It is, among other things, important for our language ability, however, in the same way that it does not make sense to call FoxP2 a language gene, nobody would connect a hereditary muscle disorder with a 'walking gene' just because this myopathy prevents upright walking (Bishop 2002: 58). A similar argument can be found in Karmiloff-Smith (1998: 392): There is a genetic defect that leads to some people beginning to lose their hearing from the age of ten and are then completely deaf at age thirty. This genetic defect causes changes in the hairs inside the ear that one requires for hearing. In this case, one would also not want to talk about a 'hearing gene'.

Fitch, Hauser & Chomsky (2005: 190) are also of the opinion that FoxP2 cannot be responsible for linguistic knowledge. For an overview of this topic, see Bishop (2002) and Dąbrowska (2004: Section 6.4.2.2) and for genetic questions in general, see Fisher & Marcus (2005).

13.8 Poverty of the Stimulus

An important argument for the innateness of the linguistic knowledge is the so-called Poverty of the Stimulus Argument (PSA) (Chomsky 1980: 34). Different versions of it can be found in the literature and have been carefully discussed by Pullum & Scholz (2002). After discussing these variants, they summarize the logical structure of the argument as follows (p. 18):

(35)
  a. Human children learn their first language either by data-driven learning or by learning supported by innate knowledge (a disjunctive premise by assumption)
  b. If children learn their first language by data-driven learning, then they could not acquire anything for which they did not have the necessary evidence (the definition of data-driven learning)
  c. However, children do in fact learn things that they do not seem to have decisive evidence for (empirical prerequisite)
  d. Therefore, children do not learn their first language by data-driven learning. (*modus tollens* of b and c)
  e. Conclusion: Children learn language through a learning process supported by innate knowledge. (disjunctive syllogism of a and d)

Pullum and Scholz then discuss four phenomena that have been claimed to constitute evidence for the fact there is innate linguistic knowledge. These are plurals as initial parts of compounds in English (Gordon 1986), sequences of auxiliaries in English (Kimball

1973), anaphoric *one* in English (Baker 1978) and the position of auxiliaries in English (Chomsky 1971: 29–33). Before I turn to these cases in Section 13.8.2, I will discuss a variant of the PSA that refers to the formal properties of phrase structure grammars.

13.8.1 Gold's Theorem

In theories of formal languages, a language is viewed as a set containing all the expressions belonging to a particular language. This kind of set can be captured using various complex rewrite grammars. A kind of rewrite grammar – so-called context-free grammars – was presented in Chapter 2. In context-free grammars, there is always exactly one symbol on the left-hand side of the rule (a so-called non-terminal symbol) and there can be more of these on the right-hand side of the rule. On the right side there can be symbols (non-terminal symbols) or words/morphemes of the language in question (terminal symbols). The words in a grammar are also referred to as vocabulary (V). Part of a formal grammar is a start symbol that is mostly depicted as S. In the literature, it has been criticized that not all expressions are sentences (see Deppermann 2006: 44). It is, however, not necessary to assume this. It is possible to call the start symbol *Utterance* and define rules that derive S, NP, VP or whatever else one wishes to class as an utterance from *Utterance*.[27]

Beginning with the start symbol, one can keep applying the phrase structure rules in a grammar until one arrives at sequences that only contain words (terminal symbols). The set of all sequences that one can generate are those expressions that form part of the language and that are licensed by the grammar. This set is a subset of all sequences or words or morphemes that can be created by arbitrary combination. The set that contains all possible sequences is referred to as V^*.

Gold (1967) has shown that in an environment E, it is not possible to solve the identification problem for any language from particular languages classes, given a finite amount of linguistic input, without additional knowledge. Gold is concerned with the identification of a language from a given class of languages. A language L counts as identified if from a given point in time t_n, a learner can identify that L is the language in question and does not change this hypothesis. This point in time is not determined in advance, however identification has to take place at some point. Gold calls this identification in the limit. The environments are arbitrary infinite sequences of sentences $\langle\ a_1, a_2, a_3, ...\rangle$, whereby each sentence in the language must occur at least once in this sequence. In order to show that the identification problem cannot be solved for even very simple language classes, Gold considers the class of languages that contain all possible sequences of words from the vocabulary V expect for one sequence: Let V be the vocabulary and x_1, x_2, x_3, ... the sequences of words from this vocabulary. The set of all strings from this vocabulary is V^*. For the class of languages in (36), which consist of all possible sequences of elements in V with the exception of one sequence, it is possible to state a

27 On page 273, we have discussed a description that corresponds to the S symbol in phrase structure grammars. If one omits the specification of head features in the description, then one gets a description of all complete phrases, that is, also *the man* or *now*.

process of how one could learn these languages from a text.

(36) $L_1 = V^* - x_1, L_2 = V^* - x_2, L_3 = V^* - x_3, \ldots$

After every input, one can guess that the language is $V^* - \sigma$, where σ stands for the alphabetically first sequence with the shortest length that has not yet been seen. If the sequence in question occurs later, than this hypothesis is revised accordingly. In this way, one will eventually arrive at the correct language.

If we expand the set of languages from which we have to choose by V^*, then our learning process will no longer work since, if V^* is the target language, then the guessing will perpetually yield incorrect results. If there were a procedure capable of learning this language class, then it would have to correctly identify V^* after a certain number of inputs. Let us assume that this input is x_k. How can the learning procedure tell us at this point that the language we are looking for is not $V^* - x_j$ for $j \neq k$? If x_k causes one to guess the wrong grammar V^*, then every input that comes after that will be compatible with both the correct ($V^* - x_j$) and incorrect (V^*) result. Since we only have positive data, no input allows us to distinguish between either of the hypotheses and provide the information that we have found a superset of the language we are looking for. Gold has shown that none of the classes of grammars assumed in the theory of formal languages (for example, regular, context-free and context-sensitive languages) can be identified after a finite amount of steps given the input of a text with example utterances. This is true for all classes of languages that contain all finite languages and at least one infinite languages. The situation is different if positive and negative data are used for learning instead of text.

The conclusion that has been drawn from Gold's results is that, for language acquisition, one requires knowledge that helps to avoid particular hypotheses from the start. Pullum (2003) criticizes the use of Gold's findings as evidence for the fact that linguistic knowledge must be innate. He lists a number of assumptions that have to be made in order for Gold's results to be relevant for the acquisition of natural languages. He then shows that each of these is not uncontroversial.

1. Natural languages could belong to the class of languages that can be learned from text as opposed to the class of context-free grammars mentioned above.

2. Learners could have information about which sequences of words are not grammatical (see p. 453–454 of Gold's essay for a similar conjecture). As has been shown since then, children do have direct negative evidence and there is also indirect negative evidence (see Section 13.8.4).

3. It is not clear whether learners really restrict themselves to exactly one grammar. Feldman (1972) has developed a learning procedure that eliminates all incorrect grammars at some point and is infinitely many times correct, however, it does not have to always choose one correct grammar and stick to the corresponding hypothesis. Using this procedure, it is possible to learn all recursively enumerative languages, that is, all languages for which there is a generative grammar. Pullum

notes that even Feldman's learning procedure could prove to be too restrictive. It could take an entire lifetime for a learner to reach the correct grammar and they could have incorrect yet increasingly better hypotheses along the way.

4. Learners could work in terms of improvements. If one allows for a certain degree of tolerance, then acquisition is easier and it even becomes possible to learn the class of recursively enumerative languages (Wharton 1974).

5. Language acquisition does not necessarily constitute the acquisition of knowledge about a particular set of sequences, that is, the acquisition of a generative grammar capable of creating this set. The situation is completely different if grammars are viewed as a set of constraints that partially describe linguistic structures, but do not necessarily constitute a concrete set of linguistic structures (for more on this point, see Section 6.7 and Chapter 14).

Furthermore, Pullum notes that it is also possible to learn the class of context-sensitive grammars with Gold's procedure with positive input only in a finite number of steps, if there is a upper bound k for the number of rules, where k is an arbitrary number. It is possible to make k so big that the cognitive abilities of the human brain would not be able to use a grammar with more rules than this. Since it is normally assumed that natural languages can be described by context-sensitive grammars, it can therefore be shown that the syntax of natural languages in Gold's sense can be learned from texts (also, see Scholz & Pullum (2002: 195–196)).

Johnson (2004) adds that there is another important point that has been overlooked in the discussion about language acquisition. Gold's problem of identifiability is different from the problem of language acquisition that has played an important role in the nativism debate. In order to make the difference clear, Johnson differentiates between identifiability (in the Goldian sense) and learnability in the sense of language acquisition. Identifiability for a language class C means that there must be a function f that for each environment E for each language L in C, determines hypothesis L as the target language in a finite amount of time.

Johnson proposes the following as the definition of *learnability* (p. 585): *A class C of natural languages is learnable iff given almost any normal human child and almost any normal linguistic environment for any language L in C, the child will acquire L (or something sufficiently similar to L) as a native language between the ages of one and five years.* Johnson adds the caveat that this definition does not correspond to any theory of learnability in psycholinguistics, but rather it is a hint in the direction of a realistic conception of what acquisition should be.

Johnson notes that in most interpretations of Gold's theorem, identifiability and learnability are viewed as one and the same and shows that this is not logically correct: The main difference between the two depends on the use of two quantifiers. Identifiability of *a* language L from a class C requires that learner converges on L in *every* environment after a finite amount of time. This time can differ greatly from environment to environment. There is not even an upper bound for the time in question. One can simply construct a sequence of environments E_1, E_2, … for L, so that a learner in the environment

E_i will not guess L earlier than the time t_i. Unlike identifiability, learnability means that there is a point in time after which in every normal environment, *every* normal child has converged on the correct language. This means that children acquire their language after a particular time span. Johnson quotes Morgan (1989: 352) claiming that children learn their native language after they have heard approximately 4,280,000 sentences. If we assume that the concept of learnability has a finite upper-bound for available time, then very few language classes can be identified in the limit. Johnson has shown this as follows: Let C be a class of languages containing L and L', where L and L' have some elements in common. It is possible to construct a text such that the first n sentences are contained both in L and in L'. If the learner has L as its working hypothesis then continue the text with sentences from L', if he has L' as his hypothesis, then continue with sentences from L. In each case, the learner has entertained a false hypothesis after n steps. This means that identifiability is not a plausible model for language acquisition.

Aside from the fact that identifiability is psychologically unrealistic, it is not compatible with learnability (Johnson 2004: 586). For identifiability, only one learner has to be found (the function f mentioned above), learnability, however, quantifies over (almost) all normal children. If one keeps all factors constant, then it is easier to show the identifiability of a language class rather than its learnability. On the one hand, identifiability quantifies universally over all environments, regardless of whether these may seem odd or of how many repetitions these may contain. Learnability, on the other hand, has (almost) universal quantification exclusively over normal environments. Therefore, learnability refers to fewer environments than identifiability, such that there are less possibilities for problematic texts that could occur as an input and render a language unlearnable. Furthermore, learnability is defined in such a way that the learner does not have to learn L exactly, but rather learn something sufficiently similar to L. In this sense, learnability is a weaker property of a language class than identifiability. Therefore, learnability does not follow from identifiability nor the reverse.

Finally, Gold is dealing with the acquisition of syntactic knowledge without taking semantic knowledge into consideration. However, children possess a vast amount of information from the context that they employ when acquiring a language (Tomasello et al. 2005). As pointed out by Klein (1986: 44), humans do not learn anything if they are placed in a room and sentences in Mandarin Chinese are played to them. Language is acquired in a social and cultural context.

Concluding, one should note that the existence of innate linguistic knowledge cannot be derived from mathematical findings about the learnability of languages.

13.8.2 Four case studies

Pullum & Scholz (2002) have investigated four prominent instances of the Poverty of the Stimulus Argument in more detail. These will be discussed in what follows Pullum and Scholz's article appeared in a discussion volume. Arguments against the article are addressed by Scholz & Pullum (2002) in the same volume. Further PoS arguments from Chomsky (1986b) and from literature in German have been disproved by Eisenberg

(1992).

13.8.2.1 Plurals in noun-noun compounding

Gordon (1986) claims that compounds in English only allow irregular plurals in compounds, that is, *mice-eater* but ostensibly not *rats-eater*. Gordon claims that compounds with irregular plurals as first element are so rare that children could not have learned the fact that such compounds are possible purely from data.

On pages 25–26, Pullum and Scholz discuss data from English that show that regular plurals can indeed occur as the first element of a compound (*chemicals-maker*, *forms-reader*, *generics-maker*, *securities-dealer*, *drinks trolley*, *rules committee*, *publications catalogue*).[28] This shows that what could have allegedly not been learned from data is in fact not linguistically adequate and one therefore does not have to explain its acquisition.

13.8.2.2 Position of auxiliaries

The second study deals with the position of modal and auxiliary verbs. Kimball (1973: 73–75) discusses the data in (37) and the rule in (38) that is similar to one of the rules suggested by Chomsky (1957: 39) and is designed to capture the following data:

(37) a. It rains.
  b. It may rain.
  c. It may have rained.
  d. It may be raining.
  e. It has rained.
  f. It has been raining.
  g. It is raining.
  h. It may have been raining.

(38) Aux → T(M)(have+en)(be+ing)

T stands for tense, M for a modal verb and *en* stands for the participle morpheme (-*en* in *been*/*seen*/… and -*ed* in *rained*). The brackets here indicate the optionality of the expressions. Kimball notes that it is only possible to formulate this rule if (37h) is well-formed. If this were not the case, then one would have to reorganize the material in rules such that the three cases (M)(have+en), (M)(be+ing) and (have+en)(be+ing) would be covered. Kimball assumes that children master the complex rule since they know that sentences such as (37h) are well-formed and since they know the order in which modal and auxiliary verbs must occur. Kimball assumes that children do not have positive evidence for the order in (37h) and concludes from this that the knowledge about the rule in (38) must be innate.

[28] Also, see Abney (1996: 7) for examples from the Wall Street Journal.

Pullum and Scholz note two problems with this PSA: First, they have found hundreds of examples, among them some from children's stories, so that the Kimball's claim that sentences such as (37h) are „vanishingly rare“ should be called into question. For PSA arguments, one should at least specify how many occurrences there are allowed to be if one still wants to claim that nothing can be learned from them (Pullum & Scholz 2002: 29).

The second problem is that is does not make sense to assume that the rule in (37h) plays a role in our linguistic knowledge. Empirical findings have shown that this rule is not descriptively adequate. If the rule in (38) is not descriptively adequate, then it cannot be explanatory adequate and therefore, one no longer has to explain how it can be acquired.

Instead of a rule such as (38), all theories discussed here currently assume that auxiliary or modal verbs embed a phrase, that is, one does not have an Aux node containing all auxiliary and modal verbs, but rather a structure for (37h) that looks as follows:

(39) It [may [have [been raining]]].

Here, the auxiliary or modal verb always selects the embedded phrase. The acquisition problem now looks completely different: A speaker has to learn the form of the head verb in the verbal projection selected by the auxiliary or modal verb. If this information has been learned, then it is irrelevant how complex the embedded verbal projections are: *may* can be combined with a non-finite lexical verb (37b) or a non-finite auxiliary (37c,d).

13.8.2.3 Reference of *one*

The third case study investigated by Pullum and Scholz deals with the pronoun *one* in English. Baker (1978: 413–425, 327–340) claims that children cannot learn that *one* can refer to constituents larger than a single word as in (40).

(40) a. I would like to tell you another funny story, but I've already told you the only *one* I know.
b. The old man from France was more erudite than the young *one*.

Baker (416–417) claims that *one* can never refer to single nouns inside of NPs and supports this with examples such as (41):

(41) * The student of chemistry was more thoroughly prepared than the one of physics.

According to Baker, learners would require negative data in order to acquire this knowledge about ungrammaticality. Since learners – following his argumentation – never have access to negative evidence, they cannot possibly have learned the relevant knowledge and must therefore already possess it.

Pullum & Scholz (2002: 33) point out that there are acceptable examples with the same structure as these examples:

(42) a. I'd rather teach linguistics to a student of mathematics than to one of any discipline in the humanities.

b. An advocate of Linux got into a heated discussion with one of Windows NT and the rest of the evening was nerd talk.

This means that there is nothing to learn with regard to the well-formedness of the structure in (41). Furthermore, the available data for acquiring the fact that *one* can refer to larger constituents is not as hopeless as Baker (p. 416) claims: There are examples that only allow an interpretation where *one* refers to a larger string of words. Pullum and Scholz offer examples from various corpora. They also provide examples from the CHILDES corpus, a corpus that contains communication with children (MacWhinney 1995). The following example is from a daytime TV show:

(43) A: "Do you think you will ever remarry again? I don't."

B: "Maybe I will, someday. But he'd have to be somebody very special. Sensitive and supportive, giving. Hey, wait a minute, where do they make guys like this?"

A: "I don't know. I've never seen one up close."

Here, it is clear that *one* cannot refer to *guys* since A has certainly already seen *guys*. Instead, it refers to *guys like this*, that is, men who are sensitive and supportive.

Once again, the question arises here as to how many instances a learner has to hear for it to count as evidence in the eyes of proponents of the PSA.

13.8.2.4 Position of auxiliaries in polar questions

The fourth PoS argument discussed by Pullum and Scholz comes from Chomsky and pertains to the position of the auxiliary in polar interrogatives in English. As shown on page 100, it was assumed in GB theory that a polar question is derived by movement of the auxiliary from the I position to the initial position C of the sentence. In early versions of Transformational Grammar, the exact analyses were different but the main point was that the highest auxiliary is moved to the beginning of the clause. Chomsky (1971: 29–33) discusses the sentences in (44) and claims that children know that they have to move the highest auxiliary verb even without having positive evidence for this.[29] If, for example, they entertained the hypothesis that one simply places the first auxiliary at the beginning of the sentence, then this hypothesis would deliver the correct result (44b) for (44a), but not for (44c) since the polar question should be (44d) and not (44e).

(44) a. The dog in the corner is hungry.

b. Is the dog in the corner hungry?

c. The dog that is in the corner is hungry.

d. Is the dog that is in the corner hungry?

[29] Examples with auxiliary movement are used in more recent PoS arguments too, for example in Berwick, Pietroski, Yankama & Chomsky (2011) and Chomsky (2013: 39). Work by Bod (2009b) is not discussed by the authors. For more on Bod's approach, see Section 13.8.3.

e. * Is the dog that in the corner is hungry?

Chomsky claims that children do not have any evidence for the fact that the hypothesis that one simply fronts the linearly first auxiliary is wrong, which is why they could pursue this hypothesis in a data-driven learning process. He even goes so far as to claim that speakers of English only rarely or even never produce examples such as (44d) (Chomsky in Piattelli-Palmarini (1980: 114–115)). With the help of corpus data and plausibly constructed examples, Pullum (1996) has shown that this claim is clearly wrong. Pullum (1996) provides examples from the Wall Street Journal and Pullum & Scholz (2002) discuss the relevant examples in more detail and add to them with examples from the CHILDES corpus showing that adult speakers cannot only produce the relevant kinds of sentences, but also that these occur in the child's input.[30] Examples for entries in CHILDES that disprove the hypothesis that the first auxiliary has to be fronted are given in (45):[31]

(45) a. Is the ball you were speaking of in the box with the bowling pin?
b. Where's this little boy who's full of smiles?
c. While you're sleeping, shall I make the breakfast?

Pullum and Scholz point out that *wh*-questions such as (45b) are also relevant if one assumes that these are derived from polar questions (see page 100 in this book) and if one wishes to show how the child can learn the structure-independent hypothesis. The base form from which (46a) is derived is (46b). If we were to front the first auxiliary in (46b), we would produce (46c).

(46) a. Where's the application Mark promised to fill out?[32]
b. the application Mark [$_{\text{AUX}}$ PAST] promised to fill out [$_{\text{AUX}}$ is] there
c. * Where did the application Mark promised to fill out is?

Evidence for the fact that (46c) is not correct can however also be found in language addressed to children. Pullum and Scholz provide the examples in (47):[33]

(47) a. Where's the little blue crib that was in the house before?
b. Where's the other dolly that was in here?
c. Where's the other doll that goes in there?

These questions have the form *Where's NP?*, where NP contains a relative clause.

In (45c), there is another clause preceding the actual interrogative, an adjunct clause containing an auxiliary as well. This sentence therefore provides evidence for falsehood of the hypothesis that the linearly first auxiliary must be fronted (Sampson 1989: 223).

[30] For more on this point, see Sampson (1989: 223). Sampson cites part of a poem by William Blake, that is studied in English schools, as well as a children's encyclopedia. These examples surely do not play a role in acquisition of auxiliary position since this order is learned at the age of 3;2, that is, it has already been learned by the time children reach school age.

[31] See Lewis & Elman (2001). Researchers on language acquisition agree that the frequency of these kind of examples in communication with children is in fact very low. See Ambridge et al. (2008: 223).

[32] From the transcription of a TV program in the CHILDES corpus.

[33] These sentences are taken from NINA05.CHA in DATABASE/ENG/SUPPES/.

In total, there are a number of attested sentence types in the input of children that would allow them to choose between the two hypotheses. Once again, the question arises as to how much evidence should be viewed as sufficient.

Pullum und Scholz's article has been criticized by Lasnik & Uriagereka (2002) and Legate & Yang (2002). Lasnik and Uriagereka argue that the acquisition problem is much bigger than presented by Pullum and Scholz since a learner without any knowledge about the language he was going to acquire could have the additional hypotheses in (48) and (49):

(48) a. Place the first auxiliary at the front of the clause.
b. Place the first auxiliary in matrix-Infl at the front of the clause.

(49) a. Place any auxiliary at the front of the clause.
b. Place any finite auxiliary at the front of the clause.

Both hypotheses in (49) would be permitted by the sentences in (50):

(50) a. Is the dog in the corner hungry?
b. Is the dog that is in the corner hungry?

They would, however, also allow sentences such as (51):

(51) * Is the dog that in the corner is hungry?

The question that must now be addressed is why all hypotheses that allow (51) should be discarded since the learners do not have any information in their natural-linguistic input about the fact that (51) is not possible. They are lacking negative evidence. If (50b) is present as positive evidence, then this by no means implies that the hypothesis in (48b) has to be the correct one. Lasnik and Uriagereka present the following hypotheses that would also be compatible with (50b):

(52) a. Place the first auxiliary in initial position (that follows a change in intonation).
b. Place the first auxiliary in initial position (that follows the first complete constituent).
c. Place the first auxiliary in initial position (that follows the first parsed semantic unit).

These hypotheses do not hold for sentences such as (53) that contain a conjunction:

(53) Will those who are coming and those who are not coming raise their hands?

The hypotheses in (52) would also allow for sentences such as (54):

(54) * Are those who are coming and those who not coming will raise their hands?

Speakers hearing sentences such as (53) can reject the hypotheses (52) and thereby rule out (54), however, it is still possible to think of analogous implausible hypotheses that are compatible with all previously discussed data.

Legate & Yang (2002) take up the challenge of Pullum and Scholz and explicitly claim how many occurrences one needs to acquire a particular phenomenon. They write the following:

> Suppose we have two independent problems of acquisition, P_1 and P_2, each of which involves a binary decision. For P_1, let F_1 be the frequency of the data that can settle P_1 one way or another, and for P_2, F_2. Suppose further that children successfully acquire P_1 and P_2 at roughly the same developmental stage. Then, under any theory that makes quantitative predictions of language development, we expect F_1 and F_2 to be roughly the same. Conversely, if F_1 and F_2 turn out significantly different, then P_1 and P_2 must represent qualitatively different learning problems.
>
> Now let P_1 be the auxiliary inversion problem. The two choices are the structure-dependent hypothesis (3b-i) and the first auxiliary hypothesis (3a-i). (Legate & Yang 2002: 155)

The position of auxiliaries in English is learned by children at the age of 3;2. According to Legate and Yang, one should also find another acquisition phenomenon that is learned at the age of 3;2. The authors focus on subject drop that is learned at 36 months (two months earlier than auxiliary movement). According to the authors, acquisition problems involve a binary decision: In the first case, one has to choose between the two hypotheses in (48). In the second case, the learner has to determine whether a language uses overt subjects. The authors assume that the use of expletives such as *there* serves as evidence for learners that the language they are learning is not one with optional subjects. They then count the sentences in the CHILDES corpus that contain *there*-subjects and estimate F_2 at 1,2 % of the sentences heard by the learner. Since, in their opinion, we are dealing with equally difficult phenomena here, sentences such as (44d) and (47) should constitute 1.2 % of the input in order for auxiliary movement to be learnable.

The authors then searched in the Nina and Adam corpora (both part of CHILDES) and note that 0,068 to 0,045 % of utterances have the form of (47) and none have the form of (44d). They conclude that this number is not sufficient as positive evidence.

Legate and Yang are right in pointing out that Pullum and Scholz's data from the Wall Street Journal are not necessarily relevant for language acquisition and also in pointing out that examples with complex subject noun phrases do not occur in the data or at least to a negligible degree. There are, however, three serious problems with their argumentation: First, there is no correlation between the occurrence of expletive subjects and the property of being a pro-drop language: Galician (Raposo & Uriagereka 1990: Section 2.5) is a pro-drop language with subject expletive pronouns, in Italian there is an existential expletive *ci*,[34] even though Italian counts as a pro-drop language, Franks

[34] However, *ci* is not treated as an expletive by all authors. See Remberger (2009) for an overview.

(1995) lists Upper and Lower Sorbian as pro-drop languages that have expletives in subject position. Since therefore expletive pronouns have nothing to do with the pro-drop parameter, their frequency is irrelevant for the acquisition of a parameter value. If there were a correlation between the possibility of omitting subjects and the occurrence of subject expletives, then Norwegian and Danish children should learn the fact there has to be a subject in their languages earlier than children learning English since expletives occur to a higher percentage of the time in Danish and Norwegian (Scholz & Pullum 2002: 220). In Danish, the constructions corresponding to *there*-constructions in English are twice as frequent. It is still unclear whether there are actually differences in rate of acquisition (Pullum 2009: 246).

Second, in constructing their Poverty of the Stimulus argument, Legate and Yang assume that there is innate linguistic knowledge (the pro-drop parameter). This argument then becomes circular since it is supposed to show that the assumption of innate linguistic knowledge is indispensable (Scholz & Pullum 2002: 220).

The third problem in Legate and Yang's argumentation is that they assume that a transformational analysis is the only possibility. This becomes clear from the following citation (Legate & Yang 2002: 153):

> The correct operation for question formation is, of course, structure dependent: it involves parsing the sentence into structurally organized phrases, and fronting the auxiliary that follows the subject NP, which can be arbitrarily long:
>
> (4) a. Is [the woman who is singing] e happy?
>   b. Has [the man that is reading a book] e eaten supper?

The analysis put forward by Chomsky (see page 100) is a transformation-based one, that is, a learner has to learn exactly what Legate and Yang describe: The auxiliary must move in front of the subject noun phrase. There are, however, alternative analyses that do not require transformations or equivalent mechanisms. If our linguistic knowledge does not contain any information about transformations, then their claim about what has to be learned is wrong. For example, one can assume as in Categorial Grammar that auxiliaries form a word class with particular distributional properties. One possible placement for them is initial positions as observed in questions, the alternative is after the subject (Villavicencio 2002: 104). There would then be the need to acquire information about whether the subject is realized to the left or to the right of its head. As an alternative to this lexicon-based analysis, one could pursue a Construction Grammar (Fillmore 1988: 44; 1999; Kay & Fillmore 1999: 18), Cognitive Grammar (Dąbrowska 2004: Chapter 9), or HPSG (Ginzburg & Sag 2000) approach. In these frameworks, there are simply two[35] schemata for the two sequences that assign different meanings according to the order of verb and subject. The acquisition problem is then that the learners have to identify the corresponding phrasal patterns in the input. They have to realize that Aux NP VP is a well-formed structure in English that has interrogative semantics. The

[35] Fillmore (1999) assumes subtypes of the Subject Auxiliary Inversion Construction since this kind of inversion does not only occur in questions.

relevant theories of acquisition in the Construction Grammar-oriented literature have been very well worked out (see Section 16.3 and 16.4). Construction-based theories of acquisition are also supported by the fact that one can see that there are frequency effects, that is, auxiliary inversion is first produced by children for just a few auxiliaries and only in later phases of development is it then extended to all auxiliaries. If speakers have learned that auxiliary constructions have the pattern Aux NP VP, then the coordination data provided by Lasnik and Uriagereka in (53) no longer pose a problem since, if we only assign the first conjunct to the NP in the pattern Aux NP VP, then the rest of the coordinate structure (*and those who are not coming*) remains unanalyzed and cannot be incorporated into the entire sentence. The hearer is thereby forced to revise his assumption that *will those who are coming* corresponds to the sequence Aux NP in Aux NP VP and instead to use the entire NP *those who are coming and those who are not coming*. It is therefore enough for acquisition to simply learn the pattern Aux NP VP first for some and then eventually for all auxiliaries in English. This has also been shown by Lewis & Elman (2001) who trained a neural network exclusively with data that did not contain NPs with relative clauses in auxiliary constructions. Relative clauses were, however, present in other structures. The complexity of the training material was increased bit by bit just as is the case for the linguistic input that children receive (Elman 1993).[36] The neural network can predict the next symbol after a sequence of words. For sentences with interrogative word order, the predictions were correct. Even the relative pronoun in (55) was predicted despite the sequence Aux Det N Relp never occurring in the training material.

(55) Is the boy who is smoking crazy?

Furthermore, the system signals an error if the network is presented with the ungrammatical sentence (56):

(56) * Is the boy who smoking is crazy?

A gerund is not expected after the relative pronoun, but rather a finite verb. The constructed neural network is of course not yet an adequate model of what is going on in our heads during acquisition and speech production.[37] The experiment shows, however, that the input that the learner receives contains rich statistical information that can be used when acquiring language. Lewis and Elman point out that the statistical information about the distribution of words in the input is not the only information that speakers have. In addition to information about distribution, they are also exposed to information about the context and can make use of phonological similarities in words.

In connection to the ungrammatical sentences in (56), it has been claimed that the fact that such sentences can never be produced shows that children already know that

[36] There are cultural differences. In some cultures, adults do not talk to children that have not attained full linguistic competence (Ochs 1982; Ochs & Schieffelin 1984) (also see Section 13.8.4). Children have to therefore learn the language from their environment, that is, the sentences that they hear reflect the full complexity of the language.

[37] See Hurford (2002: 324) and Jackendoff (2007: Section 6.2) for problems that arise for certain kinds of neural networks and Pulvermüller (2003; 2010) for an alternative architecture that does not have these problems.

grammatical operations are structure-dependent and this is why they do not entertain the hypothesis that it is simply the linearly first verb that is moved (Crain & Nakayama 1987). The claim simply cannot be verified since children do not normally form the relevant complex utterances. It is therefore only possible to illicit experimentally utterances where they could make the relevant mistakes. Crain & Nakayama (1987) have carried out such experiments. Their study has been criticized by Ambridge, Rowland & Pine (2008) since these authors could show that children do really make mistakes when fronting auxiliaries. They put the difference to the results of the first study by Crain and Nakayama down to unfortunate choice of auxiliary in Crain and Nakayama's study. Due to the use of the auxiliary *is*, the ungrammatical examples had pairs of words that never or only very rarely occur next to each other (*who running* in (57a)).

(57) a. The boy who is running fast can jump high. →
* Is the boy who running fast can jump high?

b. The boy who can run fast can jump high. →
* Can the boy who run fast can jump high?

If one uses the auxiliary *can*, this problem disappears since *who* and *run* certainly do appear together. This then leads to the children actually making mistakes that they should not have as the incorrect utterances actually violate a constraint that is supposed to be part of innate linguistic knowledge.

Estigarribia (2009) investigated English polar questions in particular. He shows that not even half of the polar questions in children's input has the form Aux NP VP (p. 74). Instead, parents communicated with their children in a simplified form and used sentences such as:

(58) a. That your tablet?

b. He talking?

c. That taste pretty good?

Estigarribia divides the various patterns into complexity classes of the following kind: FRAG (*fragmentary*), SPRED (*subject predicate*) and AUX-IN (*auxiliary inversion*). (59) shows corresponding examples:

(59) a. coming tomorrow? (FRAG)

b. you coming tomorrow? (SPRED)

c. Are you coming tomorrow? (AUX-IN)

What we see is that the complexity increases from class to class. Estigarribia suggests a system of language acquisition where simpler classes are acquired before more complex ones and the latter ones develop from peripheral modifications of more simple classes (p. 76). He assumes that question forms are learned from right to left (*right to left elaboration*), that is, (59a) is learned first, then the pattern in (59b) containing a subject in addition to the material in (59a), and then in a third step, the pattern (59c) in which an additional auxiliary occurs (p. 82). In this kind of learning procedure, no auxiliary movement is involved. This view is compatible with constraint-based analyses such as that of Ginzburg & Sag (2000). A similar approach to acquisition by Freudenthal, Pine, Aguado-Orea & Gobet (2007) will be discussed in Section 16.3.

A further interesting study has been carried out by Bod (2009b). He shows that it is possible to learn auxiliary inversion under the assumption of trees with any kind of branching even if there is no auxiliary inversion with complex noun phrases present in the input. The procedure he uses as well as the results he gains are very interesting and will be discussed in Section 13.8.3 in more detail.

In conclusion, we can say that children do make mistakes with the position of auxiliaries that they probably should not make if the relevant knowledge were innate. Information about the statistical distribution of words in the input is enough to learn the structures of complex sentences without actually having this kind of complex sentences in the input.

13.8.2.5 Summary

Pullum & Scholz (2002: 19) show what an Argument from Poverty of the Stimulus (APS) would have to look like if it were constructed correctly:

(60) APS specification schema:
   a. ACQUIRENDUM CHARACTERIZATION: describe in detail what is alleged to be known.
   b. LACUNA SPECIFICATION: identify a set of sentences such that if the learner had access to them, the claim of data-driven learning of the acquirendum would be supported.
   c. INDISPENSABILITY ARGUMENT: give reason to think that if learning were data-driven then the acquirendum could not be learned without access to sentences in the lacuna.
   d. INACCESSIBILITY EVIDENCE: support the claim that tokens of sentences in the lacuna were not available to the learner during the acquisition process.
   e. ACQUISITION EVIDENCE: give reason to believe that the the acquirendum does in fact become known to learners during child-hood.

As the four case studies have shown, there can be reasons for rejecting the acquirendum. If the acquirendum does not have to be acquired, than there is no longer any evidence for innate linguistic knowledge. The acquirendum must at least be descriptively adequate. This is an empirical question that can be answered by linguists. In three of the four PoS arguments discussed by Pullum and Scholz, there were parts, which were not descriptively adequate. In previous sections, we already encountered other PoS arguments that involve claims or arguments that cannot be upheld empirically (for example, the Subjacency Principle). For the remaining points in (60), interdisciplinary work is required: the specification of the lacuna falls into the theory of formal language (the specification of a set of utterances), the argument of indispensability is a mathematical task from the realm of theories of learning, the evidence for inaccessibility is an empirical question that can be approached by using corpora and finally the evidence for acquisition is a question for experimental developmental psychologists (Pullum & Scholz 2002: 19–20).

Pullum & Scholz (2002: 46) point out an interesting paradox with regard to (60c): Without results from mathematical theories of learning, one cannot achieve (60c). If one wishes to provide a valid Poverty of the Stimulus Argument, then this should automatically lead to improvements in theories of learning, that is, it is possible to learn more than was previously assumed.

13.8.3 Unsupervised Data-Oriented Parsing (U-DOP)

Bod (2009b) has developed an algorithm that does not require any information about word classes or relations between words contained in utterances. The only assumption that one has to make is that there is some kind of structure. The procedure consists of three steps:

1. Calculate all possible (binary-branching) trees (without category symbols) for a set of given sentences.
2. Divide these trees into sub-trees.
3. Calculate the ideal tree for each sentence.

This process will be explained using the sentences in (61):

(61) a. Watch the dog.
  b. The dog barks.

The trees that these utterances are assigned to only use the category symbol X since the categories for the relevant phrases are not (yet) known. In order to keep the example readable, the words themselves will not be given the category X, although one can of course do this. Figure 13.2 on the next page shows the trees for (61). In step two the trees are divided into subtrees. The trees in Figure 13.2 have the subtrees that can be seen in Figure 13.3 on page 472. In the third step, we now have to calculate the best tree for each utterance. For *The dog barks.*, there are two trees in the set of the subtrees that correspond exactly to this utterance. But it is also possible to build structures out of subtrees. There are therefore multiple derivations possible for *The dog barks.* all of which use the trees in Figure 13.3: One the one hand, trivial derivations that use the entire tree, and on the other, derivations that build trees from smaller subtrees. Figure 13.4 on page 473 gives an impression of how this construction of subtrees happens: If we now want to decide which of the analyses in (62) is the best, then we have to compute the probability of each tree.

(62) a. [[the dog] barks]
  b. [the [dog barks]]

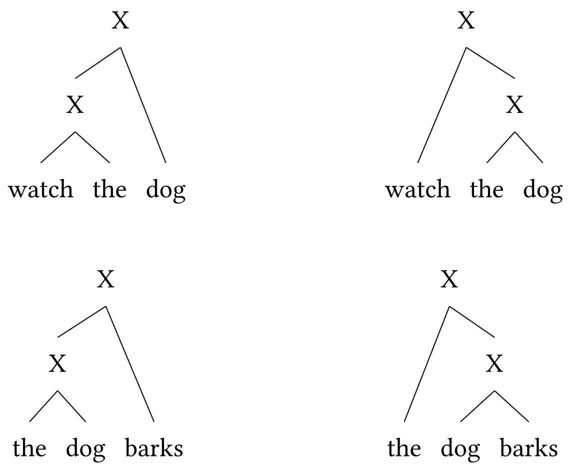


Figure 13.2: Possible binary-branching structures for *Watch the dog* and *The dog barks.*

The probability of a tree is the sum of the probabilities of all its analyses. ´There are two analyses for (62b) that can be found in Figure 13.4. The probability of the first analysis of (62b) corresponds to the probability of choosing exactly the complete tree for [the [dog barks]] from the set of all subtrees. Since there are twelve subtrees, the probability of choosing that one is 1/12. The probability of the second analysis comes from the product of the probability of the subtrees that are combined and is therefore $1/12 \times 1/12 = 1/144$. The probability of the analysis in (62b) is therefore $1/12 + (1/12 \times 1/12) = 13/144$. One can then calculate the probability of the tree in (62a) in the same way. The only difference here is that the tree for [the dog] occurs twice in the set of subtrees. Its probability is therefore 2/12. The probability of the tree [[the dog] barks] is therefore: $1/12 + (1/12 \times 2/12) = 14/144$. We have therefore extracted knowledge about plausible structures from the corpus. This knowledge can also be applied whenever one hears new utterance for which there is no complete tree. It is then possible to use already known subtrees to calculate the probabilities of possible analyses of the new utterance. Bod's model can also be combined with weights: Those sentences that were heard longer ago by the speaker, will receive a lower weight. One can thus account for the fact that children do not simultaneously have all sentences available that they have ever heard. This extension makes the UDOP model more plausible for language acquisition.

In the example above, we did not assign categories to the words. If we were to do this, then we would get the tree in Figure 13.5 on page 473 as a possible subtree. These kind of discontinuous subtrees are important if one wants to capture dependencies between elements that occur in different subtrees of a given tree. Some examples are the following sentences:

(63) a. BA carried *more* people *than* cargo in 2005.

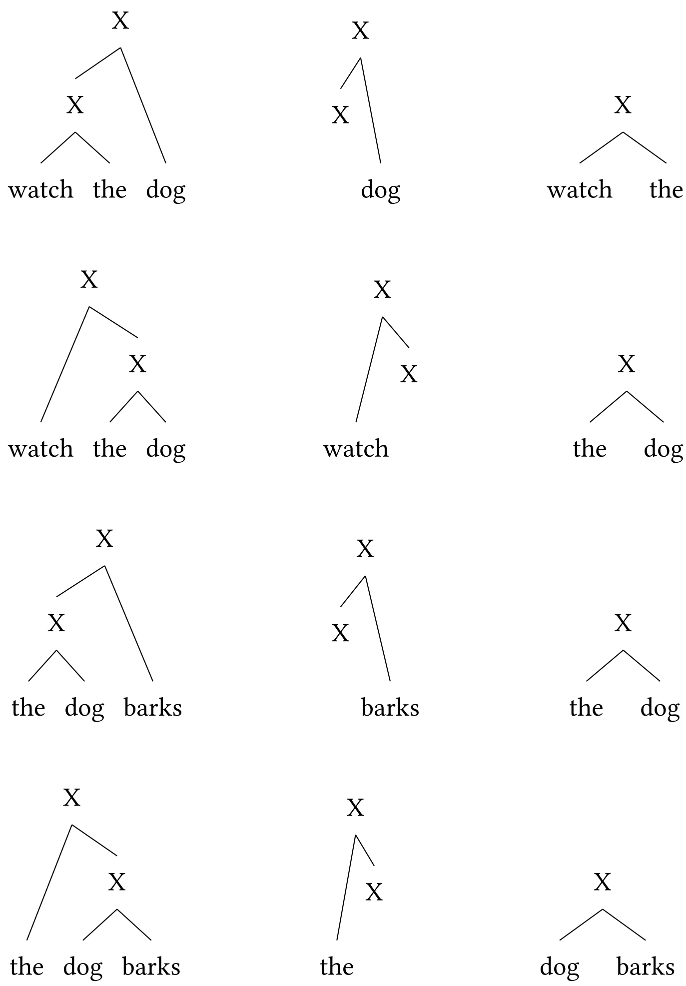


Figure 13.3: Subtrees for the trees in Figure 13.2

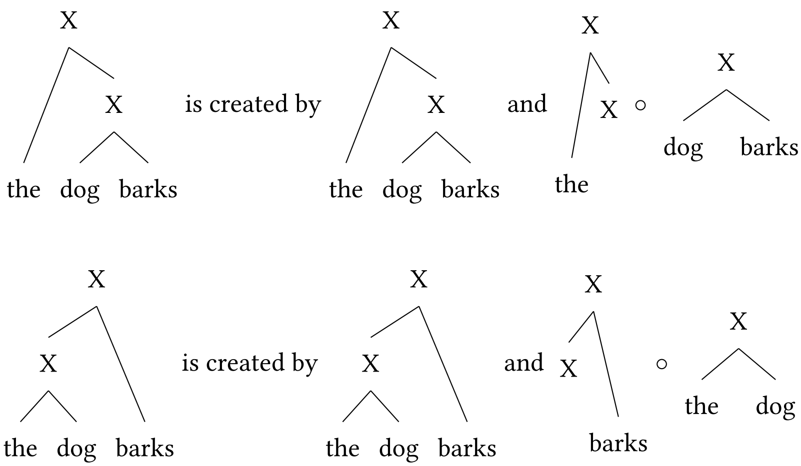


Figure 13.4: Analysis of *The dog barks* using subtrees from Figure 13.3

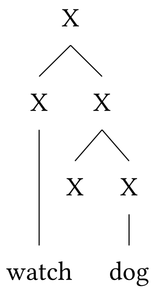


Figure 13.5: Discontinuous partial tree

b. *What's* this scratch *doing* on the table?

c. Most software *companies* in Vietnam *are* small sized.

It is then also possible to learn auxiliary inversion in English with these kind of discontinuous trees. All one needs are tree structures for both sentences in (64) in order to prefer the correct sentence (65a) over the incorrect one (65b).

(64) a. The man who is eating is hungry.

b. Is the boy hungry?

(65) a. Is the man who is eating hungry?

b. * Is the man who eating is hungry?

U-DOP can learn the structures for (64) in Figure 13.6 from the sentences in (66):

(66) a. The man who is eating mumbled.
b. The man is hungry.
c. The man mumbled.
d. The boy is eating.

Note that these sentences do not contain any instance of the structure in (65a). With

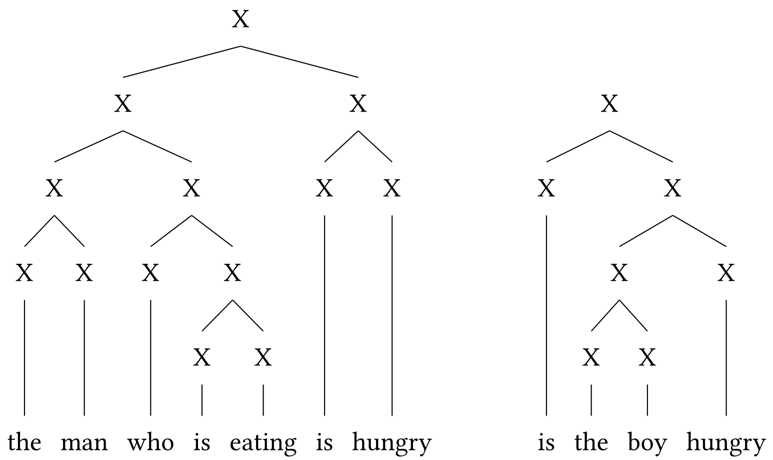


Figure 13.6: Structures that U-DOP learned from the examples in (64) and (66)

the structures learned here, it is possible to show that the shortest possible derivation for the position of the auxiliary is also the correct one: The correct order *Is the man who is eating hungry?* only requires that the fragments in Figure 13.7 on the facing page are combined, whereas the structure for * *Is the man who eating is hungry?* requires at least four subtrees from Figure 13.6 to be combined with each other. This is shown by Figure 13.8 on the next page.

The motivation for always taking the derivation that consists of the least subparts is that one maximizes similarity to already known material.

The tree for (67) containing an auxiliary too much can also be created from Figure 13.6 with just two subtrees (with the tree [$_X$ is$_X$ X] and the entire tree for *The man who is eating is hungry*).

(67) * Is the man who is eating is hungry?

Interestingly, children do produce these kind of incorrect sentences (Crain & Nakayama 1987: 530; Ambridge, Rowland & Pine 2008). However, if we consider the probabilities of the subtrees in addition to the the number of combined subparts, we get the correct result, namely (65a) and not (67). This is due to the fact that *the man who is eating*

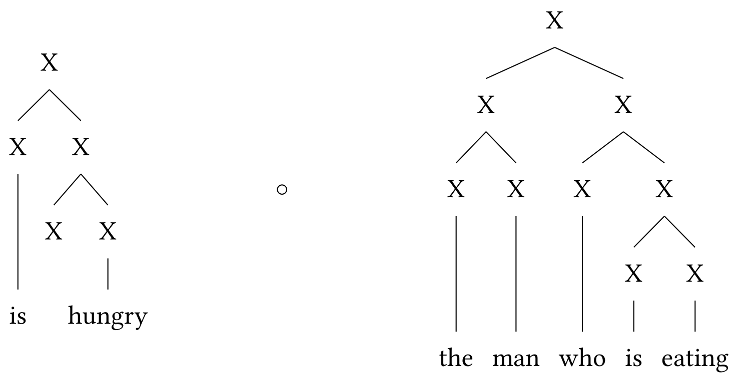


Figure 13.7: Derivation of the correct structure for combination with an auxiliary using two subtrees from Figure 13.6

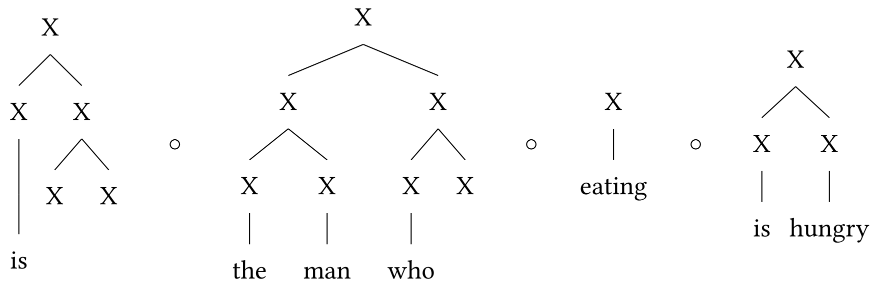


Figure 13.8: Derivation of the incorrect structure for the combination with an auxiliary using two subtrees from Figure 13.6

occurs in the corpus twice, in (65a) and in (66a). Thus, the probability of *the man who is eating* is just as high as the probability of *the man who is eating is hungry* and thus derivation in Figure 13.7 is preferred over the one for (67). This works for the constructed examples here, however one can imagine that in a realistic corpus, sequences of the form *the man who is eating* are more frequent than sequences with further words since *the man who is eating* can also occur in other contexts. Bod has applied this process to corpora of adult language (English, German and Chinese) as well as applying it to the Eve corpus from the CHILDES database in order to see whether analogy formation constitutes a plausible model for human acquisition of language. He was able to show that what we demonstrated for the sentences above also works for a larger corpus of naturally occurring language: Although there were no examples for movement of an auxiliary across a complex NP in the Eve corpus, it is possible to learn by analogy that the auxiliary from a complex NP cannot be fronted.

It is therefore possible to learn syntactic structures from a corpus without any prior knowledge about parts of speech or abstract properties of language. The only assumption that Bod makes is that there are (binary-branching) structures. The assumption of binarity is not even necessary. But if one includes flat branching structures into the calculation, the set of trees will become considerably bigger. Rens Bod only used binary-branching structures in his experiments. In his trees, X consists of two other X's or a word. We are therefore dealing with recursive structures. Therefore, Bod's work proposes a theory of the acquisition of syntactic structures that only requires recursion, something that is viewed by Hauser, Chomsky & Fitch (2002) as a basic property of language.

As shown in section 13.1.8, there is evidence that recursion is not restricted to language and thus one can conclude that it is not necessary to assume innate linguistic knowledge in order to be able to learn syntactic structures from the given input.

Nevertheless, it is important to point out something here: What Rens Bod shows is that syntactic structures can be learned. The information about the parts of speech of each word involved which are not yet included in his structures can also be derived using statistical methods.[38] In all probability, the structures that can be learned correspond to structures that surface-oriented linguistic theories would also assume. However, not all aspects of the linguistic analysis are acquired. In Bod's model, only occurrences of words in structures are evaluated. Nothing is said about whether words stand in a particular regular relationship to one another or not (for example, a lexical rule connecting a passive participle and perfect participle). Furthermore, nothing is said about how the meaning of expressions arise (are they rather holistic in the sense of Construction Grammar or projected from the lexicon?). These are questions that still concern theoretical linguists (see Chapter 23) and cannot straightforwardly be derived from the statistic distribution of words and the structures calculated from them (see Section 23.8.1 for more on this point).

A second comment is also needed: We have seen that statistical information can be used to derive the structure of complex linguistic expressions. This now begs the question of how this relates to Chomsky's earlier argumentation against statistical approaches (Chomsky 1957: 16). Abney (1996: Section 4.2) discusses this in detail. The problem with his earlier argumentation is that Chomsky referred to Markov models. These are statistical versions of finite automatons. Finite automatons can only describe type 3 languages and are therefore not appropriate for analyzing natural language. However, Chomsky's criticism cannot be applied to statistical methods in general.

13.8.4 Negative evidence

In many works that assume innate linguistic knowledge, it is claimed that children do not have access to negative evidence, that is, nobody tells them that sentences such as

[38] Computational linguistic algorithms for determining parts of speech often look at an entire corpus. But children are always dealing with just a particular part of it. The corresponding learning process must then also include a curve of forgetting. See Braine (1987: 67).

(44e) – repeated here as (68) – are ungrammatical (Brown & Hanlon 1970: 42–52; Marcus 1993).

(68) * Is the dog that in the corner is hungry?

The fact that adults do not wake up their children with the ungrammatical sentence of the day is indeed correct, however, children do in fact have access to negative evidence of various sorts. For example, Chouinard & Clark (2003) have shown that English and French speaking parents correct the utterances of their children that are not well-formed. For example, they repeat utterances where the verb was inflected incorrectly. Children can deduce from the fact that the utterance was repeated and from what was changed in the repetition that they made a mistake and Chouinard and Clark also showed that they actually do this. The authors looked at data from five children whose parents all had an academic qualification. They discuss the parent-child relationship in other cultures, too (see Ochs (1982); Ochs & Schieffelin (1984) and Marcus (1993: 71) for an overview) and refer to studies of America families with lower socio-economic status (page 660).

A further form of negative evidence is indirect negative evidence, which Chomsky (1981a: 9) also assumes could play a role in acquisition.

Goldberg (1995: Section 5.2) gives the utterance in (69a) as an example:[39]

(69) a. Look! The magician made the bird disappear.
b. * The magician disappeared the bird.

The child can conclude from the fact that adults use a more involved causative construction with *make* that the verb *disappear*, unlike other verbs such as *melt*, cannot be used transitively. An immediately instructive example for the role played by indirect negative evidence comes from morphology. There are certain productive rules that can however still not be applied if there is a word that blocks the application of the rule. An example is the *-er* nominalization suffix in German. By adding an *-er* to a verb stem, one can derive a noun that refers to someone who carries out a particular action (often habitually) (*Raucher* 'smoker', *Maler* 'painter', *Sänger* 'singer', *Tänzer* 'dancer'). However, *Stehler* 'stealer' is very unusual. The formation of *Stehler* is blocked by the existence of *Dieb* 'thief'. Language learners there have to infer from the non-existence of *Stehler* that the nominalization rule does not apply to *stehlen* 'to steal'.

Similarly, a speaker with a grammar of English that does not have any restrictions on the position of manner adverbs would expect that both orders in (70) are possible (Scholz & Pullum 2002: 206):

(70) a. call the police immediately
b. * call immediately the police

Learners can conclude indirectly from the fact that verb phrases such as (70b) (almost) never occur in the input that these are probably not part of the language. This can be modeled using the relevant statistical learning processes.

[39] Also, see Tomasello (2006a).

The examples provided here for the existence of negative evidence are more plausibility arguments. Stefanowitsch (2008) has combined corpus linguistic studies on the statistical distribution with acceptability experiments and has shown that negative evidence gained from expected frequencies correlates with acceptability judgments of speakers. This process will be discussed now briefly: Stefanowitsch assumes the following principle:

(71) Form expectations about the frequency of co-occurrence of linguistic features or elements on the basis of their individual frequency of occurrence and check these expectations against the actual frequency of co-occurrence. (Stefanowitsch 2008: 518)

Stefanowitsch works with part of the *International Corpus of English* that contains British English (ICE-GB). In this corpus, the verb *say* occurs 3,333 times and sentences with ditransitive verbs (Subj Verb Obj Obj) occur 1,824 times. The entire total of verbs in the corpus is 136,551. If all verbs occurred in all kinds of sentences with the same frequencies, then we would expect *say* to occur 44,52 times (1.824 × 3.333 / 136.551) in the ditransitive construction. But the number of actual occurrences is actually 0 since, unlike (72b), sentences such as (72a) are not used by speakers of English.

(72) a. * Dad said Sue something nice.
b. Dad said something nice to Sue.

Stefanowitsch shows that the non-occurrence of *say* in the ditransitive sentence pattern is significant. Furthermore, he investigated how acceptability judgments compare to the frequent occurrence or non-occurrence of verbs in certain constructions. In a first experiment, he was able to show that the frequent non-occurrence of elements in particular constructions correlates with the acceptability judgments of speakers, whereas this is not the case for the frequent occurrence of a verb in a construction.

In sum, we can say that indirect negative evidence can be derived from linguistic input and that it seems to play an important role in language acquisition.

13.9 Summary

It follows from all this that none of the arguments in favor of innate linguistic knowledge remains uncontroversial. This of course does not rule out that there is still innate linguistic knowledge, however, those who wish to incorporate this assumption into their theories have to take more care than was previously the case to prove that what they assume to be innate is actually part of our linguistic knowledge and that it cannot be learned from the linguistic input alone.

Comprehension questions

1. Which arguments are there for the assumption of innate linguistic knowledge?

Further reading

Pinker's book (1994) is the best written book arguing for nativist models of language.

Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett (1996) discuss all the arguments that have been proposed in favor of innate linguistic knowledge and show that the relevant phenomena can be explained differently. The authors adopt a connectionist view. They work with neuronal networks, which are assumed to model what is happening in our brains relatively accurately. The book also contains chapters about the basics of genetics and the structure of the brain, going into detail about why a direct encoding of linguistic knowledge in our genome is implausible.

Certain approaches using neuronal networks have been criticized because they cannot capture certain aspects of human abilities such as recursion of the multiple usage of words in an utterance. Pulvermüller (2010) discusses an architecture that has memory and uses this to analyze recursive structures. In the overview article, certain works are cited that show that the existence of more abstract rules or schemata of the kind theoretical linguists take for granted can be demonstrated on the neuronal level. Pulvermüller does not, however, assume that linguistic knowledge is innate (p. 173).

Pullum and Scholz have dealt with the Poverty-of-the-Stimulus argument in detail (Pullum & Scholz 2002; Scholz & Pullum 2002).

Goldberg (2006) and Tomasello (2003) are the most prominent proponents of Construction Grammar, a theory that explicitly tries to do without the assumption of innate linguistic knowledge.

14 Generative-enumerative vs. model-theoretic approaches

Generative-enumerative approaches assume that a grammar generates a set of sequences of symbols (strings of words). This is where the term Generative Grammar comes from. Thus, it is possible to use the grammar on page 55, repeated here as (1), to derive the string *er das Buch dem Mann gibt* (‘he the book the man gives’).

(1)

NP → D, N	NP → er	N → Buch
S → NP, NP, NP, V	D → das	N → Mann
	D → dem	V → gibt

Starting with the first symbol (S), symbols are replaced until one reaches a sequence of symbols only containing words. The set of all strings derived in this way is the language described by the grammar.

The following are classed as generative-enumerative approaches:

- all phrase structure grammars
- Transformational Grammars in almost all variants
- GPSG in the formalism of Gazdar, Klein, Pullum & Sag (1985)
- many variants of Categorial Grammar
- many variants of TAG
- Chomsky’s Minimalist Grammars

LFG was also originally designed to be a generative grammar.

The opposite of such theories of grammar are model-theoretic or constraint-based approaches (MTA). MTAs formulate well-formedness conditions on the expressions that the grammar describes. In Section 6.7, we already discussed a model-theoretic approach for theories that use feature structures to model phenomena. To illustrate this point, I will discuss another HPSG example: (2) shows the lexical entry for *kennst* ‘know’.

(2) Lexical entry for *kennst*:

$$\begin{bmatrix} \textsc{phon} & \langle\ \textit{kennst}\ \rangle \\ \textsc{synsem} & \begin{bmatrix} \textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textsc{dsl} & \textit{none} \\ \textit{verb} & \end{bmatrix} \\ \textsc{subcat} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}[\textsc{second},\textsc{sg}]},\ \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{ind} & \boxed{3} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textsc{event} & \boxed{3} \\ \textsc{experiencer} & \boxed{1} \\ \textsc{theme} & \boxed{2} \\ \textit{kennen} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{nonloc} & \ldots \end{bmatrix} \end{bmatrix}$$

In the description of (2), it is ensured that the PHON value of the relevant linguistic sign is ⟨ *kennst* ⟩, that is, this value of PHON is constrained. There are parallel restrictions for the features given in (2): The SYNSEM value is given. In SYNSEM, there are restrictions on the LOC and NONLOC value. In CAT, there are individual restrictions for HEAD and SUBCAT. The value of SUBCAT is a list with descriptions of dependent elements. The descriptions are given as abbreviations here, which actually stand for complex feature descriptions that also consist of feature-value pairs. For the first argument of *kennst*, a HEAD value of type *noun* is required and the PER value in the semantic index has to be *second* and the NUM value has to be *sg*. The structure sharings in (2) form a special kind of constraint. Values that are not specified in the descriptions of lexical entries can vary in accordance with the feature geometry given by the type system. In (2), neither the SLASH value of the nominative NP nor the accusative NP is fixed. This means that SLASH can either be an empty or non-empty list.

The constraints in lexical items such as (2) interact with further constraints that hold for the signs of type *phrase*. For instance, in head-argument structures, the non-head daughter must correspond to an element from the SUBCAT list of the head daughter.

Generative-enumerative and model-theoretic approaches view the same problem from different sides: the generative side only allows what can be generated by given rules, whereas the model-theoretic approach allows everything that is not ruled out by constraints.[1]

Pullum & Scholz (2001: 19–20) and Pullum (2007) list the following model-theoretic approaches:[2]

[1] Compare this to an old joke: In dictatorships, everything that is not allowed is banned, in democracies, everything that is not banned is allowed and in France, everything that is banned is allowed. Generative-enumerative approaches correspond to the dictatorships, model-theoretic approaches are the democracies and France is something that has no correlate in linguistics.

[2] See Pullum (2007) for a historical overview of Model Theoretic Syntax (MTS) and for further references.

- Non-procedural variants of Transformational Grammar of Lakoff, that formulates constraints on potential tree sequences,
- Johnson and Postal's formalism of Relational Grammar (1980)
- GPSG in the variants developed by Gazdar et al. (1988), Blackburn et al. (1993) and Rogers (1997),
- LFG in the formalism of Kaplan (1995)[3] and
- HPSG in the formalism of King (1999).

Categorial Grammars (Bouma & van Noord 1994), TAG (Rogers 1994) and Minimalist approaches (Veenstra 1998) can be formulated in model-theoretic terms.

Pullum & Scholz (2001) point out various differences between these points of view. In the following sections, I will focus on two of these differences.[4] Section 14.3 deals with ten Hacken's objection to the model-theoretic view.

14.1 Graded acceptability

Generative-enumerative approaches differ from model-theoretic approaches in how they deal with the varying degrees of acceptability of utterances. In generative-enumerative approaches, a particular string is either included in the set of well-formed expression or it is not. This means that it is not straightforwardly possible to say something about the degree of deviance: The first sentence in (3) is judged grammatical and the following three are equally ungrammatical.

(3) a. Du kennst diesen Aufsatz.
you know.2SG this.ACC essay

b. *Du kennen diesen Aufsatz.
you know.3PL this.ACC essay

c. *Du kennen dieser Aufsatz.
you know.3PL this.NOM essay

d. *Du kennen Aufsatz dieser.
you know.3PL essay this.NOM

At this point, critics of this view raise the objection that it is in fact possible to determine degrees of acceptability in (3b–d): In (3b), there is no agreement between the subject and the verb, in (3c), *dieser Aufsatz* 'this essay' has the wrong case in addition, and in (3d), *Aufsatz* 'essay' and *dieser* 'this' occur in the wrong order. Furthermore, the sentences in (4) violate grammatical rules of German, but are nevertheless still interpretable.

[3] According to Pullum (2013: Section 3.2), there seems to be a problem for model-theoretic formalizations of so-called *constraining equations*.

[4] The reader should take note here: There are differing views with regard to how generative-enumerative and MTS models are best formalized and not all of the assumptions discussed here are compatible with every formalism. The following sections mirror the important points in the general discussion.

(4) Studenten stürmen mit Flugblättern und Megafon die Mensa und rufen
students storm with flyers and megaphone the canteen and call
alle auf zur Vollversammlung in der Glashalle *zum kommen. Vielen* bleibt das
all up to plenary.meeting in the glass.hall to.the come many stays the
Essen im Mund stecken und *kommen sofort mit.*[5]
food in.the mouth stick and come immediately with
'Students stormed into the university canteen with flyers and a megaphone calling for everyone to come to a plenary meeting in the glass hall. For many, the food stuck in their throats and immediately joined them.'

Chomsky (1975: Chapter 5; 1964) tried to use a string distance function to determine the relative acceptability of utterances. This function compares the string of an ungrammatical expression with that of a grammatical expression and assigns an ungrammaticality score of 1, 2 or 3 according to certain criteria. This treatment is not adequate, however, as there are much more fine-grained differences in acceptability and the string distance function also makes incorrect predictions. For examples of this and technical problems with calculating the function, see Pullum & Scholz (2001: 29).

In model-theoretic approaches, grammar is understood as a system of well-formedness conditions. An expression becomes worse, the more well-formedness conditions it violates (Pullum & Scholz 2001: 26–27). In (3b), the person and number requirements of the lexical item for the verb *kennst* are violated. In addition, the case requirements for the object have not been fulfilled in (3c). There is a further violation of a linearization rule for the noun phrase in (3d).

Well-formedness conditions can be weighted in such a way as to explain why certain violations lead to more severe deviations than others. Furthermore, performance factors also play a role when judging sentences (for more on the distinction between performance and competence, see Chapter 15). As we will see in Section 15, constraint-based approaches work very well as performance-compatible grammatical models. If we combine the relevant grammatical theory with performance models, we will arrive at explanations for graded acceptability differences owing to performance factors.

14.2 Utterance fragments

Pullum & Scholz (2001: Section 3.2) point out that generative-enumerative theories do not assign structure to fragments. For instance, neither the string *and of the* nor the string *the of and* would receive a structure since none of these sequences is an well-formed as an utterance and they are therefore not elements of the set of sequences generated by the grammar. However, *and of the* can occur as part of the coordination of PPs in sentences such as (5) and would therefore have some structure in these cases, for example the one given in Figure 14.1 on the facing page.

(5) That cat is afraid of the dog and of the parrot.

[5] Streikzeitung der Universität Bremen, 04.12.2003, p. 2. The emphasis is mine.

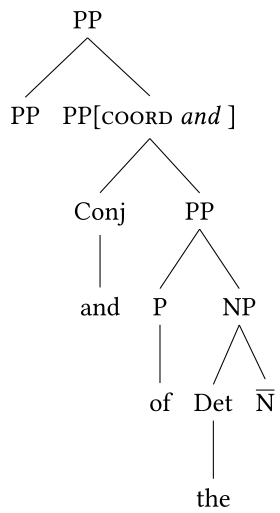


Figure 14.1: Structure of the fragment *and of the* following Pullum & Scholz (2001: 32)

As a result of the interaction of various constraints in a constraint-based grammar, it emerges that *the* is part of an NP and this NP is an argument of *of* and furthermore *and* is combined with the relevant *of*-PP. In symmetric coordination, the first conjunct has the same syntactic properties as the second, which is why the partial structure of *and of the* allows one to draw conclusions about the category of the conjunct despite this not being part of the string.

Ewan Klein noted that Categorial Grammar and Minimalist Grammars, which build up more complex expressions from simpler ones, can sometimes create these kind of fragments (Pullum 2013). For Categorial Grammars with compositional rules that allow one to combine any sequence of words to form a constituent, this is certainly the case. If one views derivations as logical proofs, as is common in some variants of Categorial Grammar, then the actual derivation is irrelevant. What matters is whether a proof can be found or not. However, if one is interested in the derivational structures, then the argument brought forward by Pullum and Scholz is still valid. For some variants of Categorial Grammar that motivate the combination of constituents based on their prosodic and information-structural properties (Steedman 1991: Section 3), the problem persists since fragments have a structure independent from the structure of the entire utterance and independent from their information-structural properties within this complete structure. This structure can be such that it is not possible to analyze it with type-raising rules and compositional rules.

In any case, this argument holds for Minimalist theories since it is not possible to have a combination of *the* with a nominal constituent if this constituent was not already built up from lexical material by Merge.

14.3 A problem for model-theoretic approaches?

Ten Hacken (2007: 237–238) discusses the formal assumptions of HPSG. In HPSG, feature descriptions are used to describe feature structures. Feature structures must contain all the features belonging to a structure of a certain type. Additionally, they have to have a maximally-specific value (see Section 6.7). Ten Hacken discusses gender properties of the English noun *cousin*. In English, gender is important in order to ensure the correct binding of pronouns (see page 274 for German):

(6) a. The man sleeps. He snores.
 b. The woman sleeps. He snores.

While *he* in (6a) can refer to *man*, this is not possible for *woman*. Ten Hacken's problem is that *cousin* is not marked with respect to gender. Thus, it is possible to use it to refer to both male and female relatives. As was explained in the discussion of the case value of *Frau* 'woman' in Section 6.7, it is possible for a value in a description to remain unspecified. Thus, in the relevant feature structures, any value can be assumed. The case of *Frau* can therefore be nominative, genitive, dative or accusative in an actual feature structure. Similarly, there are two possible genders for *cousin* corresponding to usages in (7).

(7) a. I have a cousin. He is very smart.
 b. I have a cousin. She is very smart.

Ten Hacken refers to examples such as (8) and claims that these are problematic:

(8) a. Niels has two cousins.
 b. How many cousins does Niels have?

In plural usage, it is not possible to assume that *cousins* is feminine or masculine since the set of relatives can contain either women or men. It is interesting to note that (9a) is possible in English, whereas German is forced to use (9b) to express the same meaning.

(9) a. Niels and Odette are cousins.
 b. Niels und Odette sind Cousin und Cousine.
 Niels and Odette are cousin.M and cousin.F

Ten Hacken concludes that the gender value has to remain unspecified and this shows, in his opinion, that model-theoretic analyses are unsuitable to describing language.

If we consider what exactly ten Hacken noticed, then it becomes apparent how one can account for that in a model-theoretic approach: Ten Hacken claims that it does not make sense to specify a gender value for the plural form of *cousin*. This can be captured under a model-theoretic approach in two ways. One can either assume that there are no gender features for referential indices in the plural, or one can add a gender value that plural nouns can have.

The first approach is supported by the fact that there are no inflectional differences between pronouns with regard to gender. There is therefore no reason to distinguish plural genders.

(10) a. Niels and Odette are cousins. They are very smart.
b. The cousins/brothers/sisters are standing over there. They are very smart.

No distinctions are found in plural when it comes to nominal inflection (*brothers*, *sisters*, *books*). In German, this is different. There are differences with both nominal inflection and the reference of (some) noun phrases with regard to the sexus of the referent. Examples of this are the previously mentioned examples *Cousin* 'male cousin' and *Cousine* 'female cousin' as well as forms with the suffix *-in* as in *Kindergärtnerin* 'female nursery teacher'. However, gender is normally a grammatical notion that has nothing to do with sexus. An example is the neuter noun *Mitglied* 'member'.

The question that one has to ask when discussing Ten Hacken is the following: Does gender play a role for pronominal binding in German? If this is not the case, then this feature is only relevant within the morphology component, and here the gender value is determined for each noun in the lexicon for both singular and plural. For the binding of personal pronouns, there is no gender difference in German.

(11) Die Schwestern / Brüder / Vereinsmitglieder / Geschwister stehen dort.
the sisters.F brothers.M club.members.N siblings stand there.
Sie lächeln.
They smile.
'The sisters/brothers/club members/siblings are standing there. They are smiling.'

Nevertheless, there are adverbials in German that agree in gender with the noun to which they refer (Höhle 1983: Chapter 6):

(12) a. Die Fenster wurden eins nach dem anderen geschlossen.
the windows.N were one.N after the other closed
'The windows were closed one after the other.'
b. Die Türen wurden eine nach der anderen geschlossen.
the doors.F were one.F after the other closed
'The doors were closed one after the other.'
c. Die Riegel wurden einer nach dem anderen zugeschoben.
the bolts.M were one.M after the other closed
'The bolts were closed one after the other.'

For animate nouns, it possible to diverge from the gender of the noun in question and use a form of the adverbial that corresponds to biological gender:

(13) a. Die Mitglieder des Politbüros wurden eines / einer nach dem anderen aus dem Saal getragen.
the members.N of.the politburo were one.N one.M after the other out.of the hall carried
'The members of the politburo were carried out of the hall one after another.'

b. Die Mitglieder des Frauentanzklubs verließen eines / eine nach
the members.N of.the women's.dance.club left one.N one.F after
dem / der anderen im Schutze der Dunkelheit den Keller.
the.N the.F other in.the protection of.the dark the basement
'The members of the women's dance club left the basement one after another under cover of darkness.'

This deviation from gender in favor of sexus can also be seen with binding of personal and relative pronouns with nouns such as *Weib* 'woman' (pej.) and *Mädchen* 'girl':

(14) a. „Farbe bringt die meiste Knete!" verriet ein 14jähriges türkisches
color brings the most money tells a 14-year.old Turkish
Mädchen, *die* die Mauerstückchen am Nachmittag am Checkpoint
girl.N who.F the wall.pieces in.the afternoon at Checkpoint
Charlie an Japaner und US-Bürger verkauft.[6]
Charlie on Japanese and US-citizens sells
' "Color gets the most money" said a 14-year old Turkish girl who sells pieces of the wall to Japanese and American citizens at Checkpoint Charlie.'

b. Es ist ein junges *Mädchen*, *die* auf der Suche nach CDs bei Bolzes
it is a young girl.N who.F on the search for CDs at Bolzes
reinschaut.[7]
stops.by
'It is a young girl looking for CDs that stops by Bolzes.'

For examples from Goethe, Kafka and Thomas Mann, see Müller (1999a: 417–418).

For inanimate nouns such as those in (12), agreement is obligatory. For the analysis of German, one therefore does in fact require a gender feature in the plural. In English, this is not the case since there are no parallel examples with pronouns inflecting for gender. One can therefore either assume that plural indices do not have a gender feature or that the gender value is *none*. The feature would then have a value and fulfill the formal requirements.

In general, it is clear that cases such as the one constructed by ten Hacken will never be a problem since there are either values that make sense, or there are contexts for which there is no value that makes sense and one therefore does not require the features.

I have pointed out a technical problem in connection with the model-theoretic view (Müller 1999a: Section 14.4). This problem is only technical in nature, however. I showed that spurious ambiguities arise for a particular analysis of verbal complexes in German when one resolves the values of a binary feature (FLIP). This can be avoided by the complicated stipulation of a value in certain contexts.

[6] taz, 14.06.1990, p. 6.
[7] taz, 13.03.1996, p. 11.

15 The competence/performance distinction

In Chapter 12.6.3, we discussed the distinction between competence and performance (Chomsky 1965: Chapter I.1) assumed by several theories of grammar. Theories of competence are intended to explain linguistic knowledge and performance theories are assigned the task of explaining how linguistic knowledge is used as well as why mistakes are made by speech production and comprehension. A classic example in the competence/performance discussion are cases of self-embedding. Chomsky & Miller (1963: 286) discuss the following example with recursively embedded relative clauses:

(1) (the rat (the cat (the dog chased) killed) ate the malt)

(2b) is a corresponding example in German:

(2) a. dass der Hund bellt, der die Katze jagt, die die Maus gefangen hat
that the dog.M barks that.M the cat chases that.F the mouse caught has
'that the dog that chases the cat that caught the mouse is barking.'

b. Der Hund, [$_1$ der die Katze, [$_2$ die die Maus gefangen hat, $_2$] jagt $_1$] bellt.
the dog that the cat that the mouse caught has chases barks

The examples in (1) and (2b) are entirely incomprehensible for most people. If one rearranges the material somewhat, it is possible to process the sentences and assign a meaning to them.[1] For sentences such as (2b), it is often assumed that they fall within our grammatical competence, that is, we possess the knowledge required to assign a structure to the sentence, however the processing of utterances such as (2b) exceeds language-

[1] The sentence in (2a) can be continued following the pattern that was used to create the sentence. For instance by adding *die unter der Treppe lebte, die meine Freunde repariert haben* 'who lived under the staircase which my friends repaired'. This shows that a restriction of the number of elements that depend on one head to seven (Leiss 2003: 322) does not restrict the set of the sentences that are generated or licensed by a grammar to be finite. There are at most two dependents of each head in (2a). The extraposition of the relative clauses allows the hearer to group material into processable and reducable chunks, which reduces the cognitive burden during processing.
This means that the restriction to seven dependents does not cause a finitization of recursion („Verendlichung von Rekursivität") as was claimed by Leiss (2003: 322). Leiss argued that Miller could not use his insights regarding short term memory, since he worked within Transformational Grammar rather than in Dependency Grammar. The discussion shows that dependency plays an important role, but that linear order is also important for processing.

independent abilities of our brain. In order to successfully process (2b), we would have to retain the first five noun phrases and corresponding hypotheses about the further progression of the sentence in our heads and can only begin to combine syntactic material when the verbs appear. Our brains become overwhelmed by this task. These problems do not arise when analyzing (2a) as it is possible to immediately begin to integrate the noun phrase into a larger unit.

Nevertheless, self embedding of relative clauses can also be constructed in such a way that our brains can handle them. Hans Uszkoreit gives the following example:

(3) Die Bänke, [$_1$ auf denen damals die Alten des Dorfes, [$_2$ die allen
the benches on which back.then the old.people of.the village that all
Kindern, [$_3$ die vorbeikamen $_3$], freundliche Blicke zuwarfen $_2$], lange
children that came.by friendly glances gave long
Stunden schweigend nebeneinander saßen $_1$], mussten im letzten Jahr
hours silent next.to.each.other sat must in last year
einem Parkplatz weichen.
a car.park give.way.to
'The benches on which the older residents of the village, who used to give friendly glances to all the children who came by, used to sit silently next to one another had to give way to a car park.'

Therefore, one does not wish to include in the description of our grammatical knowledge that relative clauses are not allowed to be included inside each other as in (2b) as this would also rule out (3).

We can easily accept the fact that our brains are not able to process structures past a certain degree of complexity and also that corresponding utterances then become unacceptable. The contrast in the following examples is far more fascinating:[2]

(4) a. # The patient [who the nurse [who the clinic had hired] admitted] met Jack.
b. * The patient who the nurse who the clinic had hired met Jack.

Although (4a) is syntactically well-formed and (4b) is not, Gibson & Thomas (1999) were able to show that (4b) is rated better by speakers than (4a). It does not occur to some people that an entire VP is missing. There are a number of explanations for this fact that all in some way make the claim that previously heard words are forgotten as soon as new words are heard and a particular degree of complexity is exceeded (Frazier 1985: 178; Gibson & Thomas 1999).

Instead of developing grammatical theories that treat (2b) and (4a) as unacceptable and (3) and (4b) as acceptable, descriptions are developed that equally allow (2b), (3), and (4a) (competence models) and then additionally investigate the way utterances are processed in order to find out what kinds of structures our brains can handle and what

[2] See Gibson & Thomas (1999: 227). Frazier (1985: 178) attributes the discovery of this kind of sentences to Janet Fodor.

kind of structures it cannot. The result of this research is then a performance model (see Gibson (1998), for example). This does not rule out that there are language-specific differences affecting language processing. For example, Vasishth, Suckow, Lewis & Kern (2010) have shown that the effects that arise in self-embedding structures in German are different from those that arise in the corresponding English cases such as (4): Due to the frequent occurrence of verb-final structures in German, speakers of German were able to better store predictions about the anticipated verbs into their working memory (p. 558).

Theories in the framework of Categorial Grammar, GB, LFG, GPSG and HPSG are theories about our linguistic competence.[3] If we want to develop a grammatical theory that directly reflects our cognitive abilities, then there should also be a corresponding performance model to go with a certain competence model. In the following two sections, I will recount some arguments from Sag & Wasow (2011) in favor of constraint-based theories such as GPSG, LFG and HPSG.

15.1 The derivational theory of complexity

The first point discussed by Sag & Wasow (2011) is the Derivational Theory of Complexity. In the early days of Transformational Grammar, it was assumed that transformations were cognitively real, that is, it is possible to measure the consumption of resources that transformations have. A sentence that requires more transformations than the analysis of another sentence should therefore also be more difficult for humans to process. The corresponding theory was dubbed the *Derivational Theory of Complexity* (DTC) and initial experiments seemed to confirm it (Miller & McKean 1964; Savin & Perchonock 1965; Clifton & Odom 1966) such that in 1968 Chomsky still assumed that the Derivational Theory of Complexity was in fact correct (Chomsky 1976a: 249–250).[4] Some years later however, most psycholinguists rejected the DTC. For discussion of several experiments

[3] For an approach where the parser is equated with UG, see Abney & Cole (1986: Section 3.4). For a performance-oriented variant of Minimalism, see Phillips (2003).
In Construction Grammar, the question of whether a distinction between competence and performance would be justified at all, is controversially discussed. (See Section 10.6.4.5.1.) Fanselow, Schlesewsky, Cavar & Kliegl (1999) also suggest a model – albeit for different reasons – where grammatical properties considerably affect processing properties. The aforementioned authors work in the framework of Optimality Theory and show that the OT constraints that they assume can explain parsing preferences. OT is not a grammatical theory on its own but rather a meta theory. It is assumed that there is a component GEN that creates a set of candidates. A further component EVAL then chooses the most optimal candidate from this set of candidates. GEN contains a generative grammar of the kind that we have seen in this book. Normally, an GP/MP variant or also LFG is assumed as the base grammar. If one assumes a transformational theory, then one automatically has a problem with the DTC that we will encounter in the following section. If one wishes to develop OT parsing models, then one has to make reference to representational variants of GB as the aforementioned authors seem to.

[4] In the Transformational Grammar literature, transformations were later viewed as a metaphor (Lohnstein 2014: 170, also in Chomsky 2001: Footnote 4), that is, it was no longer assumed to have a psycholinguistic reality. In *Derivation by Phase* and *On Phases*, Chomsky refers once again to processing aspects such as computational and memory load (Chomsky: 2001: 11, 12, 15; 2007: 3, 12; 2008: 138, 145, 146, 155). Also, see Marantz (2005: 440) and Richards (2015b).

that testify against the DTC, see Fodor, Bever & Garrett (1974: 320–328). One set of phenomena where the DTC makes incorrect predictions for respective analyses is that of elliptical constructions, for example (Fodor, Bever & Garrett 1974: 324): In elliptical constructions, particular parts of the utterance are left out or replaced by auxiliaries. In transformation-based approaches, it was assumed that (5b) is derived from (5a) by means of deletion of *swims* and (5c) is derived from (5b) by inserting *do*.

(5) a. John swims faster than Bob swims.
 b. John swims faster than Bob.
 c. John swims faster than Bob does.

The DTC predicts that (5b) should require more time to process than (5a), since the analysis of (5b) first requires to build up the structure in (5a) and then delete *swims*. This prediction was not confirmed.

Similarly, no difference could be identified for the pairs in (6) and (7) even though one of the sentences, given the relevant theoretical assumptions, requires more transformations for the derivation from a base structure (Fodor, Bever & Garrett 1974: 324).

(6) a. John phoned up the girl.
 b. John phoned the girl up.

(7) a. The bus driver was nervous after the wreck.
 b. The bus driver was fired after the wreck.

In (6), we are dealing with local reordering of the particle and the object. (7b) contains a passive clause that should be derived from an active clause under Transformational Grammar assumptions. If we compare this sentence with an equally long sentence with an adjective, like (7a), the passive clause should be more difficult to process. This is, however, not the case.

It is necessary to add two qualifications to Sag and Wasow's claims: If one has experimental data that show that the DTC makes incorrect predictions for a particular analysis, this does not necessarily mean that the DTC has been disproved. One could also try to find a different analysis for the phenomenon in question. For example, instead of a transformation that deletes material, one could assume empty elements for the analysis of elliptical structures that are inserted directly into the structure without deleting any material (see page 70). Data such as (5) would then be irrelevant to the discussion.[5] However, reordering such as (6b) and the passive in (7b) are the kinds of phenomena that are typically explained using transformations.

A structure building operation that begins with words and is followed by transformations, as recently assumed by theories in the Minimalist Program, is psycholinguistically implausible for sentence parsing. See Labelle (2007) and Section 15.2 for more on incremental processing.
Chomsky (2007: 6) (written later than *On Phases*) seems to adopt a constraint-based view. He writes that "a Merge-based system involves parallel operations" and compares the analysis of an utterance with a proof and explicitly mentions the competence/performance distinction.

[5] Culicover & Jackendoff (2005: Chapters 1 and 7) argue in favor of analyzing ellipsis as a semantic or pragmatic phenomenon rather than a syntactic one anyway.

The second qualification pertains to analyses for which there is a representational variant: It is often said that transformations are simply metaphors (Jackendoff: 2000: 22–23; 2007: 5, 20): For example, we have seen that extractions with a transformational grammar yield structures that are similar to those assumed in HPSG. Figure 15.1 shows cyclic movement in GB theory compared to the corresponding HPSG analysis.

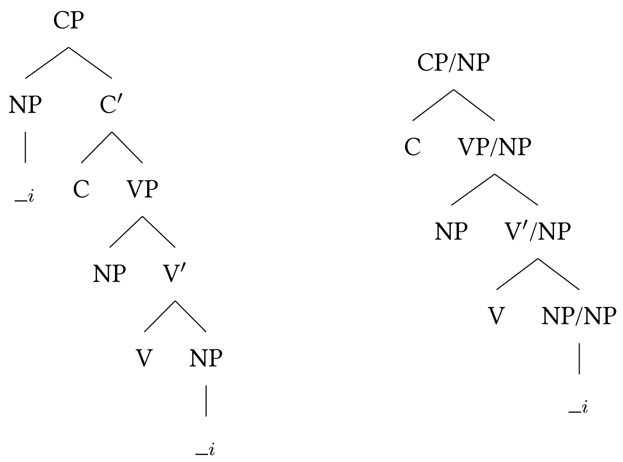


Figure 15.1: Cyclic movement vs. feature percolation

In GB, an element is moved to the specifier positions of CP (SpecCP) and can then be moved from there to the next higher SpecCP position.

(8) a. Chris$_i$, we think [$_{\text{CP}}$ $\_{i}$ Anna claims [$_{\text{CP}}$ $\_{i}$ that David saw $\_{i}$]]. (GB)
 b. Chris$_i$, we think [$_{\text{CP/NP}}$ Anna claims [$_{\text{CP/NP}}$ that David saw $\_{i}$]]. (HPSG)

In HPSG, the same effect is achieved by structure sharing. Information about a long-distance dependency is not located in the specifier node but rather in the mother node of the projection itself. In Section 21.2, we discuss various ways of eliminating empty elements from grammars. If we apply these techniques to structures such as the GB structure in Figure 15.1, then we arrive at structures where information about missing elements is integrated into the mother node (CP) and the position in SpecCP is unfilled. This fundamentally corresponds to the HPSG structure in Figure 15.1.[6] It follows from this that there are classes of phenomena that can be spoken about in terms of transformations without expecting empirical differences with regard to performance when compared to transformation-less approaches. However, it is important to note that we are dealing with an S-Structure in the left-hand tree in Figure 15.1. As soon as one assumes that this is derived by moving constituents out of other structures, this equivalence of approaches disappears.

[6] In Figure 15.1, additionally the unary branching of C′ to CP was omitted in the tree on the right so that C combines directly with VP/NP to form CP/NP.

15.2 Incremental processing

The next important point mentioned by Sag & Wasow (2011) is the fact that both processing and production of language takes places incrementally. As soon as we hear or read even the beginning of a word, we begin to assign meaning and to create structure. In the same way, we sometimes start talking before we have finished planning the entire utterance. This is shown by interruptions and self correction in spontaneous speech (Clark & Wasow 1998; Clark & Fox Tree 2002). When it comes to processing spoken speech, Tanenhaus et al. (1996) have shown that we access a word as soon as we have heard a part of it (also see Marslen-Wilson 1975). The authors of the study carried out an experiment where participants were instructed to pick up particular objects on a grid and reorganize them. Using eye-tracking measurements, Tanenhaus and colleagues could then show that the participants could identify the object in question earlier if the sound sequence at the beginning of the word was unambiguous than in cases where the initial sounds occurred in multiple words. An example for this is a configuration with a candle and candy: *candy* and *candle* both begin with *can* such that speakers could not yet decide upon hearing this sequence which lexical entry should be accessed. Therefore, there was a slight delay in accessing the lexical entry when compared to words where the objects in question did not contain the same segment at the start of the word (Tanenhaus et al. 1995: 1633).

If complex noun phrases were used in the instructions (*Touch the starred yellow square*), the participants' gaze fell on the object in question 250ms after it was unambiguously identifiable. This means that if there was only a single object with stars on it, then they looked at it after they heard *starred*. In cases where there were starred yellow blocks as well as squares, they looked at the square only after they had processed the word *square* (Tanenhaus et al. 1995: 1632). The planning and execution of a gaze lasts 200ms. From this, one can assume that hearers combine words directly and as soon as enough information is available, they create sufficient structure in order to capture the (potential) meaning of an expression and react accordingly. This finding is incompatible with models that assume that one must have heard a complete noun phrase or even a complete utterance of even more complexity before it is possible to conclude anything about the meaning of a phrase/utterance. In particular, analyses in the Minimalist Program which assume that only entire phrases or so-called phases[7] are interpreted (Chomsky (1999) and Marantz (2005: 441), who explicitly contrast the MP to Categorial Grammar) must therefore be rejected as psycholinguistically inadequate.[8,9]

[7] Usually, only CP and vP are assumed to be phases.

[8] Sternefeld (2006: 729–730) points out that in theories in the Minimalist Program, the common assumption of uninterpretable features is entirely unjustified. Chomsky assumes that there are features that have to be deleted in the course of a derivation since they are only relevant for syntax. If they are not checked, the derivation crashes at the interface to semantics. It follows from this that NPs should not be interpretable under the assumptions of these theories since they contain a number of features that are irrelevant for the semantics and have to therefore be deleted (for an overview, see Richards 2015b). As we have seen, these kinds of theories are incompatible with the facts.

[9] It is sometimes claimed that current Minimalist theories are better suited to explain generation than parsing. But these models are as implausible for generation as they are for parsing. The reason is that it is assumed

With contrastive emphasis of individual adjectives in complex noun phrases (e. g. *the BIG blue triangle*), hearers assumed that there must be a corresponding counterpart to the reference object, e. g. a small blue triangle. The eye-tracking studies carried out by Tanenhaus et al. (1996) have shown that taking this kind of information into account results in objects being identified more quickly.

Similarly, Arnold et al. (2004) have shown, also using eye-tracking studies, that hearers tend to direct their gaze to previously unmentioned objects if the interlocutor interrupts their speech with *um* or *uh*. This can be traced back to the assumption that hearers assume that describing previously unmentioned objects is more complex than referring to objects already under discussion. The speaker can create more time for himself by using *um* or *uh*.

Examples such as those above constitute evidence for approaches that assume that when processing language, information from all available channels is used and that this information is also used as soon as it is available and not only after the structure of the entire utterance or complete word group has been constructed. The results of experimental research therefore show that the hypothesis of a strictly modular organization of linguistic knowledge must be refuted. Proponents of this hypothesis assume that the output of one module constitutes the input of another without a given module having access to the inner states of another module or the processes taking place inside it. For example, the morphology module could provide the input for syntax and then this would be processed later by the semantic module. One kind of evidence for this kind of organization of linguistic knowledge that is often cited are so-called *garden path sentences* such as (9):

(9) a. The horse raced past the barn fell.
    b. The boat floated down the river sank.

These sentences cannot be processed by nearly all speakers of English since their parser is led down a garden path as it builds up a complete structure for (10a) or (10b) only then to realize that there is another verb that cannot be integrated into this structure.

(10) a. The horse raced past the barn.
     b. The boat floated down the river.

However, the actual structure of (9) contains a reduced relative clause (*raced past the barn* or *floated down the river*) that corresponds to the sentences in (11) meaningwise:

(11) a. The horse that was raced past the barn fell.
     b. The boat that was floated down the river sank.

The failure of the parser in these cases was explained by assuming that syntactic processing such as constructing a sentence from NP and VP takes place independently of the

that there is a syntax component that generates structures that are then shipped to the interfaces. This is not what happens in generation though. Usually speakers know what they want to say (at least partly), that is, they start with semantics.

processing of other constraints. As Crain & Steedman (1985) and others have shown, yet there are data that make this explanation seem less plausible: If (9a) is uttered in a relevant context, the parser is not misled. In (12), there are multiple horses under discussion and each NP is clearly identified by a relative clause. The hearer is therefore prepared for a relative clause and can process the reduced relative clause without being led down the garden path, so to speak.

(12) The horse that they raced around the track held up fine. The horse that was raced down the road faltered a bit. And the horse raced past the barn fell.

By exchanging lexical material, it is also possible to modify (9a) in such way as to ensure that processing is unproblematic without having to add additional context. It is necessary to choose the material so that the interpretation of the noun as the subject of verb in the reduced relative clause is ruled out. Accordingly, *evidence* in (13) refers to an inanimate noun. It is therefore not a possible agent of *examined.* An hypothesis with *evidence* as the agent of *examined* is therefore never created when processing this sentence (Sag & Wasow 2011).

(13) The evidence examined by the judge turned out to be unreliable.

Since processing is carried out incrementally, it is sometimes assumed that realistic grammars should be obliged to immediately assign a constituent structure to previously heard material (Ades & Steedman 1982; Hausser 1992). Proponents of this view would assume a structure for the following sentence where every word forms a constituent with the preceding material:

(14) [[[[[[[[[[[[[Das britische] Finanzministerium] stellt] dem] angeschlagenen]
the British treasury provides the downbeat
Bankensystem] des] Landes] mindestens] 200] Milliarden] Pfund] zur]
banking.system of.the country at.least 200 billion pounds to
Verfügung].
use

Pulman (1985), Stabler (1991) and Shieber & Johnson (1993: 301–308) have shown, however, that it is possible to build semantic structures incrementally using the kind of phrase structure grammars we encountered in Chapter 2. This means that a partial semantic representation for the string *das britische* ('the British') can be calculated without having to assume both of the words form a constituent in (14). Therefore, one does not necessarily need a grammar that licenses the immediate combination of words directly. Furthermore, Shieber & Johnson (1993) point out that from a purely technical point of view, synchronic processing is more costly than asynchronic processing since synchronic processing requires additional mechanisms for synchronization whereas asynchronic processing processes information as soon as it becomes available (p. 297–298). Shieber and Johnson do not clarify whether this also applies to synchronic/asynchronic processing of syntactic and semantic information. See Shieber & Johnson (1993) for incremental processing and for a comparison of Steedman's Categorial Grammar and TAG.

What kind of conclusions can we draw from the data we have previously discussed? Are there further data that can help to determine the kinds of properties a theory of grammar should have in order to count as psycholinguistically plausible? Sag, Wasow & Bender (2003) and Sag & Wasow (2011) list the following properties that a performance-compatible competence grammar should have:[10]

- surface-oriented
- model-theoretic and therefore constraint-based
- strictly lexicalist

Approaches such as CG, GPSG, LFG, HPSG, CxG and TAG are surface-oriented since they do not assume a base structure from which other structures are derived via transformations. Transformational approaches, however, require additional assumptions.[11] This will be briefly illustrated in what follows. In Section 3.1.8, we encountered the following analysis of English interrogatives:

(15) $[_{\text{CP}}$ What$_i$ $[_{\text{C}'}$ will$_k$ $[_{\text{IP}}$ Ann $[_{\text{I}'}$ $\_{}_k$ $[_{\text{VP}}$ read $\_{}_i]]]]]$.

This structure is derived from (16a) by two transformations (two applications of move α):

(16) a. Ann will read what?
b. * Will Ann read what.

The first transformation creates the order in (16b) from (16a), and the second creates (15) from (16b).

When a hearer processes the sentence in (15), he begins building structure as soon as he hears the first word. Transformations can, however, only be carried out when the entire utterance has been heard. One can, of course, assume that hearers process surface structures. However, since they begin to access semantic knowledge early into an utterance – as we have seen –, this begs the question of what we need a deep structure for at all.

In analyses such as those of (15), deep structure is superfluous since the relevant information can be reconstructed from the traces. Corresponding variants of GB have

[10] Also, see Jackendoff (2007), who considers a performance model for a constraint-based, surface-oriented linguistic theory.

[11] An exception among transformational approaches is Phillips (2003). Phillips assumes that structures relevant for phenomena such as ellipsis, coordination and fronting are built up incrementally. These constituents are then reordered in later steps by transformations. For example, in the analysis of (i), the string *Wallace saw Gromit in* forms a constituent where *in* is dominated by a node with the label P(P). This node is then turned into a PP in a subsequent step (p. 43–44).

(i) Wallace saw Gromit in the kitchen.

While this approach is a transformation-based approach, the kind of transformation here is very idiosyncratic and incompatible with other variants of the theory. In particular, the modification of constituents contradicts the assumption of Structure Preservation when applying transformations as well as the *No Tampering Condition* of Chomsky (2008). Furthermore, the conditions under which an incomplete string such as *Wallace saw Gromit in* forms a constituent are not entirely clear.

been proposed in the literature (see page 124). They are compatible with the requirement of being surface-oriented. Chomsky (1981a: 181; 1986a: 49) and Lasnik & Saito (1992: 59–60) propose analyses where traces can be deleted. In these analyses, the deep structure can not be directly reconstructed from the surface structure and one requires transformations in order to relate the two. If we assume that transformations are applied 'online' during the analysis of utterances, then this would mean that the hearer would have to keep a structure derived from previously heard material as well as a list of possible transformations during processing in his working memory. In constraint-based grammars, entertaining hypotheses about potential upcoming transformation steps is not necessary since there is only a single surface structure that is processed directly. At present, it is still unclear whether it is actually possible to distinguish between these models empirically. For Minimalist models with a large number of movements (see Figure 4.20 on page 148, for example), it should be clear that they are unrealistic since they require storage space to manage these hypotheses and we know that this is very limited in humans.

Frazier & Clifton (1996: 27) assume that a transformation-based competence grammar yields a grammar with pre-compiled rules/templates that is then used for parsing. Therefore, theorems derived from UG are used for parsing and not axioms of UG directly. Johnson (1989) also suggests a parsing system that applies constraints from different subtheories of GB as early as possible. This means that while he does assume the levels of representation D-Structure, S-Structure, LF and PF, he specifies the relevant constraints ($\overline{\mathrm{X}}$ Theory, Theta-Theory, Case Theory, …) as logical conditions that can be reorganized, evaluated logically adequate and then used in a different order for structure building.[12] Chomsky (2007: 6) also compares human parsing to working through a proof, where each step of the proof can be carried out in different orders. This view does not assume the psychological reality of levels of grammatical representation when processing language, but simply assumes that principles and structures play a role when it comes to language acquisition. As we have seen, the question of whether we need UG to explain language acquisition was not yet decided in favor of UG-based approaches. Instead, all available evidence seems to point in the opposite direction. However, even if innate linguistic knowledge does exist, the question arises as to why one would want to represent this as several structures linked via transformations when it is clear that these do not play a role for humans (especially language learners) when processing language. Approaches that can represent this knowledge by using fewer technical means, e. g. without transformations, therefore are preferable. For more on this point, see Kuhn (2007: 615).

The requirement for constraint-based grammars is supported by incremental processing and also by the ability to deduce what will follow from previously heard material. Stabler (1991) has pointed out that Steedman's argumentation with regard to incrementally processable grammars is incorrect, and instead argues for maintaining a modular view of grammar. However, he himself has developed a constraint-based grammar where syn-

[12] Stabler (1992: Section 15.7) also considers a constraint-based view, but arrives at the conclusion that parsing and other linguistic tasks should use the structural levels of the competence theory. This would again pose problems for the DTC.

tactic and semantic knowledge can be accessed at any time. He formulates both syntactic structures and the semantic representations attached to them as conjoined constraints and then presented a processing system that processes structures based on the availability of parts of syntactic and semantic knowledge. Stabler rejects models of performance that assume that one must first apply all syntactic constraints before the semantic ones can be applied. If one abandons this strict view of modularity, then we arrive at something like (17):

(17) $(\mathrm{Syn}_1 \wedge \mathrm{Syn}_2 \wedge \ldots \wedge \mathrm{Syn}_n) \wedge (\mathrm{Sem}_1 \wedge \mathrm{Sem}_2 \wedge \ldots \wedge \mathrm{Sem}_n)$

Syn_1–Syn_n stand for syntactic rules or constraints and Sem_1–Sem_n stand for semantic rules or constraints. If one so desires, it is possible to refer to the expressions in brackets as modules. Since it is possible to randomly reorder conjoined expressions , one can imagine performance models that first apply some rules from the syntax module and then, when enough information is present, respective rules from the semantic module. The order of processing could therefore be as in (18), for example:

(18) $\mathrm{Syn}_2 \wedge \mathrm{Sem}_1 \wedge \mathrm{Syn}_1 \wedge \ldots \wedge \mathrm{Syn}_n \wedge \mathrm{Sem}_2 \wedge \ldots \wedge \mathrm{Sem}_n$

If one subscribes to this view of modularity, then theories such as HPSG or CxG also have a modular structure. In the representation assumed in the HPSG variant of Pollard & Sag (1987) and Sign-Based CxG (see Section 10.6.2), the value of SYN would correspond to the syntax module, the value of SEM to the semantic module and the value of PHON to the phonology module. If one were to remove the respective other parts of the lexical entries/dominance schemata, then one would get the part of the theory corresponding exactly to the level of representation in question.[13] Jackendoff (2000) argues for this form of modularity with the relevant interfaces between the modules for phonology, syntax, semantics and further modules from other areas of cognition. Exactly what is to be gained from assuming those modules and how these could be proved empirically remains somewhat unclear to me. For skepticism with regard to the very concept of modules, see Jackendoff (2000: 22,27). For more on interfaces and modularization in theories such as LFG and HPSG, see Kuhn (2007).

In conclusion, we can say that surface-oriented, model-theoretic and strongly lexicalist grammatical theories such as CG, LFG, GPSG, HPSG, CxG and the corresponding GB/MP variants can plausibly be combined with processing models, while this is not the case for the overwhelming majority of GB/MP theories.

[13] In current theories in the Minimalist Program, an increasing amount of morphological, syntactic, semantic and information-structural information is being included in analyses (see Section 4.6.1). While there are suggestions for using feature-value pairs (Sauerland & Elbourne 2002: 290–291), a strict structuring of information as in GPSG, LFG, HPSG, CxG and variants of CG and TAG is not present. This means that there are the levels for syntax, Phonological Form and Logical Form, but the information relevant for these levels is an unstructured part of syntax, smeared all over syntactic trees.

16 Language acquisition

Linguists and philosophers are fascinated by the human ability to acquire language. Assuming the relevant input during childhood, language acquisition normally takes place completely effortlessly. Chomsky (1965: 24–25) put forward the requirement that a grammatical theory must provide a plausible model of language acquisition. Only then could it actually explain anything and would otherwise remain descriptive at best. In this section, we will discuss theories of acquisition from a number of theoretical standpoints.

16.1 Principles and Parameters

A very influential explanation of language acquisition is Chomsky's Principles and Parameters model (1981a). Chomsky assumes that there is an innate Universal Grammar that contains knowledge that is equally relevant for all languages. Languages can then vary in particular ways. For every difference between languages in the area of core grammar, there is a feature with a specific value. Normally, the value of a parameter is binary, that is, the value is either '+' or '−'. Depending on the setting of a parameter, languages have certain properties, that is, the setting of the parameter determines whether a language belongs to a particular class of languages. Parameters are assumed to influence multiple properties of a grammar simultaneously (Chomsky 1981a: 6). For example, Rizzi (1986) claims that the pro-drop parameter affects whether referential subjects can be omitted, the absence of expletives, subject extraction from clauses with complementizers (*that*-t contexts) and interrogatives and finally the possibility of realizing the subject postverbally in VO-languages (see Chomsky 1981a: Section 4.3; Meisel 1995: 12). It has been noted that there are counter-examples for all the correlations assumed.[1] Another example of a parameter is the Head Directionality Parameter discussed in Section 13.1.1. As was shown, there are languages where heads govern in different directions. In his overview article, Haider (2001) still mentions the parametrized Subjacency Principle but notes that subjacency is no longer assumed as a principle in newer versions of the theory (see Section 13.1.5.2 for more on subjacency).

Snyder (2001) discovered a correlation of various phenomena with productive root compounding as it is manifested for instance in compounding of two nouns. He argues

[1] See Haider (1994) and Haider (2001: Section 2.2) for an overview. Haider assumes that there is at least a correlation between the existence of expletive subjects and pro-drop. However, Galician is a pro-drop language with expletive subject pronouns (Raposo & Uriagereka 1990: Section 2.5). Franks (1995: 314) cites Upper and Lower Sorbian as pro-drop languages with expletive subjects. Scholz & Pullum (2002: 218) point out that there is an expletive pronoun *ci* in modern Italian although Italian is classed as a pro-drop language.

that the acquisition of complex predicate formation is connected to the acquisition of compound structures and that there is a parameter that is responsible for this type of compounding and simultaneously for the following set of phenomena:

(1)
a. John painted the house red. (resultative)
b. Mary picked the book up/picked up the book. (verb-particle)
c. Fred made Jeff leave. (*make*-causative)
d. Fred saw Jeff leave. (perceptual report)
e. Bob put the book on the table. (*put*-locative)
f. Alice sent the letter to Sue. (*to*-dative)
g. Alice sent Sue the letter. (double-object dative)

Snyder examined languages from various language groups (Afroasiatic, Austroasiatic, Austronesian, Finno-Ugric, Indo-European (Germanic, Romance, Slavic), Japanese-Korean, Niger-Kordofanian (Bantu), and Sino-Tibetan, as well as American Sign Language and the language isolate Basque). The languages that were examined either had all of these phenomena or none. This was tested with native speakers of the respective languages. In addition the claim that these phenomena are acquired once noun-noun compounds are used productively was tested for English using CHILDES data. The result was positive with the exception of the double object construction, for which an explanation was provided. The correlation of the phenomena in (1) is interesting and was interpreted as proof of the existence of a parameter that correlates several phenomena in a language. However, Son (2007) and Son & Svenonius (2008) showed that Snyder's claims on Japanese were wrong and that there are further languages like Korean, Hebrew, Czech, Malayalam, Javanese in which some of the phenomena are not correlated.

Gibson & Wexler (1994) discuss the acquisition of constituent order and assume three parameters that concern the position of the verb relative to the subject (SV vs. VS) and relative to the object(VO vs. OV) as well as the V2-property. There is no consensus in the literature about which parameters determine the make-up of languages (see Newmeyer (2005: Section 3.2) and Haspelmath (2008) for an overview and critical discussion). Fodor (1998a: 346–347) assumes that there are 20 to 30 parameters, Gibson & Wexler (1994: 408) mention the number 40, Baker (2003: 349) talks of 10 to 20 and Roberts & Holmberg (2005: 541) of 50 to 100. There is no consensus in the literature as to which parameters one should assume, how they interact and what they predict. However, it is nevertheless possible to contemplate how a grammar of an individual language could be derived from an UG with parameters that need to be set. Chomsky's original idea (1986b: Section 3.5.1) was that the child sets the value of a parameter based on the language input as soon as the relevant evidence is present from the input Gibson & Wexler (1994); Nowak et al. (2001)). At every point in time, the learner has a grammar with certain parameter settings that correspond to the input seen so far. In order to fully acquire a grammar, all parameters must be assigned a value. In theory, thirty utterances should be enough to acquire a grammar with thirty parameters if these utterances provide unambiguous evidence for a particular parameter value.

This approach has often been criticized. If the setting of a parameter leads to a learner using a different grammar, one would expect sudden changes in linguistic behavior. This is, however, not the case (Bloom 1993: 731). Fodor (1998a: 343–344) also notes the following three problems: 1) Parameters can affect things that are not visible from the perceptible word order. 2) Many sentences are ambiguous with regard to the setting of a particular parameter, that is, there are sometimes multiple combinations of parameters compatible with one utterance. Therefore, the respective utterances cannot be used to set any parameters (Berwick & Niyogi 1996; Fodor 1998b). 3) There is a problem with the interaction of parameters. Normally multiple parameters play a role in an utterance such that it can be difficult to determine which parameter contributes what and thus how the values should be determined.

Points 1) and 2) can be explained using the constituent order parameters of Gibson and Wexler: Imagine a child hears a sentences such as (2):

(2) a. Daddy drinks juice.
    b. Pappa trinkt Saft.
       daddy drinks juice

These sentences look exactly the same, even though radically different structures are assumed for each. According to the theories under discussion, the English sentence has the structure shown in Figure 3.8 on page 102 given in abbreviated form in (3a). The German sentence, on the other hand, has the structure in Figure 3.13 on page 110 corresponding to (3b):

(3) a. [$_{\text{IP}}$ [Daddy [$_{\text{I}'}$ _$_k$ [$_{\text{VP}}$ drinks$_k$ juice]]].
    b. [$_{\text{CP}}$ Pappa$_i$ [$_{\text{C}'}$ trinkt$_k$ [$_{\text{IP}}$ _$_i$ [$_{\text{I}'}$ [$_{\text{VP}}$ Saft _$_k$] _$_k$]]]].

English has the basic word order SVO. The verb forms a constituent with the object (VP) and this is combined with the subject. The parameter setting must therefore be SV, VO and −V2. German, on the other had, is analyzed as a verb-final and verb-second language and the parameter values would therefore have to be SV, OV and +V2. If we consider the sentences in (2), we see that both sentences do not differ from one another with regard to the order of the verb and its arguments.

Fodor (1998a;b) concludes from this that one first has to build a structure in order to see what grammatical class the grammar licensing the structure belongs to since one first needs the structure in (3b) in order to be able to see that the verb in the partial constituent occurs after its argument in the VP (Saft _$_k$). The question is now how one achieves this structure. A UG with 30 parameters corresponds to 2^{30} = 1,073,741,824 fully instantiated grammars. It is an unrealistic assumption that children try out these grammars successively or simultaneously.

Gibson & Wexler (1994) discuss a number of solutions for this problem: Parameters have a default value and the learner can only change a parameter value if a sentence that could previously not be analyzed can then be analyzed with the new parameter (*Greediness Constraint*). In this kind of procedure, only one parameter can be changed at a time (*Single Value Constraint*), which aims at ruling out great leaps leading to extremely different grammars (see Berwick & Niyogi 1996: 612–613, however). This reduces the processing demands, however with 40 parameters, the worst case could still be that one has to test 40 parameter values separately, that is, try to parse a sentence with 40 different grammars. This processing feat is still unrealistic, which is why Gibson & Wexler (1994: 442) additionally assume that one hypothesis is tested per input sentence. A further modification of the model is the assumption that certain parameters only begin to play a role during the maturation of the child. At a given point in time, there could be only a few accessible parameters that also need to be set. After setting these parameters, new parameters could become available.

In their article, Gibson and Wexler show that the interaction between input and parameter setting is in no way trivial. In their example scenario with three parameters, the situation can arise that a learner sets a parameter in order to analyze a new sentence, however setting this parameter leads to the fact that the target grammar cannot be acquired because only one value can be changed at a time and changes can only be made if more sentences can be analyzed than before. The learner reaches a so-called local maximum in these problematic cases.[2] Gibson and Wexler then suggest assigning a default value to particular parameters, whereby the default value is the one that will cause the learner to avoid problematic situations. For the V2 parameter, they assume '−' as the default value.

Berwick & Niyogi (1996) show that Gibson and Wexler calculated the problematic conditions incorrectly and that, if one shares their assumptions, it is possible to arrive more frequently at parameter combinations from which it is not possible to reach the target grammar by changing individual parameter values. They show that one of the problematic cases not addressed by Gibson and Wexler is −V2 (p. 609) and that the assumption of a default value for a parameter does not solve the problem as both '+' and '−' can lead to problematic combinations of parameters.[3] In their article, Berwick and Niyogi show that learners in the example scenario above (with three parameters) learn the target grammar faster if one abandons the Greediness or else the Single Value Constraint. They suggest a process that simply randomly changes one parameter if a sentence cannot be analyzed (*Random Step*, p. 615–616). The authors note that this approach does not share the problems with the local maxima that Gibson and Wexler had in their example and that it also reaches its goal faster than theirs. However, the fact that *Random Step* converges more quickly has to do with the quality of the parameter space (p. 618). Since there is no consensus about parameters in the literature, it is not possible to assess how the entire system works.

Yang (2004: 453) has criticized the classic Principles & Parameters model since abrupt switching between grammars after setting a parameter cannot be demonstrated. Instead, he proposes the following learning mechanism:

[2] If one imagines the acquisition process as climbing a hill, then the Greediness Constraint ensures that one can only go uphill. It could be the case, however, that one begins to climb the wrong hill and can no longer get back down.

[3] Kohl (1999; 2000) has investigated this acquisition model in a case with twelve parameters. Of the 4096 possible grammars, 2336 (57%) are unlearnable if one assumes the best initial values for the parameters.

(4) For an input sentence, s, the child: (i) with probability P_i selects a grammar G_i, (ii) analyzes s with G_i,
(iii) if successful, reward G_i by increasing P_i, otherwise punish G_i by decreasing P_i.

Yang discusses the example of the pro-drop and topic drop parameters. In pro-drop languages (e. g. Italian), it is possible to omit the subject and in topic drop languages (e. g. Mandarin Chinese), it possible to omit both the subject and the object if it is a topic. Yang compares English-speaking with Chinese-speaking children and notes that English children omit both subjects and objects in an early linguistic stage. He traces this back to the fact that English-speaking children start off using the Chinese grammar.

The pro-drop parameter is one of the most widely discussed parameters in the context of Principles & Parameters theory and it will therefore be discussed in more detail here. It is assumed that speakers of English have to learn that all sentences in English require a subject, whereas speakers of Italian learn that subjects can be omitted. One can observe that children learning both English and Italian omit subjects (German children too in fact). Objects are also omitted notably more often than subjects. There are two possible explanations for this: a competence-based one and a performance-based one. In competence-based approaches, it is assumed that children use a grammar that allows them to omit subjects and then only later acquire the correct grammar (by setting parameters or increasing the rule apparatus). In performance-based approaches,by contrast, the omission of subjects is traced back to the fact that children are not yet capable of planning and producing long utterances due to their limited brain capacity. Since the cognitive demands are greatest at the beginning of an utterance, this leads to subjects beings increasingly left out. Valian (1991) investigated these various hypotheses and showed that the frequency with which children learning English and Italian respectively omit subjects is not the same. Subjects are omitted more often than objects. She therefore concludes that competence-based explanations are not empirically adequate. The omission of subjects should then be viewed more as a performance phenomenon (also, see Bloom (1993)). Another argument for the influence of performance factors is the fact that articles are left out more often with subjects than with objects (31% vs. 18%, see Gerken 1991: 440). As Bloom notes, no subject article-drop parameter has been proposed so far. If we explain this phenomenon as a performance phenomenon, then it is also plausible to assume that the omittance of subjects is due to performance issues.

Gerken (1991) shows that the metrical properties of utterances also play a role: In experiments where children had to repeat sentences, they omitted the subject/article more often in subjects than the object/article in objects. Here, it made a difference whether the intonation pattern was iambic (weak-strong) or trochaic (strong-weak). It can even be observed with individual words that children leave out weak syllables at the beginning of words more often than at the end of the word. Thus, it is more probable that „giRAFFE“ is reduced to „RAFFE“ than „MONkey“ to „MON“. Gerken assumes the following for the metrical structure of utterances:

1. Every metrical foot contains exactly one strong syllable.

2. Maximally binary feet are created from left to right.

3. Metrical structure is independent of syntactic structure.

Subject pronouns in English are sentence-initial and form a iambic foot with the following strongly emphasized verb as in (5a). Object pronouns, however, can form the weak syllable of a trochaic foot as in (5b).

(5) a. she KISSED + the DOG
b. the DOG + KISSED her
c. PETE + KISSED the + DOG

Furthermore, articles in iambic feet as in the object of (5a) and the subject of (5b) are omitted more often than in trochaic feet such as with the object of (5c).

It follows from this that there are multiple factors that influence the omission of elements and that one cannot simply take the behavior of children as evidence for switching between two grammars.

Apart from what has been discussed so far, the pro-drop parameter is of interest for another reason: There is a problem when it comes to setting parameters. The standard explanation is that learners identify that a subject must occur in all English sentences, which is suggested by the appearance of expletive pronouns in the input.

As discussed on page 501, there is no relation between the pro-drop property and the presence of expletives in a language. Since the pro-drop property does not correlate with any of the other putative properties either, only the existence of subject-less sentences in the input constitutes decisive evidence for setting a parameter. The problem is that there are grammatical utterances where there is no visible subject. Examples of this are imperatives such as (6) and even utterances without an expletive such as the example in (7b) found by Valian (1991: 32) in the New York Times.

(6) a. Give me the teddy bear!
b. Show me your toy!

(7) a. She'll be a big hit. Sings like a dream.
b. Seems like she always has something twin-related perking.

The following title of a Nirvana song also comes from the same year as Valian's article:

(8) Smells like Teen Spirit.

Teen Spirit refers to a deodorant and *smell* is a verb that, both in German and English, requires a referential subject but can also be used with an expletive *it* as subject. The usage that Curt Cobain had in mind cannot be reconstructed[4], independent of the intended meaning, however, the subject in (8) is missing. Imperatives do occur in the input children have and are therefore relevant for acquisition. Valian (1991: 33) says the following about them:

[4] See http://de.wikipedia.org/wiki/Smells_Like_Teen_Spirit. 18.04.2010.

> What is acceptable in the adult community forms part of the child's input, and is also part of what children must master. The utterances that I have termed "acceptable" are not grammatical in English (since English does not have pro subjects, and also cannot be characterized as a simple VP). They lack subjects and therefore violate the extended projection principle (Chomsky 1981a), which we are assuming.
>
> Children are exposed to fully grammatical utterances without subjects, in the form of imperatives. They are also exposed to acceptable utterances which are not fully grammatical, such as [(8a)], as well as forms like, "Want lunch now?" The American child must grow into an adult who not only knows that overt subjects are grammatically required, but also knows when subjects can acceptably be omitted. The child must not only acquire the correct grammar, but also master the discourse conditions that allow relaxation of the grammar. (Valian 1991: 33)

This passage turns the relations on their head: We cannot conclude from the fact that a particular grammatical theory is not compatible with certain data, that these data should not be described by this theory, instead we should modify the incompatible grammar or, if this is not possible, we should reject it. Since utterances with imperatives are entirely regular, there is no reason to categorize them as utterances that do not follow grammatical rules. The quotation above represents a situation where a learner has to acquire two grammars: one that corresponds to the innate grammar and a second that partially suppresses the rules of innate grammar and also adds some additional rules.

The question we can pose at this point is: How does a child distinguish which of the data it hears are relevant for which of the two grammars?

Fodor (1998a: 347) pursues a different analysis that does not suffer from many of the aforementioned problems. Rather than assuming that learners try to find a correct grammar among a billion others, she instead assumes that children work with a single grammar that contains all possibilities. She suggests using parts of trees (*treelets*) rather than parameters. These treelets can also be underspecified and in extreme cases, a treelet can consist of a single feature (Fodor 1998b: 6). A language learner can deduce whether a language has a given property from the usage of a particular treelet. As an example, she provides a VP treelet consisting of a verb and a prepositional phrase. This treelet must be used for the analysis of the VP occurring in *Look at the frog*. Similarly, the analysis of an interrogative clause with a fronted *who* would make use of a treelet with a *wh*-NP in the specifier of a complementizer phrase (see Figure 3.7 on page 102). In Fodor's version of Principles and Parameters Theory, this treelet would be the parameter that licenses *wh*-movement in (overt) syntax. Fodor assumes that there are defaults that allow a learner to parse a sentence even when no or very few parameters have been set. This allows one to learn from utterances that would have otherwise not been usable since there would have been multiple possible analyses for them. Assuming a default can lead to misanalyses, however: Due to a default value, a second parameter could be set because an utterance was analyzed with a treelet t_1 and t_3, for example, but t_1 was not suited to the particular language in question and the utterance should have instead been analyzed with the non-default treelet t_2 and the treelet t_{17}. In this acquisition model, there must therefore be the possibility to correct poor decisions in the parameter setting process. Fodor therefore assumes that there is a frequency-based degree of activation for parameters (p. 365): Treelets that are often used in analyses have a high degree of activation, whereas those used less often have a lower degree of activation. In this way, it is not necessary to assume a particular parameter value to the exclusion of others.

Furthermore, Fodor proposes that parameters should be structured hierarchically, that is, only if a parameter has a particular value does it then make sense to think about specific other parameter values.

Fodor's analysis is – as she herself notes (Fodor 2001: 385) – compatible with theories such as HPSG and TAG. Pollard & Sag (1987: 147) characterize UG as the conjunction of all universally applicable principles:

(9) $\text{UG} = \text{P}_1 \wedge \text{P}_2 \wedge \ldots \wedge \text{P}_n$

As well as principles that hold universally, there are other principles that are specific to a particular language or a class of languages. Pollard and Sag give the example of the constituent ordering principle that only holds for English. English can be characterized as follows if one assumes that P_{n+1}–P_m are language-specific principles L_1–L_p a complete list of lexical entries and R_1–R_q a list of dominance schemata relevant for English.

(10) $\text{English} = \text{P}_1 \wedge \text{P}_2 \wedge \ldots \wedge \text{P}_m \wedge (\text{L}_1 \vee \ldots \vee \text{L}_p \vee \text{R}_1 \vee \ldots \vee \text{R}_q)$

In Pollard and Sag's conception, only those properties of language that equally hold for all languages are part of UG. Pollard and Sag do not count the dominance schemata as part of this. However, one can indeed also describe UG as follows:

(11) $\text{UG} = \text{P}_1 \wedge \text{P}_2 \wedge \ldots \wedge \text{P}_n \wedge (\text{R}_{en-1} \vee \ldots \vee \text{R}_{en-q} \vee \text{R}_{de-1} \vee \ldots \vee \text{R}_{de-r} \vee \ldots)$

P_1–P_n are, as before, universally applicable principles and R_{en-1}–R_{en-q} are the (core) dominance schemata of English and R_{de-1}–R_{de-r} are the dominance schemata in German. The dominance schemata in (11) are combined by means of disjunctions, that is, not every disjunct needs to be realizable in a specific language. Principles can make reference to particular properties of lexical entries and rule out certain phrasal configurations. If a language only contains heads that are marked for final-position in the lexicon, then grammatical rules that require a head in initial position as their daughter can never be combined with these heads or their projections. Furthermore, theories with a type system are compatible with Fodor's approach to language acquisition because constraints can easily be underspecified. As such, constraints in UG do not have to make reference to all properties of grammatical rules: Principles can refer to feature values, the language-specific values themselves do not have to already be contained in UG. Similarly, a super type describing multiple dominance schemata that have similar but language-specific instantiations can also be part of UG, however the language-specific details remain open and are then deduced by the learner upon parsing (see Ackerman & Webelhuth 1998: Section 9.2). The differences in activation assumed by Fodor can be captured by weighting the constraints: The dominance schemata R_{en-1}–R_{en-q} etc. are sets of feature-value pairs as well as path equations. As explained in Chapter 15, one can add weights to such constraints and also to sets of constraints. In Fodor's acquisition model, given a German input, the weights for the rules of English would be reduced and those for the German rules would be increased. Note that in this acquisition scenario, there are no triggers for parameter setting unlike in Fodor's model. Furthermore, properties that were previously disjunctively specified as part of UG will now be acquired directly. Using the treelet t_{17} (or rather a possibly underspecified dominance schema), the parameter value '+' for a parameter P_5 is not derived but rather the activation potential of t_{17} is increased such that t_{17} will be prioritized for future analyses.

16.2 Principles and the lexicon

A variant of the UG-driven theory of language acquisition would be to assume that principles are so general that they hold for all languages and individual languages simply differ with regard to their lexicon. Principles then refer to properties of combined entities. Parameters therefore migrate from principles into the lexicon (Chomsky 1999: 2). See Mensching & Remberger (2011) for a study of Romance languages in this model and Son & Svenonius (2008: 395) for an analysis of Snyder's examples that were discussed in the previous subsection.

At this point, one can observe an interesting convergence in these approaches: Most of the theories discussed here assume a very general structure for the combination of heads with their arguments. For example, in Categorial Grammar and the Minimalist Program, these are always binary functor-argument combinations. The way in which constituents can be ordered in a particular language depends on the lexical properties of the combined elements.

The question that is being discussed controversially at present is whether the spectrum of lexical properties is provided by UG (Chomsky 2007: 6–7) and whether there are other areas of the language aside from the general combinatorial possibilities that may require a different description (see Section 23.10 on phrasal constructions).

In Section 16.1, I have shown what theories of acquisition assuming innate language specific knowledge can look like and also that variants of such acquisition theories are compatible with all the theories of grammar we have discussed. During this discussion, one should have kept in the back of ones mind the question of whether it makes sense at all to assume that English children use parts of a Chinese grammar during some stages of their acquisition process (as suggested by Yang (2004: 453)), or whether the relevant phenomena can be explained in different ways. In the following, I will present some alternative approaches that do not presuppose innate language specific knowledge, but instead assume that language can simply be acquired from the input. The following section will deal with pattern-based approaches and Section 16.4 will discuss lexically-oriented variants of input-based language acquisition.

16.3 Pattern-based approaches

Chomsky (1981a: 7–8) proposed that languages can be divided into a core area and a periphery. The core contains all regular aspects of language. The core grammar of a language is seen as an instantiation of UG. Idioms and other irregular parts of language are then part of the periphery. Critics of the Principles & Parameters model have pointed out that idiomatic and irregular constructions constitute a relatively large part of our language and that the distinction, both fluid and somewhat arbitrary, is only motivated theory-internally (Jackendoff 1997: Chapter 7; Culicover 1999; Ginzburg & Sag 2000: 5; Newmeyer 2005: 48; Kuhn 2007: 619). For example, it possible to note that there are interactions between various idioms and syntax (Nunberg, Sag & Wasow 1994). Most idioms in German with a verbal component allow the verb to be moved to initial position (12b), some allow that parts of idioms can be fronted (12c) and some can undergo passivization (12d).

(12) a. dass er ihm den Garaus macht
that he him the GARAUS makes
'that he finishes him off (kills him).'

b. Er macht ihm den Garaus.
he makes him the GARAUS
'He finishes him off.'

c. In Amerika sagte man der Kamera nach, die größte Kleinbildkamera der Welt zu sein. Sie war laut Schleiffer am Ende der Sargnagel der Mühlheimer Kameraproduktion. *Den Garaus machte* ihr die Diskussion um die Standardisierung des 16-Millimeter-Filmformats, an dessen Ende die DIN-Norm 19022 (Patrone mit Spule für 16-Millimeter-Film) stand, die im März 1963 zur Norm wurde.[5]
in America says one the camera after the biggest compact.camera of.the world to be. she was according.to Schleiffer at.the end the coffin.nail of.the Mühlheim camera.production the GARAUS made her the discussion around the standardization of.the 16-millimeter-film.format at whose end the DIN-norm 19022 cartridge with cleaning.fluid for 16-millimeter-film stood that in March 1963 to.the norm became
'In America, one says that this camera was the biggest compact camera in the world. According to Schleiffer, it was the last nail in the coffin for camera production in Mühlheim. What finished it off was the discussion about standardizing the 16 millimeter format, which ended in the DIN-Norm 19022 (cartridge with fluid for 16 millimeter film) that became the norm in March 1963.'

d. in Heidelberg wird „parasitären Elementen“ unter den Professoren *der*
in Heidelberg are parasitic elements among the professors the

[5] Frankfurter Rundschau, 28.06.1997, S. 2.

Garaus gemacht[6]
GARAUS made
'In Heidelberg, „parasitic elements" among professors are being killed off'

It is assumed of the periphery and the lexicon that they are not components of UG (Chomsky 1986b: 150–151; Fodor 1998a: 343) and that they are acquired using other learning methods – namely inductively directly from the input. The question posed by critics is now why these methods should not work for regular aspects of the language as well (Abney 1996: 20; Goldberg 2003a: 222; Newmeyer 2005: 100; Tomasello: 2006c: 36; 2006b: 20): The areas of the so-called 'core' are by definition more regular then components of the periphery, which is why they should be easier to learn.

Tomasello (2000; 2003) has pointed out that a Principles & Parameters model of language acquisition is not compatible with the observable facts. The Principles and Parameters Theory predicts that children should no longer make mistakes in a particular area of grammar once they have set a particular parameter correctly (see Chomsky 1986b: 146, Radford 1990: 21–22 and Lightfoot 1997: 175). Furthermore, it is assumed that a parameter is responsible for very different areas of grammar (see the discussion of the pro-drop parameter in Section 16.1). When a parameter value is set, then there should be sudden developments with regard to a number of phenomena (Lightfoot 1997: 174). This is, however, not the case. Instead, children acquire language from utterances in their input and begin to generalize from a certain age. Depending on the input, they can reorder certain auxiliaries and not others, although movement of auxiliaries is obligatory in English.[7] One argument brought against these kinds of input-based theories is that children produce utterances that cannot be observed to a significant frequency in the input. One much discussed phenomenon of this kind are so called *root infinitives* (RI) or *optional infinitives* (OI) (Wexler 1998). These are infinitive forms that can be used in non-embedded clauses (*root sentences*) instead of a finite verb. Optional infinitives are those where children use both a finite (13a) and non-finite (13b) form (Wexler 1998: 59):

(13) a. Mary likes ice cream.
b. Mary like ice cream.

Wijnen, Kempen & Gillis (2001: 656) showed that Dutch children use the order object infinitive 90 % of the time during the two-word phase although these orders occur in less than 10 % of their mother's utterances that contained a verb. Compound verb forms, e. g. with a modal in initial position (*Willst du Brei essen?* 'want you to eat porridge?') that contain another instance of this pattern only occurred in 30 % of the input containing a verb (Wijnen, Kempen & Gillis 2001: 647). At first glance, there seems to be a discrepancy between the input and the child's utterances. However, this deviation could also be explained by an utterance-final bias in learning (Wijnen et al. 2001; Freudenthal, Pine & Gobet 2006). A number of factors can be made responsible for the salience of verbs

[6] Mannheimer Morgen, 28.06.1999, Sport; Schrauben allein genügen nicht.
[7] Here, Yang's process does not help to combine grammars with a particular probability since one would have to assume that the child uses different grammars for different auxiliaries, which is highly unlikely.

at the end of an utterance: 1) restrictions of the infant brain. It has been shown that humans (both children and adults) forget words during the course of an utterance, that is, the activation potential decreases. Since the cognitive capabilities of small children are restricted, it is clear why elements at the end of an utterance have an important status. 2) Easier segmentation at the end of an utterance. At the end of an utterance, part of the segmentation problem for hearers disappears: The hearer first has to divide a sequence of phonemes into individual words before he can understand them and combine them to create larger syntactic entities. This segmentation is easier at the end of an utterance since the word boundary is already given by the end of the utterance. Furthermore according to Wijnen, Kempen & Gillis (2001: 637), utterance-final words have an above average length and do bear a pitch accent. This effect occurs more often in language directed at children.

Freudenthal, Pine, Aguado-Orea & Gobet (2007) have modeled language acquisition for English, German, Dutch and Spanish. The computer model could reproduce differences between these languages based on input. At first glance, it is surprising that there are even differences between German and Dutch and between English and Spanish with regard to the use of infinitives as German and Dutch have a very similar syntax (SOV+V2). Similarly, English and Spanish are both languages with SVO order. Nevertheless, children learning English make OI mistakes, whereas this is hardly ever the case for children learning Spanish.

Freudenthal, Pine, Aguado-Orea & Gobet (2007) trace the differences in error frequencies back to the distributional differences in each language: The authors note that 75 % of verb final utterances[8] in English consist of compound verbs (finite verb + dependent verb, e. g. *Can he go?*), whereas this is only the case 30 % of the time in Dutch.

German also differs from Dutch with regard to the number of utterance-final infinitives. Dutch has a progressive form that does not exist in Standard German:

(14) Wat ben je aan het doen?
what are you on the do
'What are you doing?'

Furthermore, verbs such as *zitten* 'to sit', *lopen* 'to run' and *staan* 'to stand' can be used in conjunction with the infinitive to describe events happening in that moment:

(15) Zit je te spelen?
sit you to play
'Are you sitting and playing?'

Furthermore, there is a future form in Dutch that is formed with *ga* 'go'. These factors contribute to the fact that Dutch has 20 % more utterance-final infinitives than German.

Spanish differs from English in that it has object clitics:

[8] For English, the authors only count utterances with a subject in third person singular since it is only in these cases that a morphological difference between the finite and infinitive form becomes clear.

(16) (Yo) Lo quiero.
I it want
'I want it.'

Short pronouns such as *lo* in (16) are realized in front of the finite verb so that the verb appears in final position. In English, the object follows the verb, however. Furthermore, there are a greater number of compound verb forms in the English input (70 %) than in Spanish (25 %). This is due to the higher frequency of the progressive in English and the presence of *do*-support in question formation.

The relevant differences in the distribution of infinitives are captured correctly by the proposed acquisition model, whereas alternative approaches that assume that children posses an adult grammar using infinitives rather than finite forms cannot explain the gradual nature of this phenomenon.

Freudenthal, Pine & Gobet (2009) could even show that input-based learning is superior to other explanations for the distribution of NPs and infinitives. They can explain why this order is often used with a modal meaning (e. g. *to want*) in German and Dutch (Ingram & Thompson 1996). In these languages, infinitives occur with modal verbs in the corresponding interrogative clauses. Alternative approaches that assume that the linguistic structures in question correspond to those of adults and only differ from them in that a modal verb is not pronounced cannot explain why not all utterances of object and verb done by children learning German and Dutch do have a modal meaning. Furthermore, the main difference to English cannot be accounted for: In English, the number of modal meanings is considerably less. Input-based models predict this exactly since English can use the dummy verb *do* to form questions:

(17) a. Did he help you?
b. Can he help you?

If larger entities are acquired from the end of an utterance, then there would be both a modal and non-modal context for *he help you*. Since German and Dutch normally do not use the auxiliary *tun* 'do', the relevant endings of utterances are always associated with modals contexts. One can thereby explain why infinitival expressions have a modal meaning significantly more often in German and Dutch than in English.

Following this discussion of the arguments against input-based theories of acquisition, I will turn to Tomasello's pattern-based approach. According to Tomasello (2003: Section 4.2.1), a child hears a sentence such as (18) and realizes that particular slots can be filled freely (also see Dąbrowska (2001) for analogous suggestions in the framework of Cognitive Grammar).

(18) a. Do you want more juice/milk?
b. Mommy is gone.

From these utterances, it is possible to derive so-called pivot schemata such as those in (19) into which words can then be inserted:

(19) a. more ___ → more juice/milk

b. ___ gone → mommy/juice gone

In this stage of development (22 months), children do not generalize using these schemata, these schemata are instead construction islands and do not yet have any syntax (Tomasello et al. 1997). The ability to use previously unknown verbs with a subject and an object in an SVO order is acquired slowly between the age of three and four (Tomasello 2003: 128–129). More abstract syntactic and semantic relations only emerge with time: When confronted with multiple instantiations of the transitive construction, the child is then able to generalize:

(20) a. [$_{\mathrm{S}}$ [$_{\mathrm{NP}}$ The man/the woman] sees [$_{\mathrm{NP}}$ the dog/the rabbit/it]].
b. [$_{\mathrm{S}}$ [$_{\mathrm{NP}}$ The man/the woman] likes [$_{\mathrm{NP}}$ the dog/the rabbit/it]].
c. [$_{\mathrm{S}}$ [$_{\mathrm{NP}}$ The man/the woman] kicks [$_{\mathrm{NP}}$ the dog/the rabbit/it]].

According to Tomasello (2003: 107), this abstraction takes the form [Sbj TrVerb Obj]. Tomasello's approach is immediately plausible since one can recognize how abstraction works: It is a generalization about reoccurring patterns. Each pattern is then assigned a semantic contribution. These generalizations can be captured in inheritance hierarchies (see page 206) (Croft 2001: 26). The problem with these kinds of approach, however, is that they cannot explain the interaction between different areas of phenomena in the language: It is possible to represent simple patterns such as the use of transitive verbs in (20), but transitive verbs interact with other areas of the grammar such as negation. If one wishes to connect the construction one assumes for the negation of transitive verbs with the transitive construction, then one arrives at a problem since this is not possible in inheritance hierarchies.

(21) The woman did not kick the dog.

The problem is that the transitive construction has a particular semantic contribution but that negated transitive construction has the opposite meaning. The values of SEM features would therefore be contradictory. There are technical tricks to avoid this problem, however, since there are a vast number of these kinds of interactions between syntax and semantics, this kind of technical solution will result in something highly implausible from a cognitive perspective (Müller 2006; 2007c;b; 2010b; Müller & Wechsler 2014a). For discussion of Croft's analysis, see Section 23.4.1.

At this point, proponents of pattern-based analyses might try and argue that these kinds of problems are only the result of a poor/inadequate formalization and would rather do without a formalization (Goldberg 2009: Section 5). However, this does not help here as the problem is not the formalization itself, rather the formalization allows one to see the problem more clearly.

An alternative to an approach built entirely on inheritance is a TAG-like approach that allows one to insert syntactic material into phrasal constructions. A relevant proposal was discussed in Section 10.6.3. Bergen & Chang (2005: 170) working in Embodied Construction Grammar suggest an Active-Ditransitive Construction with the form [RefExpr Verb RefExpr RefExpr], where RefExpr stands for a referential expression and there can be a gap between the first RefExpr and the verb. In this way, it is possible to analyze (22a,b), whereas (22c) is ruled out:

(22) a. Mary tossed me a drink.
b. Mary happily tossed me a drink.
c. * Mary tossed happily me a drink.

While the compulsory adjacency of the verb and the object correctly predicts that (22c) is ruled out, the respective constrain also rules out coordinate structures such as (23):

(23) Mary tossed me a juice and Peter a water.

Part of the meaning of this sentence corresponds to what the ditransitive construction contributes to *Mary tossed Peter a water*. There is, however, a gap between *tossed* and *Peter*. Similarly, one can create examples where there is a gap between both objects of a ditransitive construction:

(24) He showed me and bought for Mary the book that was recommended in the Guardian last week.

In (24), *me* is not adjacent to *the book* …. It is not my aim here to argue for a coordination analysis. Coordination is a very complex phenomenon for which most theories do not have a straightforward analysis (see Section 23.6.2). Instead, I would simply like to point out that the fact that constructions can be realized discontinuously poses a problem for approaches that claim that language acquisition is exclusively pattern-based. The point is the following: In order to understand coordination data in a language, a speaker must learn that a verb which has its arguments somewhere in the sentence has a particular meaning together with these arguments. The actual pattern [Sbj V Obj1 Obj2] can, however, be interrupted in all positions. In addition to the coordination examples, there is also the possibility of moving elements out of the pattern either to the left or the right. In sum, we can say that language learners have to learn that there is a relation between functors and their arguments. This is all that is left of pattern-based approaches but this insight is also covered by the selection-based approaches that we will discuss in the following section.

A defender of pattern-based approaches could perhaps object that there is a relevant construction for (24) that combines all material. This means that one would have a construction with the form [Sbj V Obj1 Conj V PP Obj2]. It would then have to be determined experimentally or with corpus studies whether this actually makes sense. The generalization that linguists have found is that categories with the same syntactic properties can be coordinated (N, $\overline{\text{N}}$, NP, V, $\overline{\text{V}}$, VP, …). For the coordination of verbs or verbal projections, it must hold that the coordinated word groups require the same arguments:

(25) a. Er [arbeitet] und [liest viele Bücher].
he works and reads many books
b. Er [kennt und liebt] diese Schallplatte.
he knows and loves this record

c. Er [zeigt dem Jungen] und [gibt der Frau] die Punk-Rock-CD.
   he shows the boy and gives the woman the punk rock CD

d. Er [liebt diese Schallplatte] und [schenkt ihr ein Buch].
   he loves this record and gives her a book

In an approach containing only patterns, one would have to assume an incredibly large number of constructions and so far we are only considering coordinations that consist of exactly two conjuncts. However, the phenomenon discussed above is not only restricted to coordination of two elements. If we do not wish to abandon the distinction between competence and performance (see Chapter 15), then the number of conjuncts is not constrained at all (by the competence grammar):

(26) Er [kennt, liebt und verborgt] diese Schallplatte.
     he knows loves and hides this record

It is therefore extremely unlikely that learners have patterns for all possible cases in their input. It is much more likely that they draw the kinds of generalizations that linguists have from the data occurring in their input: Word groups with the same syntactic properties can be coordinated. If this turns out to be true, then all that is left for pattern-based approaches is the assumption of discontinuously realized constructions and thus a dependency between parts of constructions that states that they do not have to be immediately adjacent to one another. The acquisition problem is then the same as for selection-based approaches that will be the topic of the following section: What ultimately has to be learned are dependencies between elements or valencies (see Behrens (2009: 439), the author reaches the same conclusion following different considerations).

16.4 Selection-based approaches

I will call the alternative to pattern-based approaches *selection-based*. A selection-based approach has been proposed by Green (2011).

The generalizations about the pattern in (20) pertain to the valency class of the verb. In Categorial Grammar, the pattern [Sbj TrVerb Obj] would correspond to the lexical entry (s\np)/np (for the derivation of a sentence with this kind of lexical entry, see Figure 8.3 on page 243). A TAG tree for *likes* was given on page 398. Here, one can see quite clearly that lexical entries provide the configurations for sentences in these models. Unlike pattern-based approaches, these analyses allow enough room for semantic embedding: The lexical entries in Categorial Grammar can be combined with adjuncts, and elementary trees in TAG also allow for adjunction to the relevant nodes.

Now, we face the question of how the jump from a pivot schema to a lexical entry with an argument structure takes place. In Tomasello's approach, there is no break between them. Pivot schemata are phrasal patterns and [Sbj TrVerb Obj] is also a phrasal pattern. Both schemata have open slots into which certain elements can be inserted. In selection-based approaches, the situation is similar: The elements that are fixed in the pivot schema are functors in the selection-based approach. Green (2011) proposes a theory of acquisition in HPSG that can do without UG. For the two-word phase, she assumes that *where's* is the head of an utterance such as (27) that selects *Robin* as its argument.

(27) Where's Robin?

This means that, rather than assuming that there is a phrasal pattern *Where's* X? with an empty slot X for a person or thing, she assumes that there is a lexical entry *Where's*, which contains the information that it needs to be combined with another constituent. What needs to be acquired is the same in each case: There is particular material that has to be combined with other material in order to yield a complete utterance.

In her article, Green shows how long-distance dependencies and the position of English auxiliaries can be acquired in later stages of development. The acquisition of grammar proceeds in a monotone fashion, that is, knowledge is added – for example, knowledge about the fact that material can be realized outside of the local context – and previous knowledge does not have to be revised. In her model, mistakes in the acquisition process are in fact mistakes in the assignment of lexical entries to valency classes. These mistakes have to be correctable.

In sum, one can say that all of Tomasello's insights can be applied directly to selection-based approaches and the problems with pattern-based approaches do not surface with selection-based approaches. It is important to point out explicitly once again here that the selection-based approach discussed here also is a construction-based approach. Constructions are just lexical and not phrasal. The important point is that, in both approaches, words and also more complex word groups are pairs of form and meaning and can be acquired as such.

In Chapter 23, we will discuss pattern-based approaches further and we will also explore areas of the grammar where phrasal patterns should be assumed.

16.5 Summary

We should take from the preceding discussion that models of language acquisition that assume that a grammar is chosen from a large set of grammars by setting binary parameters are in fact inadequate. All theories that make reference to parameters have in common that they are purely hypothetical since there is no non-trivial set of parameters that all proponents of the model equally agree on. In fact there is not even a trivial one.

In a number of experiments, Tomasello and his colleagues have shown that, in its original form, the Principles & Parameters model makes incorrect predictions and that language acquisition is much more pattern-based than assumed by proponents of P&P analyses. Syntactic competence develops starting from verb islands. Depending on the frequency of the input, certain verbal constructions can be mastered even though the same construction has not yet been acquired with less frequent verbs.

The interaction with other areas of grammar still remains problematic for pattern-based approaches: In a number of publications, it has been shown that the interaction of phenomena that one can observe in complex utterances can in fact not be explained with phrasal patterns since embedding cannot be captured in an inheritance hierarchy. This problem is not shared by selection-based approaches. All experimental results and insights of Tomasello can, however, be successfully extended to selection-based approaches.

Further reading

Meisel (1995) gives a very good overview of theories of acquisition in the Principles & Parameters model.

Goldberg (2006) and Tomasello (2003) are the most prominent proponents of Construction Grammar, a theory that explicitly tries to do without the assumption of innate linguistic knowledge.

An overview of different theories of acquisition can be found in Klann-Delius (2008).

17 Binary branching

We have seen that the question of what kind of branching structures have has received differing treatment in various theories. Classical $\overline{\mathrm{X}}$ 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 treat the question of branching in a similar 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 17.1 to start with and therefore, a learner would not have to rule out the corresponding hypotheses.

(1) Mummy must leave now.

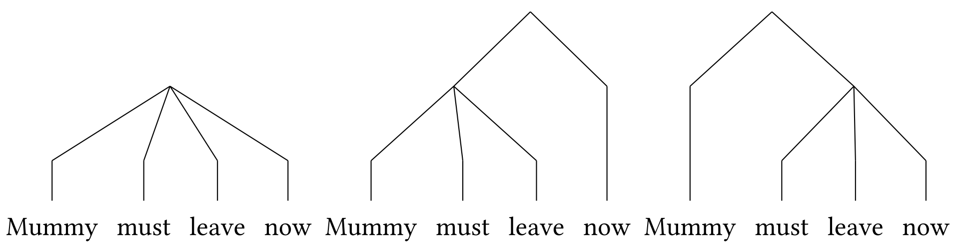


Figure 17.1: Structures with partial flat-branching

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 hypothesis: If all structures are binary-branching, then it is not possible to straightforwardly account for sentences with ditransitive verbs in $\overline{\text{X}}$ Theory. In $\overline{\text{X}}$ Theory, it is assumed that a head is combined with all its complements at once (see Section 2.5).

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 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 91), 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 (valency information that are ordered in a particular way, see Section 9.6).

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.

(3) a. [$_{\text{VP}}$ 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 17.2 on the next page, 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 derivation. In both analyses in Figure 17.2, 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, however, as 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

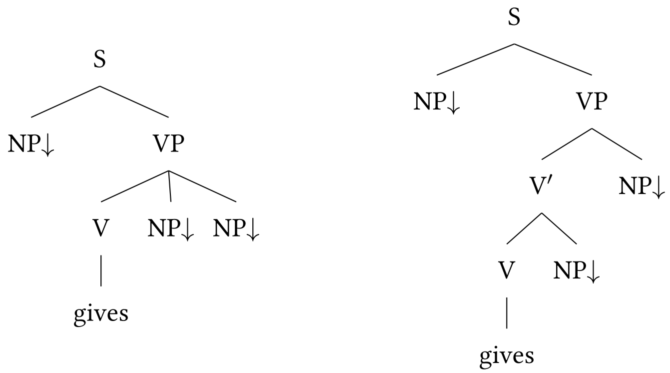


Figure 17.2: Flat and binary-branching elementary trees

from binding data for one of the structures in Figure 17.2 (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, a 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 as part of 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 17.2 are wrong. This is not correct, however, as it is the specification of the 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 17.2. 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 397). For the treatment of adjuncts in English, binary-branching structures therefore do not make any incorrect predictions.

18 Generative capacity and grammatical formalisms

In several of the preceding chapters, the complexity hierarchy for formal languages was mentioned. The simplest languages are so-called regular languages (Type 3), they are followed by those described as context-free grammars (Type 2), then those grammars which are context-sensitive (Type 1) and finally we have unrestricted grammars (Type 0) that create recursive languages, which are the most complicated class. In creating theories, a conscious effort was made to use the same formal means that correspond to what one can actually observe in natural language. This led to the abandonment of unrestricted Transformational Grammar since this has generative power of Type 0 (see page 87). GPSG was deliberately designed in such a way as to be able to analyze just the context-free languages and not more. In the mid-80s, it was shown that natural languages have a higher complexity than context-free languages (Shieber 1985; Culy 1985). It is now assumed that so-called *mildly context sensitive* grammars are sufficient for analyzing natural languages. Researchers working on TAG are working on developing variants of TAG that fall into exactly this category. Similarly, it was shown for different variants of Stabler's *Minimalist Grammars* (see Section 4.6.4 and Stabler (2001; 2011)) that they have a mildly context-sensitive capacity (Michaelis 2001). Peter Hellwig's Dependency Unification Grammar is also mildly context-sensitive (Hellwig 2003: 595).

LFG and HPSG, as well as Chomsky's theory in *Aspects*, fall into the class of Type 0 languages (Berwick 1982; Johnson 1988). The question at this point is whether it is an ideal goal to find a descriptive language that has exactly the same power as the object it describes. Carl Pollard (1996a) once said that it would be odd to claim that certain theories in physics were not adequate simply because they make use of tools from mathematics that are too powerful.[1] It is not the descriptive language that should constrain the theory but rather the theory contains the restrictions that must hold for the objects in question. This is the view that Chomsky (1981b: 277, 280) takes. Also, see Berwick (1982: Section 4), Kaplan & Bresnan (1982: Section 8) on LFG and Johnson (1988: Section 3.5) on the *Off-Line Parsability Constraint* in LFG and attribute-value grammars in general.

There is of course a technical reason to look for a grammar with the lowest level of complexity possible: We know that it is easier for computers to processes grammars

[1] If physicists required the formalism to constrain the theory
Editor: Professor Einstein, I'm afraid we can't accept this manuscript of yours on general relativity.
Einstein: Why? Are the equations wrong?
Editor: No, but we noticed that your differential equations are expressed in the first-order language of set theory. This is a totally unconstrained formalism! Why, you could have written down ANY set of differential equations! (Pollard 1996a)

with lower complexity than more complex grammars. For the relevant calculations, it is always a question of the 'worst case', that is, how long a program needs in the least favorable case to get a result for a certain part of the grammar from a certain class. This begs the question if the worst case is actually relevant. For example, some grammars that allow discontinuous constituents perform less favorably in the worst case than normal phrase structure grammars that only allow for combinations of contiguous strings (Reape 1991: Section 8). As I have shown in Müller (2004c), a parser that builds up larger units starting from words (a bottom-up parser) is far less efficient when processing a grammar assuming a verb movement analysis than is the case for a bottom-up parser that allows for discontinuous constituents. This has to do with the fact that verb traces do not contribute any phonological material and a parser cannot locate them without further machinery. It is therefore assumed that a verb trace exists in every position in the string that in most cases does not contribute to an analysis of the complete input. Since the verb trace is not specified with regard to its valency information, it can be combined with any material in the sentence, which results in an enormous computational load. On the other hand, if one allows discontinuous constituents, then one can do without verb traces and the computational load is thus reduced. In the end, the analysis using discontinuous constituents was eventually discarded due to linguistic reasons (Müller 2005b;c; 2007b), however, the investigation of the parsing behavior of both grammars is still interesting as it shows that worst case properties are not always informative.

I will discuss another example of the fact that language-specific restrictions can restrict the complexity of a grammar: Gärtner & Michaelis (2007: Section 3.2) assume that Stabler's Minimalist Grammars (see Section 4.6.4) with extensions for late adjunction and extraposition are actually more powerful than mildly context-sensitive. If one bans extraction from adjuncts (Frey & Gärtner 2002: 46) and also assumes the Shortest Move Constraint (see footnote 29 on page 163), then one arrives at a grammar that is mildly-context sensitive (Gärtner & Michaelis 2007: 178). The same is true of grammars with the Shortest Move constraint and a constraint for extraction from specifiers.

Whether extraction takes places from a specifier or not depends on the organization of the particular grammar in question. In some grammars, all arguments are specifiers (Kratzer (1996: 120–123), also see Figure 19.2 on page 530). A ban on extraction from specifiers would imply that extraction out of arguments would be impossible. This is, of course, not true in general. Normally, subjects are treated as specifiers (also by Frey & Gärtner (2002: 44)). It is often claimed that subjects are islands for extraction (see Grewendorf 1989: 35, 41; G. Müller: 1996b: 220; 1998: 32, 163; Sabel 1999: 98; Fanselow 2001: 422). Several authors have noted, however, that extraction from subjects is possible in German (see Dürscheid 1989: 25; Haider 1993: 173; Pafel 1993; Fortmann 1996: 27; Suchsland 1997: 320; Vogel & Steinbach 1998: 87; Ballweg 1997: 2066; Müller 1999a: 100–101; De Kuthy 2002: 7). The following data are corpus examples:

(1) a. [Von den übrigbleibenden Elementen]$_i$ scheinen [die Determinantien $\__i$] die wenigsten Klassifizierungsprobleme aufzuwerfen.[2]

[2] In the main text of Engel (1970: 102).

b. [Von den Gefangenen]$_i$ hatte eigentlich [keine _$_i$] die Nacht der Bomben überleben sollen.[3]

c. [Von der HVA]$_i$ hielten sich [etwa 120 Leute _$_i$] dort in ihren Gebäuden auf.[4]

d. [Aus dem „Englischen Theater“]$_i$ stehen [zwei Modelle _$_i$] in den Vitrinen.[5]

e. [Aus der Fraktion]$_i$ stimmten ihm [viele _$_i$] zu darin, dass die Kaufkraft der Bürger gepäppelt werden müsse, nicht die gute Laune der Wirtschaft.[6]

f. [Vom Erzbischof Carl Theodor Freiherr von Dalberg]$_i$ gibt es beispielsweise [ein Bild _$_i$] im Stadtarchiv.[7]

g. [Gegen die wegen Ehebruchs zum Tod durch Steinigen verurteilte Amina Lawal]$_i$ hat gestern in Negeria [der zweite Berufungsprozess _$_i$] begonnen.[8]

h. [Gegen diese Kahlschlagspolitik]$_i$ finden derzeit bundesweit [Proteste und Streiks _$_i$] statt.[9]

i. [Von den beiden, die hinzugestoßen sind], hat [einer _$_i$] eine Hacke, der andere einen Handkarren.[10]

j. ein Plan, [gegen den]$_i$ sich nun [ein Proteststurm _$_i$] erhebt[11]

k. Dagegen$_i$ jedoch regt sich jetzt [Widerstand _$_i$]: […][12]

l. [Aus der Radprofiszene]$_i$ kennt ihn [keiner _$_i$] mehr[13]

m. [Über das chinesische Programm der Deutschen Welle] tobt dieser Tage [ein heftiger Streit _$_i$].[14]

This means that a ban on extraction from specifiers cannot hold for German. As such, it cannot be true for all languages.

We have a situation that is similar to the one with discontinuous constituents: Since it is not possible to integrate the ban on extraction discussed here into the grammar formalism, it is more powerful than what is required for describing natural language. However, the restrictions in actual grammars – in this case the restrictions on extraction from specifiers in the relevant languages – ensure that the respective language-specific grammars have a mildly context-sensitive capacity.

[3] Bernhard Schlink, *Der Vorleser*, Diogenes Taschenbuch 22953, Zürich: Diogenes Verlag, 1997, p. 102.
[4] Spiegel, 3/1999, p. 42.
[5] Frankfurter Rundschau, quoted from De Kuthy (2001: 52).
[6] taz, 16.10.2003, p. 5.
[7] Frankfurter Rundschau, quoted from De Kuthy (2002: 7).
[8] taz, 28.08.2003, p. 2.
[9] Streikaufruf, Universität Bremen, 03.12.2003, p. 1.
[10] Murakami Haruki, *Hard-boiled Wonderland und das Ende der Welt*, suhrkamp taschenbuch, 3197, 2000, Translation by Annelie Ortmanns and Jürgen Stalph, p. 414.
[11] taz, 30.12.2004, p. 6.
[12] taz, 02.09.2005, p. 18.
[13] taz, 04.07.2005, p. 5.
[14] taz, 21.10.2008, p. 12.

19 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 construction of phrases inaccessible for adjacent or higher heads, that is, *glaubt* 'believe' in (1) selects a sentential argument but it cannot look inside this sentential argument.

(1) a. Karl glaubt, dass morgen seine Schwester kommt.
Karl believes that tomorrow his sister comes

b. Karl glaubt, dass sie morgen kommt.
Karl believes that she tomorrow comes

'Karl believes that his sister is coming tomorrow.'

Thus, *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 a NP with a particular case marking.

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 62, repeated here as (2), states that a ditransitive verb can occur with three noun phrases with the relevant case:

(2) 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 a 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 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 487). 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 have to therefore 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 a 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 296, 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 higher CP. It is assumed in GB theory that heads can look inside their arguments, at least they can see the elements inside their 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, an analysis for case assignment in infinitive constructions has been proposed as well, in which the case-assigning verb governs into the embedded phrase and assigns case to the element in SpecIP. Figure 19.1 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 case to the subject (cf. page 112), *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 is 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 19.2 on page 530 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 unlike in LFG and HPSG where complex feature bundles

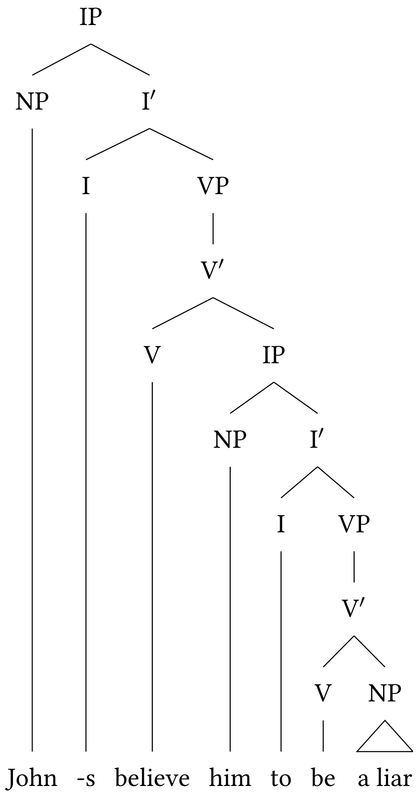


Figure 19.1: Analysis of the AcI construction with *Exceptional Case Marking*

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 Kratzer's proposal, case is asigned non-locally 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 the 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

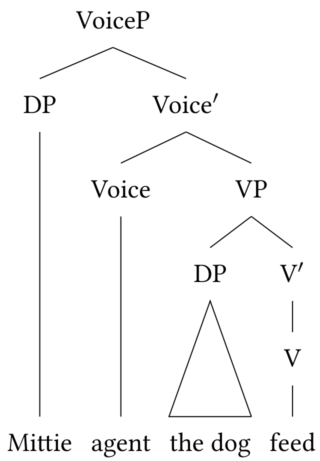


Figure 19.2: Analysis of structures with a transitive verb following Kratzer

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 (3):

(3) 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

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 objects (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 rule out 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:

(4) a. She/he 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 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:

(5) 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 (6). 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 (6). This means that one has to be able to formulate constraints affecting larger portions of syntactic structure.

(6) Ich glaube, mich / #dich tritt ein Pferd.[2]
I think me you kicks a horse
'Well, blow me down!'

Theories of grammar with extended locality domains do not have any problems with this kind of data.[3] 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 19.3 on the next page shows the tree for *kick the bucket* as it is used in (7a).

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

Since TAG trees can be separated again by adjunction, it is possible to insert elements between the parts of an idiom as in (7b,c) and thus explain the flexibility of idioms with regard to adjunction and embedding.[4] Depending on whether the lexical rules for the

[1] Also, see Sag & Pollard (1991: 89).
[2] Richter & Sailer (2009: 311).
[3] Or more carefully put: They do not have any serious problems since the treatment of idioms in their variety is by no means trivial (Sailer 2000).
[4] Interestingly, variants of Embodied CxG are strikingly similar to TAG. The Ditransitive Construction that was discussed on page 328 allows for additional material to occur between the subject and the verb.

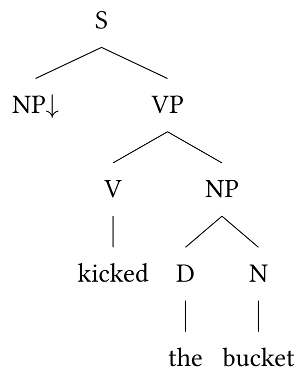


Figure 19.3: Elementary tree for *kick the bucket*

passive and long-distance dependencies can be applied, the idiom can occur in the relevant variants.

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 19.4 on the facing page shows a pertinent example from Abeillé & Schabes (1989: 7). The ban on adjunction is marked by a subscript NA.

The question that also arises for other theories is whether the efforts that have been made in deriving 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 parallel to the argumentation in the previous section. The argument there

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 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 a lie and the adjective has to be interpreted accordingly. The relevant tree should therefore contain nodes that contributes 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 326) or an NP tree can have narrow scope over the noun (*all alleged murderers*).

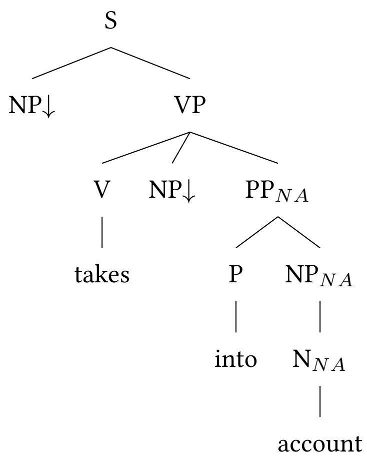


Figure 19.4: Elementary tree for *take into account*

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 accessible in principle, 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 301. 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 note on the list is crossed out. At the end of 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; 2008). Depictive predicates say something about the state of a person or object during the event expressed by a verb:

(8) a. Er sah sie nackt.[5]
he saw her naked
b. He saw her naked.

In (8), the depictive adjective can either refer to the subject or the object. However, there is a strong preference for readings where the referent noun precedes the depictive predicate (Lötscher 1985: 208). Figure 19.5 shows analyses for the sentences in (9):

(9) 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'

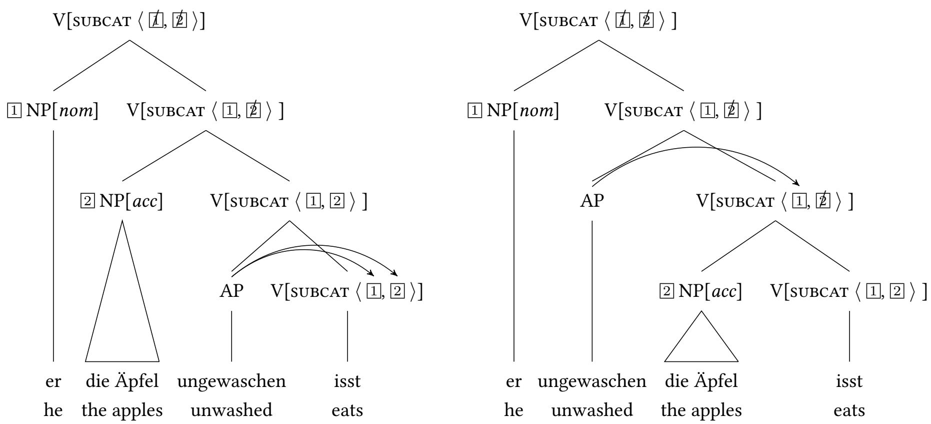


Figure 19.5: Analysis of *dass er die Äpfel ungewaschen isst* and *dass er ungewaschen die Äpfel isst*

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 referent noun can be captured by assuming a restriction that states that the referent noun must not yet have been realized.

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

(10) a. John [[$_{\text{VP}}$ ate the apples$_i$] unwashed$_i$].
b. You can't [[$_{\text{VP}}$ give them$_i$ injections] unconscious$_i$].[6]

[5] Haider (1985a: 94).
[6] Simpson (2005a: 17).

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 referent noun is inside the VP. English differs from German in that depictives can refer to both realized and unrealized 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 was available. In LFG, information about grammatical function, case and similar properties are 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 non-local access to phonological information or the part of speech, then this has to be explicitly encoded in the f-structure (see Bresnan (1982b: 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 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). 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. dative vs. accusative, 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.

20 Recursion

Every theory in this book can deal with self-embedding in language. 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 an sentence embedding or relative clause construction. A sentence embedding construction could have the form [Sbj that-Verb that-S]. A *that*-clause can then be inserted into the 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 (1):

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

In (1), 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. Constituents of the category *that*-S are also embedded in (1). 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 (2):

(2) [$_{\text{NP}}$ des Kragens [$_{\text{NP}}$ des Mantels [$_{\text{NP}}$ der Vorsitzenden]]]
the collar of.the coat of.the chair
'the collar of the coat of the chair of the committee'

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 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:

(3) [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 (3), the verb would have to be fixed in LTAG. But in Goldberg's analysis, verbs can be inserted into independently existent constructions (see Section 23.1). 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 shown in (1) and (2) – also by insertion to substitution nodes, thereby creating recursive structures.

21 Empty elements

This chapter deals with empty elements, I first discuss the general attitude of various research traditions towards empty elements and then show how they can be eliminated from grammars (Section 21.2). Section 21.3 discusses empty elements that were suggested in order to facilitate semantic interpretation. Section 21.4 discusses possible motivation for empty elements with a special focus on cross-linguistic comparison and the final Section 21.5 shows that certain accounts with transformations, lexical rules, and empty elements can be translated into each other.

21.1 Views on empty elements

One point that is particularly controversial among proponents of the theories discussed in this book is the question of whether one should assume empty elements or not. The discussion of empty elements is already quite old. There was already some investigation in 1961 with reference to phrase structure grammars (Bar-Hillel, Perles & Shamir 1961). The discussion of the status of empty elements has carried on ever since (see Löbner 1986; Wunderlich 1987; 1989; von Stechow 1989; Haider 1997a; Sag 2000; Bouma, Malouf & Sag 2001a; Levine & Hukari 2006; Müller 2014c, for example). There are sometimes empirical differences between analyses that assume empty elements and those that do not. But this is not often the case. Since empty elements often feature prominently in the argumentation for or against particular theories, I will discuss their use in somewhat more detail here.

In GB theory, empty elements were assumed for traces of movement (verb movement and fronting of phrases) as well as for deleted elements in elliptical constructions. Starting with the analysis of Larson (1988), more and more empty heads have been introduced, in order to ensure uniformity of structures and certain semantic interpretations (binding and scope). Other examples of an empty element that was introduced in order to maintain generalizations are empty expletives of Coopmans (1989: 734) and Postal (2004: Chapter 1). These fill the subject position in inversion structures in English, where the position preceding the verb is occupied by a PP and not by an overt subject NP. Similarly, Grewendorf (1993: 1311) assumes that the subject position in impersonal passives and passives without subject movement is in fact occupied by an empty expletive. See also Newmeyer (2005: 91) and Lohnstein (2014: 180) for this assumption with regard to the passive in German. Sternefeld (2006: Section II.3.3.3) assumes that there is an empty expletive subject in impersonal passives and subjectless sentences such as (1).

(1) a. Mir graut.
me.DAT scares
'I am scared.'

b. Mich dürstet.
me.ACC is.thirsty
'I am thirsty.'

On page 164, we discussed Stabler's proposal for the analysis of sentences with intransitive verbs. Since following Chomsky (2008: 146), the element that first merges with a head is the complement, intransitive verbs pose a problem for the theory. This problem is solved by Stabler by assuming that intransitive verbs are combined with an empty object (Veenstra 1998: 61, 124). Since these silent elements do not contribute to the meaning of an expression, we are also dealing with empty expletive pronouns.

In other theories, there are researchers that reject empty elements as well as those who assume them. In Categorial Grammar, Steedman follows analyses that do without empty elements. As Pollard (1988) has shown, one requires various kinds of type raising for NPs or a correspondingly high number of complex entries for relative pronouns (see Section 8.5.3). On the other hand, König (1999) uses traces. In GPSG, there is the traceless analysis of extraction by Uszkoreit (1987: 76–77) that we discussed in Section 5.4, but there is also the analysis of Gazdar, Klein, Pullum & Sag (1985: 143) that uses traces. In LFG, there are both analyses with traces (Bresnan 2001: 67) and those without (see Section 7.3 and Section 7.5). Many of the phrasal analyses in HPSG are borne out of the wish to avoid empty elements (see Section 23.10). An example for this is the relative clause analysis by Sag (1997) that replaces the empty relativizer in Pollard & Sag (1994) with a corresponding phrasal rule. Bender (2000) and Sag, Wasow & Bender (2003: 464) assume a silent copula. Another attempt to eliminate empty elements from HPSG was to handle long-distance dependencies not by traces but rather in the lexicon (Bouma, Malouf & Sag 2001a). As Levine & Hukari (2006) could show, however, theories of extraction that introduce long-distance dependencies lexically have problems with the semantic interpretation of coordinate structures. For a suggestion of how to solve these problems, see Chaves (2009). There are many TAG analyses without silent elements in the lexicon (see Section 12.5 and Kroch (1987), for example), however there are variants of TAG such as that of Kallmeyer (2005: 194), where a trace is assumed for the reordering of constituents in sentences with a verbal complex. Rambow (1994: 10–11) assumes an empty head in every verb phrase. In Dependency Grammar, Mel'čuk (1988: 303) and Engel (2014) assume empty elements for determiners and for coordinate structures, but Groß & Osborne (2009: 73) reject empty elements (with the exception of ellipsis, Osborne (2016)).

No empty elements are assumed in Construction Grammar (Michaelis & Ruppenhofer 2001: 49–50; Goldberg 2003a: 219; Goldberg 2006: 10), the related Simpler Syntax (Culicover & Jackendoff 2005) as well as in Cognitive Grammar.[1] The argumentation against empty elements runs along the following lines:

1. There is no evidence for invisible objects.

[1] However, Fillmore (1988: 51) did not rule them out.

2. There is no innate linguistic knowledge.

3. Therefore, knowledge about empty elements cannot be learned, which is why they cannot be assumed as part of our grammar.

This begs the question of whether all the premises on which the conclusion is based actually hold. If we consider an elliptical construction such as (2), then it is clear that a noun has been omitted:

(2) Ich nehme den roten Ball und du den blauen.
    I take the.ACC red.ACC ball and you the.ACC blue.ACC
    'I'll take the red ball and you take the blue (one).'

Despite there being no noun in *den blauen* 'the blue', this group of words behaves both syntactically and semantically just like a noun phrase. (2) is of course not necessarily evidence for the fact that there are empty elements, because one could simply say that *den blauen* is a noun phrase consisting only of an article and an adjective (Wunderlich 1987).

Similar to the fact that it is understood that a noun is missing in (2), speakers of English know that something is missing after *like*:

(3) Bagels, I like.

Every theory of grammar has to somehow account for these facts. It must be represented in some way that *like* in (3) behaves just like a verb phrase that is missing something. One possibility is to use traces. Bar-Hillel, Perles & Shamir (1961: 153, Lemma 4.1) have shown that it is possible to turn phrase structure grammars with empty elements into those without any. In many cases, the same techniques can be applied to the theories presented here and we will therefore discuss the point in more detail in the following section.

21.2 Eliminating empty elements from grammars

It is possible to turn a grammar with empty elements (also called *epsilon*) into a grammar without these by removing all categories that can be rewritten by an epsilon in every rule that uses such categories and then add the respective rules without the empty elements to the grammar. The following example has an epsilon rule for np. One therefore has to replace all rules containing the symbol np with new rules without this np symbol. (5) shows the result of this conversion of the grammar in (4):

(4) $\overline{\text{v}} \rightarrow$ np, v
    $\overline{\text{v}} \rightarrow$ np, pp, v
    np $\rightarrow \epsilon$

(5) $\overline{\text{v}} \rightarrow$ np, v
$\overline{\text{v}} \rightarrow$ v
$\overline{\text{v}} \rightarrow$ np, pp, v
$\overline{\text{v}} \rightarrow$ pp, v

This can also lead to cases where all elements on the right-hand side of a rule are removed. Thus, what one has done is actually create a new empty category and then one has to apply the respective replacement processes again. We will see an example of this in a moment. Looking at the pair of grammars in (4)–(5), it is clear that the number of rules has increased in (5) compared to (4) despite the grammars licensing the same sequences of symbols. The fact that an NP argument can be omitted is not expressed directly in (5) but instead is implicitly contained in two rules.

If you apply this procedure to the HPSG grammar in Chapter 9, then there is no specific category such as NP for the trace. The trace simply has to be compatible with a non-head daughter. As the examples in (6) show, adjuncts, arguments and parts of verbal complexes can be extracted.

(6) a. Er$_i$ liest t$_i$ die Berichte.
he reads the reports

b. Oft$_i$ liest er die Berichte t$_i$ nicht.
often reads he the reports not
'Often, he does not read the reports.'

c. Lesen$_i$ wird er die Berichte t$_i$ müssen.
read will he the reports must
'He will have to read the reports.'

The relevant elements are combined with their head in a special schemata (Head-Argument Schema, Head-Adjunct Schema, Predicate Complex Schema. See Chapter 9 for the first two schemata, the Predicate Complex Schema is motivated in detail in Müller (2007b: Chapter 15). If one wishes to do without traces, then one also needs an additional schema for the fronting of adjuncts, arguments and parts of predicate complexes. The combination of a head with a trace is given in Figure 21.1 on the facing page. The trace-less analysis is shown in Figure 21.2 on the next page. In Figure 21.1, the element in the SUBCAT list of *kennen* is identified with the SYNSEM value of the trace $\boxed{4}$. The lexical entry of the trace prescribes that the LOCAL value of the trace should be identical to the element in the INHER|SLASH list.

The Non-Local Feature Principle (page 293) ensures that the SLASH information is present on the mother node. Since an argument position is filled in Head-Argument Structures, the accusative object is no longer contained in the SUBCAT list of the mother node.

Figure 21.2 shows the parallel trace-less structure. The effect that one gets by combining a trace in argument position in Head-Argument structures, would be represented directly on the mother node in Figure 21.2: The LOCAL value of the accusative object was

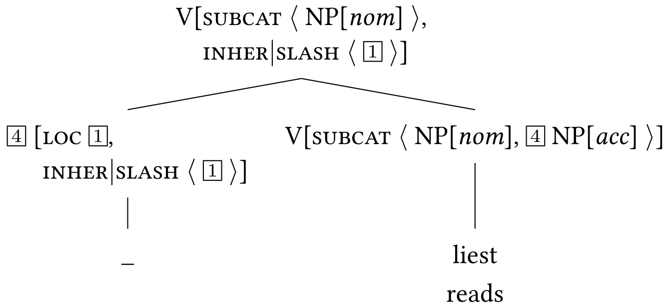


Figure 21.1: Introduction of information about long-distance dependencies with a trace

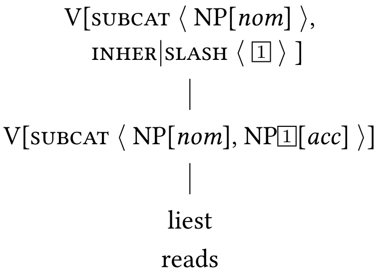


Figure 21.2: Introduction of information about long-distance dependencies using a unary projection

identified with the element in INHER|SLASH on the mother node and the accusative object does not occur in the valence list any more.

The grammar presented in Chapter 9 contains another empty element: a verb trace. This would then also have to be eliminated.

(7) a. Er$_i$ liest$_j$ t$_i$ die Berichte t$_j$.
he reads the reports

b. Oft$_i$ liest$_j$ er die Berichte t$_i$ nicht t$_j$.
often reads he the reports not
'Often, he does not read the reports.'

c. Lesen$_i$ wird$_j$ er die Berichte t$_i$ müssen t$_j$.
read will he the reports must
'He will have to read the reports.'

Figure 21.3 on the following page shows the combination of a verb trace with an accusative object. The verb trace is specified such that the DSL value is identical to the LOCAL value of the trace. Since DSL is a head feature, the corresponding value is also

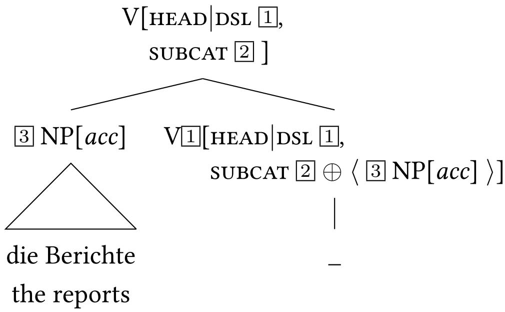


Figure 21.3: Analysis of verb position with verb trace

present on the mother node. Figure 21.4 shows the structures that we get by omitting the empty node. This structure may look odd at first sight since a noun phrase is pro-

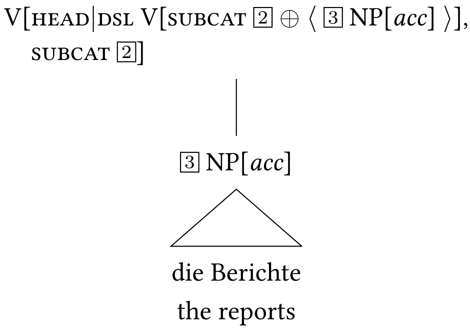


Figure 21.4: Analysis of verb position using a unary projection

jected to a verb (see page 231 for similar verb-less structures in LFG). The information about the fact that a verb is missing in the structure is equally contained in this structure as in the structure with the verb trace. It is the DSL value that is decisive for the contexts in which the structure in Figure 21.4 can appear. This is identical to the value in Figure 21.3 and contains information that a verb that requires an accusative object is missing in the structure in question. Until now, we have seen that traces of extraction can be removed from the grammar by stipulating three additional rules. Similarly, three new rules are needed for the verb trace. Unfortunately, it does not stop here as the traces can also interact. For example, the NP in the tree in Figure 21.4 could be an extraction trace. Therefore, the combination of traces can result in more empty elements that then also have to be eliminated. Since we have three schemata, we will have three new empty elements if we combine the non-head daughter with an extraction trace and the head daughter with a verb trace. (8) shows these cases:

(8) a. Er$_i$ [schläft$_j$ t$_i$ t$_j$]. (Extraction trace (argument) + verb trace)
he sleeps
'He is sleeping.'

b. Jetzt$_i$ [schlaf$_j$ t$_i$ t$_j$]! (Extraction trace (adjunct) + verb trace)
now sleep
'Go to sleep now!'

c. Geschlafen$_i$ [wird$_j$ t$_i$ t$_j$]! (Extraction trace (complex) + verb trace)
slept is
'Now is time to sleep!'

These three new traces can occur as non-head daughters in the Head-Argument Schema and thus one would require three new schemata for head-argument structures. Using these schemata, it then becomes possible to analyze the sentences in (8).

Six further schemata are required for the examples in (9) and (10) since the three new traces can each occur as heads in head-argument structures (9) and head-adjunct structures (10):

(9) a. Den Aufsatz$_i$ liest$_j$ [er t$_i$ t$_j$].
the essay reads he
'He is reading the essay.'

b. Oft$_i$ liest$_j$ er [ihn t$_i$ t$_j$].
often reads he it
'He often reads it.'

c. Lesen$_i$ wird$_j$ er [ihn t$_i$ t$_j$].
read will he it
'He will read it.'

(10) a. Den Aufsatz$_i$ liest$_j$ er [nicht t$_i$ t$_j$].
the essay reads he not
'He isn't reading the essay.'

b. Oft$_i$ liest$_j$ er ihn [nicht t$_i$ t$_j$].
often reads he it not
'He often doesn't read it'

c. Lesen$_i$ wird$_j$ er ihn [nicht t$_i$ t$_j$].
reads will he it not
'He won't read it.'

The price of eliminating two empty elements is therefore the introduction of twelve new rules. These rules are not particularly transparent and it is not immediately obvious why the mother node describes a linguistic object that follows general grammatical laws. For example, there are no heads in the structures following the pattern in Figure 21.4. Since there is no empirical difference between the theoretical variant with twelve additional schemata and the variant with two empty elements, one should prefer the theory that makes fewer assumptions (Occam's Razor) and that is the theory with two empty elements.

One might think that the problem discussed here is just a problem specific to HPSG not shared by trace-less analyses such as the LFG approach that was discussed in Section 7.5. If we take a closer look at the rule proposed by Dalrymple (2006: Section 2.2), we see that the situation in LFG grammars is entirely parallel. The brackets around the category symbols mark their optionality. The asterisk following the PP means that any number of PPs can occur in this position.

(11) $\mathrm{V}' \rightarrow (\mathrm{V})\ (\mathrm{NP})\ \mathrm{PP}^*$

This means that (11) is a shorthand for rules such as those in (12):

(12)
a. $\mathrm{V}' \rightarrow \mathrm{V}$
b. $\mathrm{V}' \rightarrow \mathrm{V}\ \mathrm{NP}$
c. $\mathrm{V}' \rightarrow \mathrm{V}\ \mathrm{NP}\ \mathrm{PP}$
d. $\mathrm{V}' \rightarrow \mathrm{V}\ \mathrm{NP}\ \mathrm{PP}\ \mathrm{PP}$
e. …
f. $\mathrm{V}' \rightarrow \mathrm{NP}$
g. $\mathrm{V}' \rightarrow \mathrm{NP}\ \mathrm{PP}$
h. $\mathrm{V}' \rightarrow \mathrm{NP}\ \mathrm{PP}\ \mathrm{PP}$
i. …

Since all the elements on the right-hand side of the rule are optional, (11) also stands for (13):

(13) $\mathrm{V}' \rightarrow \epsilon$

Thus, one does in fact have an empty element in the grammar although the empty element is not explicitly listed in the lexicon. This follows from the optionality of all elements on the right-hand side of a rule. The rule in (12f) corresponds to the schema licensed by the structure in Figure 21.4. In the licensed LFG structure, there is also no head present. Furthermore, one has a large number of rules that correspond to exactly the schemata that we get when we eliminate empty elements from an HPSG grammar. This fact is, however, hidden in the representational format of the LFG rules. The rule schemata of LFG allow for handy abbreviations of sometimes huge sets of rules (even infinite sets when using '*').

Pollard (1988) has shown that Steedman's trace-less analysis of long-distance dependencies is not without its problems. As discussed in Section 8.5.3, a vast number of recategorization rules or lexical entries for relative pronouns are required.

21.3 Empty elements and semantic interpretation

In this section, I discuss an analysis that assumes empty elements in order to allow for different readings of particular sentences. I then show how one can use so-called underspecification approaches to do without empty elements.

Sentences such as (14) are interesting since they have multiple readings (see Dowty 1979: Section 5.6) and it is not obvious how these can be derived.

(14) dass Max alle Fenster wieder öffnete
that Max all windows again opened
'that Max opened all the windows again'

There is a difference between a repetitive and a restitutive reading: For the repetitive reading of (14), Max has to have opened every window at least once before, whereas the restitutive reading only requires that all windows were open at some point, that is, they could have been opened by someone else.

These different readings are explained by dividing the predicate *öffnen′* 'open' into at least two sub-predicates. Egg (1999) suggests the division into CAUSE and *offen′* 'open':

(15) CAUSE(x, *open′*(y))

This means that there is a CAUSE operator that has scope over the relation *offen′* 'open'. Using this kind of division, it is possible to capture the varying scope of *wieder* 'again': In one of the readings, *wieder* scopes over CAUSE and it scopes over *offen′* in the other. If we assume that *öffnen* has the meaning in (15), then we still have to explain how the adverb can modify elements of a word's meaning, that is, how *wieder* 'again' can refer to *offen′* 'open'. Von Stechow (1996: 93) developed the analysis in Figure 21.5 on the following page. AgrS and AgrO are functional heads proposed for subject and object agreement in languages like Basque and have been adopted for German (see Section 4.6). Noun phrases have to be moved from the VoiceP into the specifier position of the AgrS and AgrO heads in order to receive case. T stands for Tense and corresponds to Infl in Chapter 3. What is important is that there is the Voice head and the separate representation of *offen* 'open' as the head of its own phrase. In the figure, everything below Voice′ corresponds to the verb *öffnen*. By assuming a separate Voice head that contributes causative meaning, it becomes possible to derive both readings in syntax: In the reading with narrow scope of *wieder* 'again', the adverb is adjoined to the XP and has scope over open(x). In the reading with wide scope, the adverb attaches to VoiceP or some higher phrase and therefore has scope over CAUSE(BECOME(open(x))).

Jäger & Blutner (2003) point out that this analysis predicts that sentences such as (16) only have the repetitive reading, that is, the reading where *wieder* 'again' has scope over CAUSE.

(16) dass Max wieder alle Fenster öffnete
that Max again all windows opened

This is because *wieder* precedes *alle Fenster* and therefore all heads that are inside VoiceP. Thus, *wieder* can only be combined with AgrOP or higher phrases and therefore has (too)

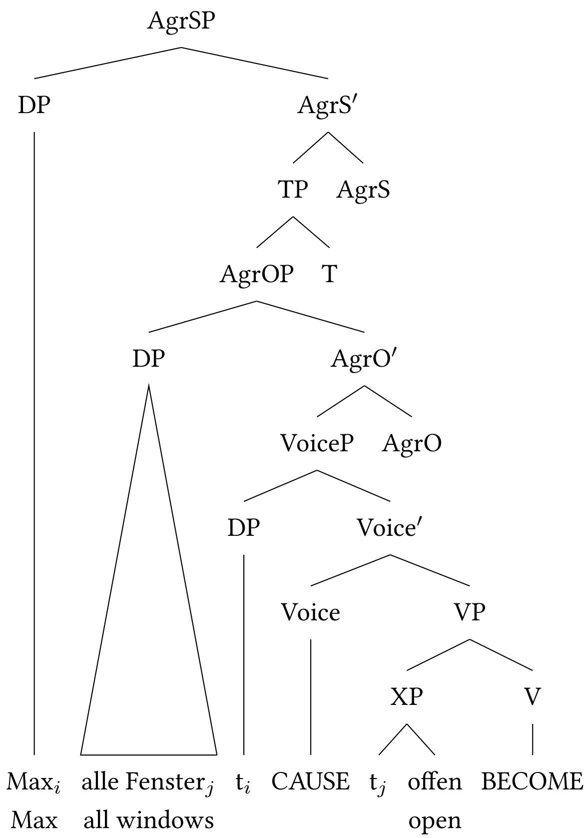


Figure 21.5: Decomposition in syntactic structures

wide scope. (16) does permit a restitutive reading, however: All windows were open at an earlier point in time and Max reestablishes this state.

Egg (1999) develops an analysis for these *wieder* cases as part of *Constraint Language for Lambda-Structures* (CLLS). CLLS is an underspecification formalism, that is, no logical formula are given but instead expressions that describe logical formulae. Using these kind of expressions, it is possible to leave scope relations underspecified. I have already mentioned Minimal Recursion Semantics (MRS) (Copestake, Flickinger, Pollard & Sag 2005) in several chapters of this book. As well as CLLS, MRS together with Underspecified Discourse Representation Theory (Reyle 1993; Frank & Reyle 1995) and *Hole Semantics* (Bos 1996; Blackburn & Bos 2005) all belong to the class of underspecification formalisms. See Baldridge & Kruijff (2002) for an underspecification analysis in Categorial Grammar and Nerbonne (1993) for an early underspecification analysis in HPSG. In the following, I will reproduce Egg's analysis in MRS notation.

Before we turn to (14) and (16), let us consider the more simple sentence in (17):

(17) dass Max alle Fenster öffnete
that Max all windows opened
'that Max opened all the windows.'

This sentence can mean that in a particular situation, it is true of all windows that Max opened them. A less readily accessible reading is the one in which Max causes all of the windows to be open. It is possible to force this reading if one rules out the first reading through contextual information (Egg 1999):

(18) Erst war nur die Hälfte der Fenster im Bus auf, aber dann öffnete
first was only the half of.the windows in.the bus open but then opened
Max alle Fenster.
Max all windows
'At first, only half of the windows in the bus were open, but then Max opened all of the windows.'

Both readings under discussion here differ with regard to the scope of the universal quantifier. The reading where Max opens all the windows himself corresponds to wide scope in (19a). The reading where some windows could have already been open corresponds to (19b):

(19) a. $\forall$ x *window′*(x) $\rightarrow$ CAUSE(*max′*, *open′*(x))
b. CAUSE(*max′*, $\forall$ x *window′*(x) $\rightarrow$ *open′*(x))

Both of these readings can be represented as underspecified in a dominance graph such as the one given in Figure 21.6. Each relation in Figure 21.6 has a name that one can

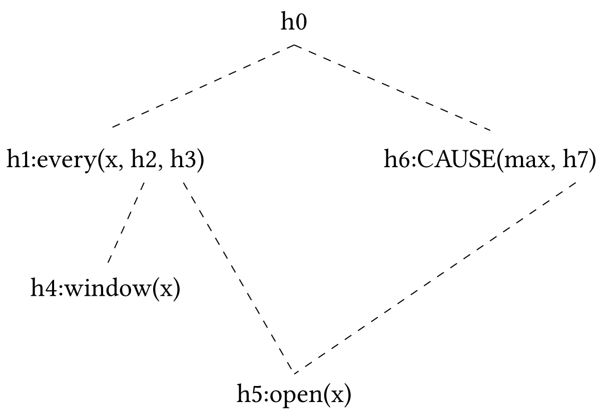


Figure 21.6: Dominance graph for *Max alle Fenster öffnete*

use to refer to the relation or „grasp" it. These names are referred to as *handle*. The dominance graph states that $h0$ dominates both $h1$ and $h6$ and that $h2$ dominates $h4$, $h3$ dominates $h5$, and $h7$ dominates $h5$. The exact scopal relations are underspecified: The universal quantifier can have scope over CAUSE or CAUSE can have scope over the universal quantifier. Figures 21.7 and 21.8 shows the forms with resolved scope. The fact

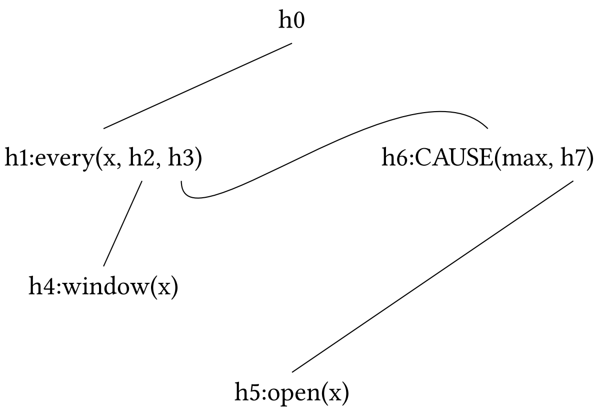


Figure 21.7: Dominance graph for the reading ∀ x window(x) → CAUSE(max,open(x)).

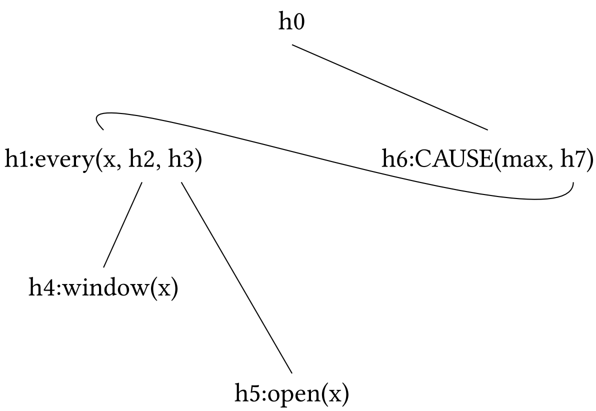


Figure 21.8: Graph for te reading CAUSE(max, ∀ x window(x) → open(x)).

that the quantifier dominates $h4$ is determined the lexical entry of the quantifier. The fact that the quantifier dominates $h5$ does not have to be made explicit in the analysis since the quantifier binds a variable in the relation belonging to $h5$ (x). The dominance relation between $h7$ and $h5$ is always determined in the lexicon since CAUSE and *offen′* 'open' both belong to the semantic contribution of a single lexical entry.

The exact syntactic theory that one adopts for this analysis is, in the end, not of great importance. I have chosen HPSG here. As Figure 21.9 shows, the analysis of *alle Fenster öffnet* contains a simple structure with a verb and an object. This structure does not

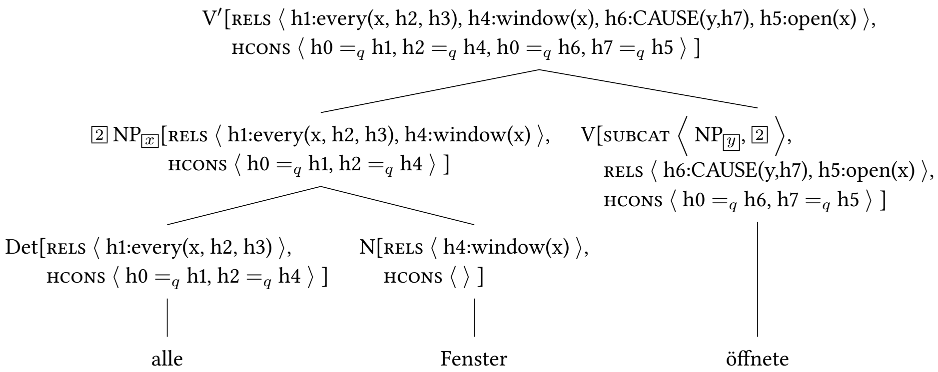


Figure 21.9: MRS analysis of *alle Fenster öffnete*

differ from the one that would be assumed for *alle Kinder kennt* ‘all children know’. The only difference comes from the meaning of the individual words involved. As shown in Section 9.1.6, relations between individual words are passed on upwards. The same happens with scopal restrictions. These are also represented in lists. HCONS stands for *handle constraints*. $=_q$ in h0 $=_q$ h6 stand for the equality *modulo* quantifier scope.

Egg lists the following readings for the sentence in (16) – repeated here as (20):

(20) dass Max wieder alle Fenster öffnete
that Max again all windows opened
‘that Max opened all the windows again’

1. Max opened every window and he had already done that at least once for each window (*again′*(∀(CAUSE(open))); repetitive)
2. Max caused every window to be open and he had done that at least once before (*again′*(CAUSE(∀(open))); repetitive)
3. At some earlier point in time, all windows were simultaneously open and Max re-established this state (CAUSE(*wieder′*(∀(offen))); restitutive)

These readings correspond to the dominance graph in Figure 21.10 on the next page. Figure 21.11 on the following page shows the graph for (14) – repeated here as (21):

(21) dass Max alle Fenster wieder öffnete
that Max all windows again opened

To derive these dominance graphs from the ones without *wieder* ‘again’, all one has to do is add the expression h8:wieder(h9) and the dominance requirements that demand

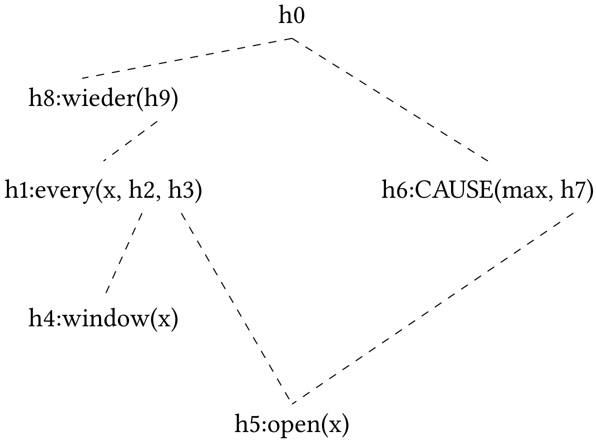


Figure 21.10: Dominance graph for *Max wieder alle Fenster öffnete*

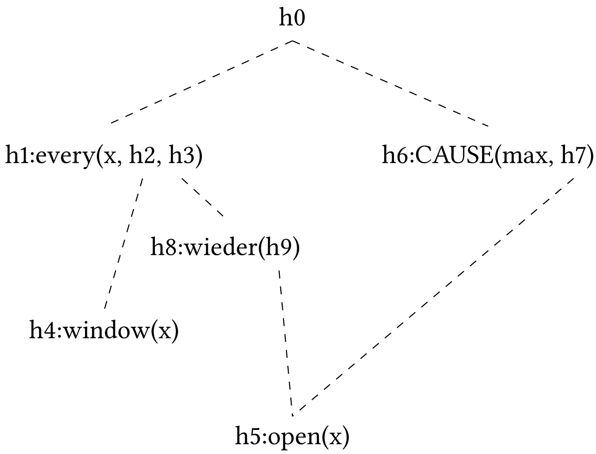


Figure 21.11: Dominance graph for *Max alle Fenster wieder öffnete*

that $h9$ dominates quantifiers occurring to the right of *wieder* and that it is dominated by quantifiers to the left of *wieder*.

It is therefore unproblematic to derive the relevant readings for modification by *wieder* without empty elements for CAUSE and BECOME. Similarly, the meaning of a word is decomposed but the decomposed meaning is assigned to a single element, the verb. By underspecification of the scopal relations in the lexicon, the relevant readings can then be derived.

21.4 Evidence for empty elements

As previously discussed, grammarians agree that both linguists and speakers notice when there is a constituent missing from a string of words. For cases where it can be shown that analyses with or without traces are indistinguishable empirically, then one can assume empty elements. Nevertheless, the learnability argument put forward by Construction Grammarians has some validity: If one assumes that there is no innate linguistic knowledge, then it is not possible to motivate empty elements with data from other languages. This means that just because Basque shows object agreement, this does not mean that one can assume an empty head for object agreement in a grammar of German (AgrO) as for instance von Stechow (1996) and Meinunger (2000) do. Since there is no object agreement in German, there would be no way for the child to learn the fact that there is an AgrO head. Knowledge about AgrO must therefore be innate. Since the assumption of innate linguistic knowledge is controversial (see Chapter 13), any theory that uses cross-linguistic data to motivate the use of empty elements is on shaky ground.

Cross-linguistic considerations can only be drawn upon if there are no empirical differences between multiple alternative analyses. In this case, one should follow Occam's Razor and choose that analysis which is compatible with analyses of other languages.

21.5 Transformations, lexical rules, and empty elements

In the discussion of the passive in the framework of TAG, it became clear that lexical rules correspond to particular transformations, namely those which have some relation to a lexical entry (*lexically governed transformations*, Dowty (1978), for the discussion of transformations and lexical rules, see Bresnan (1978); Bresnan & Kaplan (1982)). In respective variants of TAG, lexical rules establish a relation between a lexical item for an active tree with a lexical item of a passive tree. Both the active and passive tree can be extended by adjunction.

In theories such as Categorial Grammar, the situation is similar: Since the direction in which a functor expects to find its argument is fixed for languages such as English, the lexical item stands for an entire tree. Only the attachment of adjuncts is not yet specified in lexical items. The positions in the tree where the adjuncts can occur depend on the properties of the adjuncts. In Section 8.4, we have seen suggestions of how to treat languages with free constituent order. If the direction of combination is not fixed in the lexicon, then the lexical item can occur in a number of trees. If we compare lexical rules that can be applied to these kind of lexical items with transformations, we see that lexical rules create relations between different sets of trees.

In HPSG analyses, this works in a similar away: Lexical rules relate lexical items with differing valency properties to each other. In HPSG grammars of English, there is normally a schema that licenses a VP containing the verb and all its complements as well as a schema that connects the subject to the VP (Pollard & Sag 1994: 39). In the lexical items for finite verbs, it is already determined what the tree will look like in the end. As in Categorial Grammar, adjuncts in HPSG can be combined with various intermediate projections. Depending on the dominance schemata used in a particular grammar, the lexical item will determine the constituent structure in which it can occur or allow for multiple structures. In the grammar of German proposed in Chapter 9, it is possible to analyze six different sequences with a lexical item for a ditransitive verb, that is, the lexical item can – putting adjuncts aside – occur in six different structures with verb-final order. Two sequences can be analyzed with the passive that only has two arguments. As in Categorial Grammar, sets of licensed structure are related to other sets of licensed structures. In HPSG theorizing and also in Construction Grammar, there have been attempts to replace lexical rules with other mechanisms since their 'status is dubious and their interaction with other analyses is controversial' (Bouma, Malouf & Sag 2001a: 19). Bouma et al. (2001a) propose an analysis for extraction that, rather than connecting lexical items with differing valence lists, establishes a relation between a subset of a particular list in a lexical item and another list in the same lexical item. The results of the two alternative analyses are shown in (22) and (23), respectively:

(22) a. $\begin{bmatrix} \text{SUBCAT} & \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SLASH} & \langle\rangle \end{bmatrix}$

b. $\begin{bmatrix} \text{SUBCAT} & \langle \text{NP}[\textit{nom}] \rangle \\ \text{SLASH} & \langle \text{NP}[\textit{acc}] \rangle \end{bmatrix}$

In (22), (22a) is the basic entry and (22b) is related to (22a) via a lexical rule. The alternative analysis would only involve specifying the appropriate value of the ARG-ST feature[2] and the SUBCAT and SLASH value is then derived from the ARG-ST value using the relevant constraints. (23) shows two of the licensed lexical items.

(23) a. $\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SUBCAT} & \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SLASH} & \langle\rangle \end{bmatrix}$

b. $\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \\ \text{SUBCAT} & \langle \text{NP}[\textit{nom}] \rangle \\ \text{SLASH} & \langle \text{NP}[\textit{acc}] \rangle \end{bmatrix}$

[2] ARG-ST stands for *Argument Structure*. The value of ARG-ST is a list containing all the arguments of a head. For more on ARG-ST, see Section 9.6.

If we want to eliminate lexical rules entirely in this way, then we would require an additional feature for each change.[3] Since there are many interacting valency-changing processes, things only work out with the stipulation of a large number of auxiliary features. The consequences of these analyses have been discussed in detail in Müller (2007b: Section 7.5.2.2). A parallel problem arises for inheritance-based approaches for argument structure-changing processes: They also require auxiliary features since it is not possible to model embedding only with inheritance. See Section 10.2.

The claim that the status of lexical rules is dubious must be refuted: There are worked-out formalizations of lexical rules (Meurers 2001; Copestake & Briscoe 1992; Lascarides & Copestake 1999) and their interaction with other analyses is not controversial. Most HPSG implementations make use of lexical rules and the interaction of a number of rules and constraints can be easily verified by experiments with implemented fragments.

Jackendoff (1975) presents two possible conceptions of lexical rules: In one variant, the lexicon contains all words in a given language and there are just redundancy rules saying something about how certain properties of lexical entries behave with regard to properties of other lexical entries. For example, *les-* 'read-' and *lesbar* 'readable' would both have equal status in the lexicon. In the other way of thinking of lexical rules, there are a few basic lexical entries and the others are derived from these using lexical rules. The stem *les-* 'read-' would be the basic entry and *lesbar* would be derived from it. In HPSG, the second of the two variants is more often assumed. This is equivalent to the assumption of unary rules. In Figure 9.8 on page 285, this has been shown accordingly: The verb *kennt* 'knows' is mapped by a lexical rule to a verb that selects the projection of an empty verbal head. With this conception of lexical rules, it is possible to remove lexical rules from the grammar by assuming binary-branching structures with an empty head rather than unary rules. For example, in HPSG analyses of resultative constructions lexical rules such as (24) have been proposed (Verspoor 1997; Wechsler 1997; Wechsler & Noh 2001; Müller 2002a: Chapter 5).

(24) [dass] Peter den Teich leer fischt
that Peter the pond empty fishes
'that Peter fishes the pond empty'

In my own analysis, a lexical rule connects a verb used intransitively to a verb that selects an accusative object and a predicate. Figure 21.12 on the next page shows the corresponding tree. If we consider what (24) means, then we notice that the fishing act causes the pond to become empty. This causation is not contained in either of the basic lexical entries for the words in (24). In order for this information to be present in the semantic representation of the entire expression, it has to be added by means of lexical rule. The lexical rule says: If a verb is used with an additional predicate and accusative object, then the entire construction has a causative meaning.

Figure 21.13 on page 557 shows how a lexical rule can be replaced by an empty head.The

[3] Alternatively, one could assume a very complex relation that connects ARG-ST and SUBCAT. But this would then have to re-deliver the result of an interaction of a number of phenomena where the interaction of these phenomena would not be captured in a transparent way.

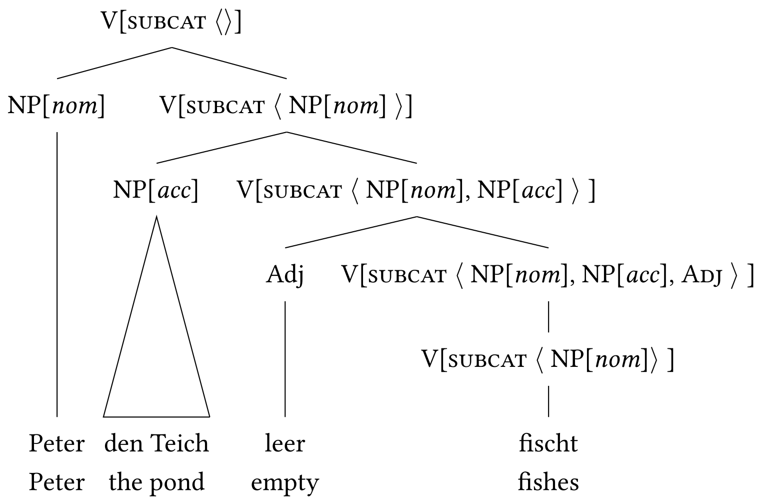


Figure 21.12: Analysis of the resultative construction with a lexical rule

empty head requires the intransitive verb and additionally an adjective, an accusative object and a subject. The subject of *fischt* 'fishes' must of course be identical to the subject that is selected by the combination of *fischt* and the empty head. This is not shown in the figure. It is possible, however, to establish this identity (see Hinrichs & Nakazawa 1994). The causative semantics is contributed by the empty head in this analysis. The trick that is being implemented here is exactly what was done in Section 21.2 just in the opposite direction: In the previous section, binary-branching structures with an empty daughter were replaced by unary-branching structures. In this section, we have replaced unary-branching structures with binary-branching structures with an empty daughter.[4]

We have therefore seen that certain transformations can be replaced by lexical rules and also that lexical rules can be replaced by empty heads. The following chapter deals with the question of whether phenomena like extraction, scrambling, and passive should be described with the same tool as in GB/Minimalism or with different tools as in LFG and HPSG.

[4] Here, we are discussing lexical rules, however this transformation trick can also be applied to other unary rules. Semanticists often use rules for type shifting. For example, a rule that turns a referential NP such as *a trickster* in (i.a) into a predicative one (i.b) (Partee 1987).

(i) a. A trickster laughs.
    b. He is a trickster.

These changes can be achieved by a unary rule that is applied to an NP or with a special empty head that takes an NP as its argument. In current Minimalist approaches, empty heads are used (Ramchand 2005: 370), in Categorial Grammar and HPSG unary-branching rules are more common (Flickinger 2008: 91–92; Müller 2009c; 2012).

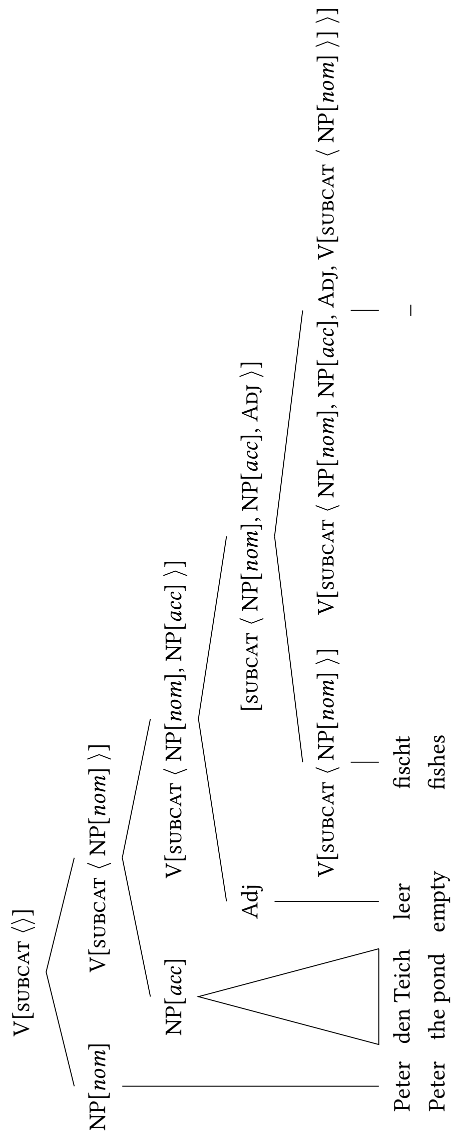


Figure 21.13: Analysis of the resultative construction with an empty head

22 Extraction, scrambling, and passive: one or several descriptive devices?

An anonymous reviewer suggested to discuss one issue in which transformational theories differ from theories like LFG and HPSG. The reviewer claimed that Transformational Grammars use just one tool for the description of active/passive alternations, scrambling, and extraction, while theories like LFG and HPSG use different techniques for all three phenomena. If this claim would be correct and if the analyses would make correct predictions, the respective GB/Minimalism theories would be better than their competitors, since the general aim in science is to develop theories that need a minimal set of assumptions. I already commented on the analysis of passive in GB in Section 3.4, but I want to extend this discussion here and include a Minimalist analysis and one from Dependency Grammar.

The task of any passive analysis is to explain the difference in argument realization in examples like (1):

(1) a. She beats him.
  b. He was beaten.

In these examples about chess the accusative object of *beat* is realized as the nominative in (1b). In addition it can be observed that the position of the elements is different: While *him* is realized postverbally in object position in (1a), it is realized preverbally in (1b). In GB this is explained by a movement analysis. It is assumed that the object does not get case in passive constructions and hence has to move into the subject position where case is assigned by the finite verb. This analysis is also assumed in Minimalist work as for instance in David Adger's textbook (2003). Figure 22.1 on the next page shows his analysis of (2):

(2) Jason was killed.

TP stands for Tense Phrase and corresponds to the IP that was discussed in Chapter 3. PassP is a functional head for passives. *v*P is a special category for the analysis of verb phrases that was originally introduced for the analysis of ditransitives Larson (1988) and VP is the normal VP that consists of verb and object. In Adger's analysis the verb *kill* moves from the verb position in VP into the head position of *v*, the passive auxiliary *be* moves from the head position of PassP to the head position of the Tense Phrase. Features like Infl are 'checked' in connection with such movements. The exact implementation of these checking and valuing operations does not matter here. What is important is that *Jason* moves from the object position to a position that was formerly known as the

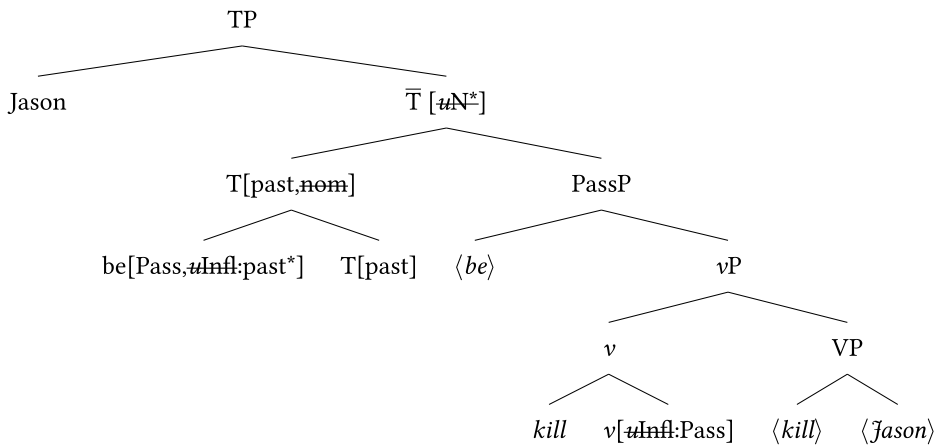


Figure 22.1: Adger's Minimalist movement-based analysis of the passive

specifier position of T (see Footnote 25 on page 158 on the notion of specifier). All these analyses assume that the participle cannot assign accusative to its object and that the object has to move to another position to get case or check features. How the fact that the participle cannot assign case is represented formally was hardly ever made explicit in the GB literature. The following is a list of statements that can be found in the literature:

(3) a. We shall assume that a passivized verb loses the ability to assign structural ACCUSATIVE case to its complement. (Haegeman 1994: 183)

b. das Objekt des Aktivsatzes wird zum Subjekt des Passivsatzes, weil die passivische Verbform keinen Akkusativ-Kasus regieren kann (Akk-Kasus-Absorption). (Lohnstein 2014: 172)

In addition it is sometimes said that the external theta role is absorbed by the verb morphology (Jaeggli 1986; Haegeman 1994: 183). Now, what would it entail if we made this explicit? There is some lexical item for verbs like *beat*. The active form has the ability to assign accusative to its object, but the passive form does not. Since this is a property that is shared by all transitive verbs (by definition of the term transitive verb), this is some regularity that has to be captured. One way to capture this is the assumption of a special passive morpheme that suppresses the agent and changes something in the case specification of the stem it attaches too. How this works in detail was never made explicit. Let us compare this morpheme-based analysis with lexical rule-based analyses: As was explained in Section 21.5, empty heads can be used instead of lexical rules in those cases in which the phonological form of the input and the output does not differ. So for instance lexical rules that license additional arguments as for instance in resultative constructions can be replaced by an empty head. However, as was explained in Section 9.2, lexical rules are also used to model morphology. This is also true for Construction Grammar (see Gert Booij's work on Construction Morphology (2010), which is in many way similar to Riehemann's work in HPSG (1993; 1998)). In the case of the passive lexical rule, the participle morphology is combined with the stem and the subject is suppressed on the respective valency list. This is exactly what is described in the GB/MP literature. The respective lexical rule for the analysis of *ge-lieb-t* 'loved' is depicted in Figure 22.2 to the left. The morpheme-based analysis is shown to the right. To keep things simple, I

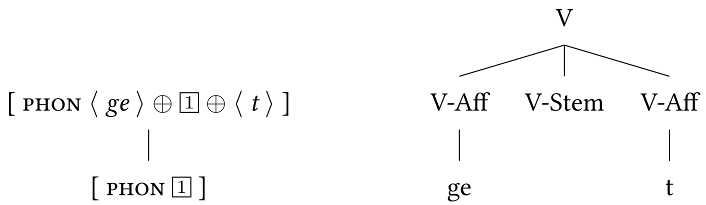


Figure 22.2: Lexical rule-based/constructionist vs. morpheme-based analysis

assume a flat analysis, but those who insist on binary branching structures would have to come up with a way of deciding whether the *ge-* or the *-t* is combined first with the stem and in which way selection and percolation of features takes place. Independent of how morphology is done, the fact has to be represented that the inflected form (the top node in both figures) has different properties than the verb stem. In the morpheme-based world the morpheme is responsible for suppressing the agent and changing the case assignment properties, in the lexical rule/construction world this is done by the respective lexical rule. There is no difference in terms of needed tools and necessary stipulations.

The situation in Minimalist theories is a little bit different. For instance, (Adger 2003: 229, 231) writes the following:

> Passives are akin to unaccusatives, in that they do not assign accusative case to their object, and they do not appear to have a thematic subject. […] Moreover, the idea that the function of this auxiliary is to select an unaccusative little *v*P simultaneously explains the lack of accusative case and the lack of a thematic subject. (Adger 2003: 229, 231)

So this is an explicit statement. The relation between a stem and a passive participle form that was assumed in GB analyses is now a verb stem that is combined with two different versions of little *v*. Which *v* is chosen is determined by the governing head a functional Pass head or a Perf head. This can be depicted as in Figure 22.3 on the next page. When *kill* is used in the perfect or in the passive, it is spelled out as *killed*. If it is used in the active with a 3rd person singular subject it is spelled out as *kills*. This can be compared with a lexical analysis, as for instance the one that is assumed in HPSG. The analysis is shown in Figure 22.4 on the following page. The left figure shows a lexical item that is licensed by a lexical rule that is applied to the stem *kill-*. The stem has two elements in its argument structure list and for the active forms the complete argument structure list

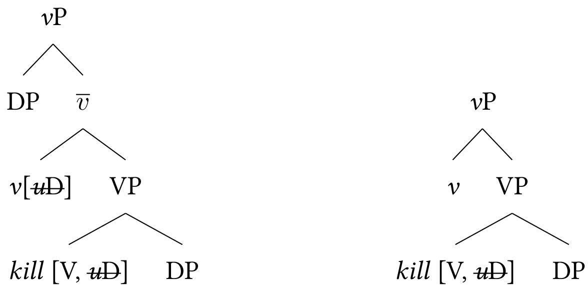


Figure 22.3: Analysis of the passive in a Minimalist theory involving little *v*

V[SPR ⟨ $\boxed{1}$ ⟩, COMPS ⟨ $\boxed{2}$ ⟩, ARG-ST ⟨ $\boxed{1}$ NP[*str*], $\boxed{2}$ NP[*str*] ⟩]	V[SPR ⟨ $\boxed{2}$ ⟩, COMPS ⟨ ⟩, ARG-ST ⟨ $\boxed{2}$ NP[*str*] ⟩]
|	|
V[ARG-ST ⟨ $\boxed{1}$ NP[*str*], $\boxed{2}$ NP[*str*] ⟩]	V[ARG-ST ⟨ $\boxed{1}$ NP[*str*], $\boxed{2}$ NP[*str*] ⟩]

Figure 22.4: Lexical rule-based analysis of the passive in HPSG

is shared between the licenced lexical item and the stem. The first element of the ARG-ST list is mapped to SPR and the other elements to COMPS (in English). Passive is depicted in the right figure: The first element of the ARG-ST with structural case is suppressed and since the element that was the second element in the ARG-ST list of the stem ($\boxed{2}$) is now the first elelement, this item is mapped to SPR. See Section 9.2 for passive in HPSG and Section 9.6 for comments on ARG-ST and the differences between German and English.

The discussion of the Figures 22.3 and 22.4 are a further ilustration of a point made in Section 21.5: Lexical rules can be replaced by empty heads and vice versa. While HPSG says there are stems that are related to inflected forms and corresponding to the inflection the arguments are realized in a certain way, Minimalist theories assume two variants of little *v* that differ in their selection of arguments. Now, the question is: Are there empirical differences between the two approaches? I guess there are differences if one considers the question of language acquisition. What children can acquire from data is that there are various inflected forms and that they are related somehow. What remains questionable is whether they really would be able to detect empty little *v*s. One could claim of course that children operate with chunks of structures such as the ones in Figure 22.3. But then a verb would be just a chunk consisting of little *v* and V and having some open slots. This would be indistinguishable from what the HPSG analysis assumes.

As far as the 'lexical rules as additional tool' aspect is concerned the discussion is closed, but note that the standard GB/Minimalism analyses differ in another way from

LFG and HPSG analyses, since they assume that passive has something to do with movement, that is, that the same mechanisms are used that are used for non-local dependencies.[1] This works for languages like English in which the object has to be realized in postverbal position in the active and in preverbal position in the passive, but it fails for languages like German in which the order of constituents is rather free. Lenerz (1977: Section 4.4.3) discussed the examples in (44) on page 114 – which are repeated here as (4) for convenience:

(4) a. weil das Mädchen dem Jungen den Ball schenkt
because the girl the.DAT boy the.ACC Ball gives
'because the girl gives the ball to the boy'

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

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

While both orders in (4b) and (4c) are possible, the one with dative–nominative order in (4b) is the unmarked one. There is a strong linearization rule in German that demands animate NPs to be serialized before inanimate ones (Hoberg 1981: 46). This linearization rule is unaffected by passivization. Theories that assume that passive is movement either have to assume that the passive of (4a) is (4c) and (4b) is derived from (4c) by a further reordering operation (which would be implausible since usually one assumes that more marked constructions require more transformations) or they would have to come up with other explanations for the fact that the subject of the passive sentence has the same position as the object in active sentences. As was already explained in Section 3.4 one such explanation is to assume an empty expletive subject that is placed in the position to which nominative is assigned and to somehow connect this expletive element to the subject in object position. While this somehow works, it should be clear that the price for rescuing a movement-based analysis of passive is rather high: One has to assume an empty expletive element, that is, something that neither has a form nor a meaning. The existence of such an object could not inferred from the input unless it is assumed that the structures in which it is assumed are given. So a rather rich UG would have to be assumed.

The question to be asked here is: Why does the movement-based analysis have these problems and why does the valency-based analysis does not have them? The cause of the problem is that the analysis of the passive mixes two things: The fact that SVO languages like English encode subjecthood positionally and the fact that the subject is suppressed in passives. If these two things are separated the problem disappears. The fact that

[1] There is another option in Minimalist theories. Since Agree can check features non-locally, T can assign nominative to an embedded element. So, in principle the object may get nominative in the VP without moving to T. However, Adger (2003: 368) assumes that German has a strong EPP feature on T, so that the underlying object has to move to the specifier of T. This is basically the old GB analysis of passive in German with all its conceptual problems and disadvantages.

the object of the active sentence in (1a) is realized as the subject in (1b) is explained by the assumption that the first NP on the argument structure list with structural case is realized as subject and mapped to the respective valency feature: SPR in English. Such mappings can be language specific (see Section 9.6 and Müller (2016) where I discuss Icelandic, which is an SVO language with subjects with lexical case) and the schemata that license head-argument combinations are language- or rather language class-specific as well.

In what follows I discuss another set of examples that are sometimes seen as evidence that makes a movement-based analysis necessary. The examples in (5) are instances of the so-called remote passive (Höhle 1978: 175–176).[2]

(5) a. daß er auch von mir zu überreden versucht wurde.[3]
   that he.NOM also from me to persuade tried got
   'that an attempt to persuade him was also made by me.'
  b. weil der Wagen oft zu reparieren versucht wurde.
   because the car.NOM often to repair tried was
   'because many attempts were made to repair the car.'

What is interesting about these examples is that the subject is the underlying object of a deeply embedded verb. This seems to suggest that the object is extracted out of the verb phrase. So the analysis of (5b) would be (6):

(6) weil [$_{IP}$ der Wagen$_i$ [$_{VP}$ oft [$_{VP}$ [$_{VP}$ [$_{VP}$ _$_i$ zu reparieren] versucht] wurde].
  because the car.NOM often to repair tried was

While this is a straight-forward explanation of the fact that (5b) is grammatical, another explanation is possible as well. In the HPSG analysis of German (and Dutch) it is assumed that verbs like those in (5b) form a verbal complex, that is, *zu reparieren versucht wurde* 'to repair tried was' forms one unit. When two or more verbs form a complex, the highest verb attracts the arguments from the verb it embeds (Hinrichs & Nakazawa 1989b; 1994; Bouma & van Noord 1998). A verb like *versuchen* 'to try' selects a subject, an infinitive with *zu* 'to' and all complements that are selected by this infinitive. In the analysis of (7), *versuchen* 'to try' selects for its subject, the object of *reparieren* 'to repair' and *zu reparieren* 'to repair'.

(7) weil er den Wagen zu reparieren versuchen will
  because he.NOM the.ACC car to repair try wants
  'because he wants to try to repair the car'

Now if the passive lexical rule applies to *versuch-*, it suppresses the first argument of *versuch-* with structural case, which is the subject of *versuch-*. The next argument of *versuch-* is the object of *zu reparieren*. Since this element is the first NP with structural case it gets nominative as in (5b). So, this shows that there is an analysis of the remote

[2] See Müller (2002a: Section 3.1.4.1) and Wurmbrand (2003b) for corpus examples.
[3] Oppenrieder (1991: 212).

passive that does not rely on movement. Since movement-based analyses were shown to be problematic and since there are no data that cannot be explained without movement, analyses without movement have to be preferred.

This leaves us with movement-based accounts of local reordering (scrambling). The reviewer suggested that scrambling, passive, and non-local extraction may be analyzed with the same mechanism. It was long thought that scope facts made the assumption of movement-based analyses of scrambling necessary, but it was pointed out by Kiss (2001: 146) and Fanselow (2001: Section 2.6) that the reverse is true: Movement-based accounts of scrambling make wrong predictions with regard to available quantifier scopings. I discussed the respective examples in Section 3.5 already and will not repeat the discussion here. The conclusion that has to be drawn from this is that passive, scrambling, and long distance extraction are three different phenomena that should be treated differently. The solution for the analysis of the passive that is adopted in HPSG is based on an analysis by Haider (1986a), who worked within the GB framework. The 'scrambling-as-base generation' approach to local reorderings that was used in HPSG right from the beginning (Gunji 1986) is also adopted by some practitioners of GB/Minimalism, e. g. Fanselow (2001).

Having discussed the analyses in GB/Minimalism, I now turn to Dependency Grammar. Groß & Osborne (2009) suggest that *w*-fronting, topicalization, scrambling, extraposition, splitting, and also the remote passive should be analyzed by what they call *rising*. The concept was already explained in Section 11.5. The Figures 22.5 and 22.6 show examples for the fronting and the scrambling of an object. Groß and Osborne assume

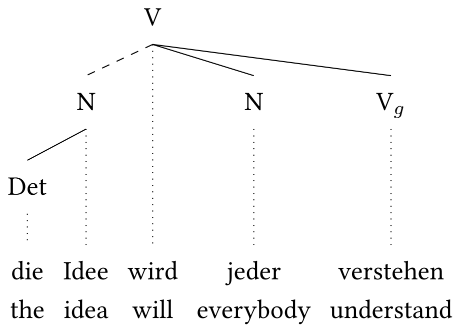


Figure 22.5: Analysis of *Die Idee wird jeder verstehen.* 'Everybody will understand the idea.' involving rising

that the object depends on the main verb in sentences with auxiliary verbs, while the subject depends on the auxiliary. Therefore the object *die Idee* 'the idea' and the object *sich* 'himself' have to rise to the next higher verb in order to keep the structures projective. Figure 22.7 on the next page shows the analysis of the remote passive. The object of *zu reparieren* 'to repair' rises to the auxiliary *wurde* 'was'.

Groß and Osborne use the same mechanism for all these phenomena, but it should be

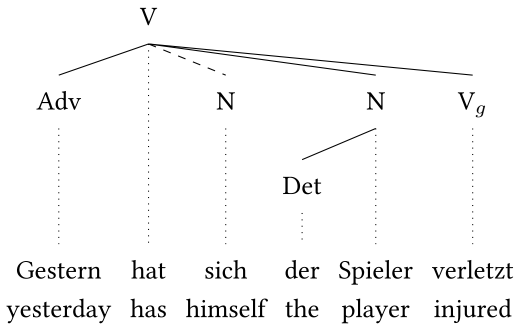


Figure 22.6: Analysis of *Gestern hat sich der Spieler verletzt.* ‘Yesterday, the player injured himself.’ involving rising

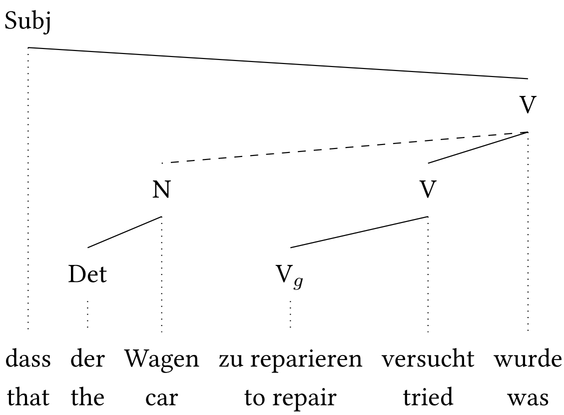


Figure 22.7: Analysis of the remote passive *dass der Wagen zu reparieren versucht wurde* ‘that it was tried to repair the car’ involving rising

clear that there have to be differences on the exact implementation. Groß and Osborne say that English does not have scrambling, while German does. If this is to be captured, there must be a way to distinguish the two phenomena, since if this was not possible one would predict that English has scrambling as well, since both German and English do allow long distance frontings. Groß & Osborne (2009: 58) assume that object nouns that raise must take the nominative. But if the kind of rising that they assume for remote passives is identical to the one that they assume for scrambling, they would predict that *den Wagen* gets nominative in (8) as well:

(8) dass den Wagen niemand repariert hat
that the.ACC car nobody.NOM repaired has
‘that nobody repaired the car’

Since *den Wagen* 'the car' and *repariert* 'repaired' are not adjacent, *den Wagen* has to rise to the next higher head in order to allow for a projective realization of elements. So in order to assign case properly, one has to take the arguments into account that are governed by the head to which a certain element rises. Since the auxiliary *hat* 'has' already governs a nominative, the NP *den Wagen* has to be realized in the accusative. An analysis that assumes that both the accusative and the nominative depend on *hat* 'has' in (8) is basically the verbal complex analysis that is assumed in HPSG and some GB variants.

But note, that this does not extend to nonlocal dependencies. Case is assigned locally by verbs or verbal complexes, but not to elements that come from far away. The long distance extraction of NPs is more common in southern variants of German and there are only a few verbs that do not take a nominative argument themselves. The examples below involve *dünken* 'to think', which governs an accusative and a sentential object and *scheinen* 'to seem', which governs a dative and a sentential object. If (9a) is analyzed with *den Wagen* rising to *dünkt*, one might expect that *den Wagen* 'the car' gets nominative since there is no other element in the nominative. However, (8b) is entirely out.

(9) a. Den Wagen dünkt mich, dass er repariert.
the.ACC car thinks me.ACC that he.NOM repairs
'I think that he repairs the car'

b. * Der Wagen dünkt mich, dass er repariert.
the.NOM car thinks me.ACC that he.NOM repairs

Similarly there is no agreement between the fronted element and the verb to which it attaches:

(10) a. Mir scheint, dass die Wagen ihm gefallen.
me.DAT.1PL seems.3SG that the cars.3PL him please.3PL

b. Die Wagen scheint mir, dass ihm gefallen.
the cars.3PL seem.3SG me.DAT that him please.3PL

c. * Die Wagen scheinen mir, dass ihm gefällt.
the cars.3PL seem.3PL me.DAT that him pleases.3SG

d. * Die Wagen scheinen mir, dass ihm gefallen.
the cars.3PL seem.3PL me.DAT that him please.3PL

This shows that scrambling/remote passive and extraction should not be dealt with by the same mechanism or if they are dealt with by the same mechanism one has to make sure that there are specialized variants of the mechanism that take the differences into account. I think what Groß and Osborne did is recode the attachment relations of phrase structure grammars. *die Idee* 'the idea' has some relation to *wird jeder verstehen* 'will everybody understand' in Figure 22.5, as it does in GB, LFG, GPSG, HPSG, and other similar frameworks. In HPSG *die Idee* 'the idea' is the filler in a filler-head configuration. The remote passive and local reorderings of arguments of auxiliaries, modal verbs, and other verbs that behave similarly are explained by verbal complex formation where all non-verbal arguments depend on the highest verb.

Concluding this chapter it can be said that local reorderings and long distance dependencies are two different things that should be described with different tools (or there should be further constraints that differ for the respective phenomena when the same tool is used). Similarly movement-based analyses of the passive are problematic since passive does not necessarily imply reordering.

23 Phrasal vs. lexical analyses

This section deals with a rather crucial aspect when it comes to the comparison of the theories described in this book: valency and the question whether sentence structure or rather syntactic structure in general is determined by lexical information or whether syntactic structures have an independent live (and meaning) and lexical items are just inserted into them. Roughly speaking, frameworks like GB/Minimalism, LFG, CG, HPSG, and DG are lexical, while GPSG and Construction Grammar (Goldberg 1995; 2003a; Tomasello 2003; 2006b; Croft 2001; Michaelis 2006) are phrasal approaches. This categorization reflects tendencies, but there are non-lexical approaches in Minimalism (Borer's exoskeletal approach, 2003) and LFG (Alsina 1996; Asudeh et al. 2008; 2013) and there are lexical approaches in Construction Grammar (Sign-Based Construction Grammar, see Section 10.6.2). The phrasal approach is wide-spread also in frameworks like Cognitive Grammar (Dąbrowska 2001; Langacker 2009: 169) and Simpler Syntax (Culicover & Jackendoff 2005; Jackendoff 2008) that could not be discussed in this book.

The question is whether the meaning of an utterance like (1a) is contributed by the verb *give* and the structure is needed for the NPs around the verb does not contribute any meaning or whether there is a phrasal pattern X Verb Y Z that contributes some 'ditransitive meaning' whatever this is.[1]

(1) a. Peter gives Mary the book.
    b. Peter fishes the pond empty.

Similarly the question is how the constituents in (1b) are licensed. This sentence is interesting since it has a resultative meaning that is not part of the meaning of the verb *fish*: Peter's fishing causes the pond to become empty. Nor is this additional meaning part of the meaning of any other item in the sentence. On the lexical account there is a lexical rule that licenses a lexical item that selects for *Peter*, *the pond*, and *empty*. This lexical item also contributes the resultative meaning. On the phrasal approach it is assumed that there is a pattern Subj V Obj Obl. This pattern contributes the resultative meaning, while the verb that is inserted into this pattern just contributes the meaning it would

[1] Note that the prototypical meaning is a transfer of possession in which Y receives Z from X, but the reverse holds in (i.b):

(i) a. Er gibt ihr den Ball.
       he.NOM gives her.DAT the.ACC ball
    b. Er stiehlt ihr den Ball.
       he.NOM steals her.DAT the ball
       'He steals the ball from her.'

have outside this construction, e. g. the meaning that *fish* would have in an intransitive construction. I call such phrasal approaches *plugging approaches*, since lexical items are plugged into ready-made structures that do most of the work.

In what follows I will examine these proposals in more detail and argue that the lexical approaches to valency are the correct ones. The discussion will be based on earlier work of mine (Müller 2006; 2007c; 2010b) and work that I did together with Steve Wechsler (Müller & Wechsler 2014a;b). Some of the sections in Müller & Wechsler (2014a) started out as translations of Müller (2013b), but the material was reorganized and refocused due to intense discussions with Steve Wechsler. So rather than using a translation of Section 11.11 of Müller (2013b), I use parts of Müller & Wechsler (2014a) here and add some subsections that have to be left out of the article due to space restrictions (Subsections 23.3.6 and 23.7.3). Because there have been misunderstandings in the past (e. g. Boas (2014), see Müller & Wechsler (2014b)), a disclaimer is necessary here. This section is not an argument against Construction Grammar. As was mentioned above Sign-Based Construction Grammar is a lexical variant of Construction Grammar and hence compatible with what I believe to be correct. This section is also not against phrasal constructions in general, since there are phenomena that seem to be best captured with phrasal constructions. These are discussed in detail in Subsection 23.10. What is argued against in the following subsections is a special kind of phrasal constructions, namely phrasal argument structure constructions. I believe that all phenomena that have to do with valence and valence alternations should be treated lexically.

23.1 Some putative advantages of phrasal models

In this section we examine certain claims to purported advantages of phrasal versions of Construction Grammar over lexical rules. Then in the following section we turn to positive arguments for lexical rules.

23.1.1 Usage-based theories

For many practitioners of Construction Grammar, their approach to syntax is deeply rooted in the ontological strictures of *usage-based* theories of language (Langacker 1987; Goldberg 1995; Croft 2001; Tomasello 2003). Usage-based theorists oppose the notion of 'linguistic rules conceived of as algebraic procedures for combining symbols that do not themselves contribute to meaning' (Tomasello 2003: 99). All linguistic entities are symbolic of things in the realm of denotations; 'all have communicative significance because they all derive directly from language use' (*ibid*). Although the formatives of language may be rather abstract, they can never be divorced from their functional origin as a tool of communication. The usage-based view of constructions is summed up well in the following quote:

> The most important point is that constructions are nothing more or less than patterns of usage, which may therefore become relatively abstract if these patterns

> include many different kinds of specific linguistic symbols. But never are they empty rules devoid of semantic content or communicative function. (Tomasello 2003: 100)

Thus constructions are said to differ from grammatical rules in two ways: they must carry meaning; and they reflect the actual 'patterns of usage' fairly directly.

Consider first the constraint that every element of the grammar must carry meaning, which we call the *semiotic dictum*. Do lexical or phrasal theories hew the most closely to this dictum? Categorial Grammar, the paradigm of a lexical theory (see Chapter 8), is a strong contender: it consists of meaningful words, with only a few very general combinatorial rules such as X/Y * Y = X. Given the rule-to-rule assumption those combinatorial rules specify the meaning of the whole as a function of the parts. Whether such a rule counts as meaningful in itself in Tomasello's sense is not clear.

What does seem clear is that the combinatorial rules of Construction Grammar, such as Goldberg's Correspondence Principle for combining a verb with a construction (1995: 50), have the same status as those combinatorial rules:

(2) The Correspondence Principle: Each participant that is lexically profiled and expressed must be fused with a profiled argument role of the construction. If a verb has three profiled participant roles, then one of them may be fused with a non-profiled argument role of a construction. (Goldberg 1995: 50)

Both verbs and constructions are specified for participant roles, some of which are *profiled*. Argument profiling for verbs is 'lexically determined and highly conventionalized' (Goldberg 1995: 46). Profiled argument roles of a construction are mapped to direct grammatical functions, i. e., SUBJ, OBJ, or OBJ2. By the Correspondence Principle the lexically profiled argument roles must be direct, unless there are three of them, in which case one may be indirect.[2] With respect to the semiotic dictum, the Correspondence Principle has the same status as the Categorial Grammar combinatorial rules: a meaningless algebraic rule that specifies the way to combine meaningful items.

Turning now to the lexicalist syntax we favor, some elements abide by the semiotic dictum while others do not. Phrase structure rules for intransitive and transitive VPs (or the respective HPSG ID schema) do not. Lexical valence structures clearly carry meaning since they are associated with particular verbs. In an English ditransitive, the first object expresses the role of 'intended recipient' of the referent of the second object. Hence *He carved her a toy* entails that he carved a toy with the intention that she receive it. So the lexical rule that adds a benefactive recipient argument to a verb adds meaning. Alternatively, a phrasal ditransitive construction might contribute that 'recipient' meaning.[3] Which structures have meaning is an empirical question for us.

[2] We assume that the second sentence of (2) provides for exceptions to the first sentence.

[3] In Section 23.2.1 we argue that the recipient should be added in the lexical argument structure, not through a phrasal construction. See Wechsler (1991: 111–113; 1995: 88–89) for an analysis of English ditransitives with elements of both constructional and lexical approaches. It is based on Kiparsky's notion of a *thematically restricted positional linker* (1987; 1988).

In contrast, in Construction Grammar meaning is assumed *a priori* for all constructions. But while the ditransitive construction plausibly contributes meaning, no truth-conditional meaning has yet been discovered for either the intransitive or bi-valent transitive constructions. Clearly the constructionist's evidence for the meaningfulness of *certain* constructions such as the ditransitive does not constitute evidence that *all* phrasal constructions have meaning. So the lexical and phrasal approaches seem to come out the same, as far as the semiotic dictum is concerned.

Now consider the second usage-based dictum, that the elements of the grammar directly reflect patterns of usage, which we call *the transparency dictum*. The Construction Grammar literature often presents their constructions informally in ways that suggest that they represent surface constituent order patterns: the transitive construction is 'X VERB Y' (Tomasello) or 'Subj V Obj' (Goldberg 1995; 2006)[4]; the passive construction is 'X *was* VERB*ed by* Y' (Tomasello 2003: 100) or 'Subj aux Vpp (PPby)' (Goldberg 2006: 5). But a theory in which constructions consist of surface patterns was considered in detail and rejected by (Müller, 2006: Section 2), and does not accurately reflect Goldberg's actual theory.[5] The more detailed discussions present *argument structure constructions*, which are more abstract and rather like the lexicalists' grammatical elements (or perhaps an LFG f-structure): the transitive construction resembles a transitive valence structure (minus the verb itself); the passive construction resembles the passive lexical rule.

With respect to fulfilling the desiderata of usage-based theorists, we do not find any significant difference between the non-lexical and lexical approaches.

23.1.2 Coercion

Researchers working with plugging proposals usually take coercion as showing the usefulness of phrasal constructions. For instance, Anatol Stefanowitsch (Lecture in the lecture series *Algorithmen und Muster –- Strukturen in der Sprache*, 2009) discussed the example in (3):

(3) Das Tor zur Welt Hrnglb öffnete sich ohne Vorwarnung und verschlang [sie] … die Welt Hrnglb wird von Magiern erschaffen, die Träume zu Realität formen können, aber nicht in der Lage sind zu träumen. Haltet aus, Freunde. Und ihr da draußen, bitte träumt ihnen ein Tor.[6]

[4] Goldberg et al. (2004: 300) report about a language acquisition experiment that involves an SOV pattern. The SOV order is mentioned explicitly and seen as part of the construction.

[5] This applies to argument structure constructions only. In some of her papers Goldberg assumes that very specific phrase structural configurations are part of the constructions. For instance in her paper on complex predicates in Persian (Goldberg 2003b) she assigns V^0 and $\overline{V}$ categories. See Müller (2010b: Section 4.9) for a critique of that analysis.

[6] http://www.elbenwaldforum.de/showflat.php?Cat=&Board=Tolkiens_Werke& Number=1457418&page=3&view=collapsed&sb=5&o=&fpart=16. 27.02.2010.
'The gate to the world Hrnglb opened without warning and swallowed them. The world Hrnglb is created by magicians that can form reality from dreams but cannot dream themselves. Hold out, friends! And you out there, please, dream a gate for them.'

The crucial part is *bitte träumt ihnen ein Tor* 'Dream a gate for them'. In this fantasy context the word *träumen*, which is intransitive, is forced into the ditransitive construction and therefore gets a certain meaning. This forcing of a verb corresponds to overwriting or rather extending properties of the verb by the phrasal construction.

In cases in which the plugging proposals assume that information is over-written or extended, lexical approaches assume mediating lexical rules. Briscoe & Copestake (1999: Section 4) have worked out a lexical approach in detail.[7] They discuss the ditransitive sentences in (4), which either correspond to the prototypical ditransitive construction (4a) or deviate from it in various ways.

(4) a. Mary gave Joe a present.
  b. Joe painted Sally a picture.
  c. Mary promised Joe a new car.
  d. He tipped Bill two pounds.
  e. The medicine brought him relief.
  f. The music lent the party a festive air.
  g. Jo gave Bob a punch.
  h. He blew his wife a kiss.
  i. She smiled herself an upgrade.

For the non-canonical examples they assume lexical rules that relate transitive (*paint*) and intransitive (*smile*) verbs to ditransitive ones and contribute the respective semantic information or the respective metaphorical extension. The example in (4i) is rather similar to the *träumen* example discussed above and is also analyzed with a lexical rule (page 509). Briscoe and Copestake note that this lexical rule is much more restricted in productivity than other lexical rules that were suggested by them. They take this as motivation for developing a representational format in which lexical items (including those that are derived by lexical rules) are associated with probabilities, so that differences in productivity of various patterns can be captured.

Looking narrowly at such cases, it is hard to see any rational grounds for choosing between the phrasal analysis and the lexical rule. But if we broaden our view, the lexical rule approach can be seen to have much wider application. Coercion is a very general pragmatic process, occurring in many contexts where no construction seems to be responsible (Nunberg 1995). Nunberg cites many cases such as the restaurant waiter asking *Who is the ham sandwich?* (Nunberg 1995: 115). Copestake & Briscoe (1992: 116) discuss the conversion of terms for animals to mass nouns (see also Copestake & Briscoe (1995: 36–43)). Example (5) is about a substance, not about a cute bunny.

(5) After several lorries had run over the body, there was rabbit splattered all over the road.

[7] Kay (2005), working in the framework of CxG, also suggests unary constructions.

The authors suggest a lexical rule that maps a count noun onto a mass noun. This analysis is also assumed by Fillmore (1999: 114–115). Such coercion can occur without any syntactic context: one can answer the question *What's that stuff on the road?* or *What are you eating?* with the one-word utterance *Rabbit.* Some coercion happens to affect the complement structure of a verb, but this is simply a special case of a more general phenomenon that has been analyzed by rules of systematic polysemy.

23.1.3 Aspect as a clause level phenomenon

Alsina (1996), working in the framework of LFG, argues for a phrasal analysis of resultative constructions based on the aspectual properties of sentences, since aspect is normally viewed as a property that is determined by the sentence syntax. Intransitive verbs such as *bark* refer to activities, a resultative construction with the same verb, however, stands for an accomplishment (an extended change of state). Alsina supports this with the following data:

(6) a. (*) The dog barked in five minutes.
b. The dog barked the neighbors awake in five minutes.

The latter sentence means that the *barking* event was completed after five minutes. A reading referring to the time span of the event is not available for (6a). If (6a) is grammatical at all, then a claim is being made about the time frame in which the event begun.

If we now consider examples such as (21b), however, we see that Alsina's argumentation is not water-tight since the resultative meaning is already present at the word-level in nominalizations. As the examples in (7) show, this contrast can be observed in nominal constructions and is therefore independent of the sentence syntax:

(7) a. weil sie die Nordsee in fünf Jahren leer fischten
because they the North.Sea in five years empty fished
'because they fished the North Sea (until it was) empty in five years'
b. # weil sie in fünf Jahren fischten
because they in five years fished
c. das Leerfischen der Nordsee in fünf Jahren
the empty.fishing of.the North.Sea in five years
d. # das Fischen in fünf Jahren
the fishing in five years

23.1.4 Simplicity and polysemy

Much of the intuitive appeal of the plugging approach stems from its apparent simplicity relative to the use of lexical rules. But the claim to greater simplicity for Construction Grammar is based on misunderstandings of both lexical rules and Construction Grammar (specifically of Goldberg's (1995, 2006) version). It draws the distinction in the wrong place and misses the real differences between these approaches. This argument from simplicity is often repeated and so it is important to understand why it is incorrect.

Tomasello (2003) presents the argument as follows. Discussing first the lexical rules approach, Tomasello (2003: 160) writes that

> One implication of this view is that a verb must have listed in the lexicon a different meaning for virtually every different construction in which it participates […]. For example, while the prototypical meaning of *cough* involves only one participant, the cougher, we may say such things as *He coughed her his cold*, in which there are three core participants. In the lexical rules approach, in order to produce this utterance the child's lexicon must have as an entry a ditransitive meaning for the verb *cough*. (Tomasello 2003: 160)

Tomasello (2003: 160) then contrasts a Construction Grammar approach, citing Fillmore et al. (1988), Goldberg (1995), and Croft (2001). He concludes as follows:

> The main point is that if we grant that constructions may have meaning of their own, in relative independence of the lexical items involved, then we do not need to populate the lexicon with all kinds of implausible meanings for each of the verbs we use in everyday life. The construction grammar approach in which constructions have meanings is therefore both much simpler and much more plausible than the lexical rules approach. (Tomasello 2003: 161)

This reflects a misunderstanding of lexical rules, as they are normally understood. There is no implausible sense populating the lexicon. The lexical rule approach to *He coughed her his cold* states that when the word *coughed* appears with two objects, the whole complex has a certain meaning. See Müller (2006: 876). Furthermore we explicitly distinguish between listed elements (lexical entries) and derived ones. The general term subsuming both is *lexical item*.

The simplicity argument also relies on a misunderstanding of a theory Tomasello advocates, namely the theory due to Goldberg (1995; 2006). For his argument to go through, Tomasello must tacitly assume that verbs can combine freely with constructions, that is, that the grammar does not place extrinsic constraints on such combinations. If it is necessary to also stipulate which verbs can appear in which constructions then the claim to greater simplicity collapses: each variant lexical item with its 'implausible meaning' under the lexical rule approach corresponds to a verb-plus-construction combination under the phrasal approach.

Passages such as the following may suggest that verbs and constructions are assumed to combine freely:[8]

> Constructions are combined freely to form actual expressions as long as they can be construed as not being in conflict (invoking the notion of construal is intended to allow for processes of accommodation or coercion). (Goldberg 2006: 22)

[8] The context of these quotes makes clear that the verb and the argument structure construction are considered constructions. See Goldberg (2006: 21, ex. (2)).

> Allowing constructions to combine freely as long as there are no conflicts, allows for the infinitely creative potential of language. […] That is, a speaker is free to creatively combine constructions as long as constructions exist in the language that can be combined suitably to categorize the target message, given that there is no conflict among the constructions. (Goldberg 2006: 22)

But in fact Goldberg does not assume free combination, but rather that a verb is 'conventionally associated with a construction' (Goldberg 1995: 50): verbs specify their participant roles and which of those are obligatory direct arguments (*profiled*, in Goldberg's terminology). In fact Goldberg herself (2006: 211) argues against Borer's 2003 putative assumption of free combination on the grounds that Borer is unable to account for the difference between *dine* (intransitive), *eat* (optionally transitive), and *devour* (obligatorily transitive).[9] Despite Tomasello's comment above, Construction Grammar is no simpler than the lexical rules.

The resultative construction is often used to illustrate the simplicity argument. For example, Goldberg (1995: Chapter 7) assumes that the same lexical item for the verb *sneeze* is used in (8a) and (8b). It is simply inserted into different constructions:

(8) a. He sneezed.
    b. He sneezed the napkin off the table.

The meaning of (8a) corresponds more or less to the verb meaning, since the verb is used in the Intransitive Construction. But the Caused-Motion Construction in (8b) contributes additional semantic information concerning the causation and movement: His sneezing caused the napkin to move off the table. *sneeze* is plugged into the Caused Motion Construction, which licenses the subject of *sneeze* and additionally provides two slots: one for the theme (*napkin*) and one for the goal (*off the table*). The lexical approach is essentially parallel, except that the lexical rule can feed further lexical processes like passivization (*The napkin was sneezed off the table*), and conversion to nouns or adjectives (see Sections 23.2.2 and 23.6).

In a nuanced comparison of the two approaches, Goldberg (1995: 139–140) considers again the added recipient argument in *Mary kicked Joe the ball*, where *kick* is lexically a 2-place verb. She notes that on the constructional view, 'the composite fused structure involving both verb and construction is stored in memory'. The verb itself retains its original meaning as a 2-place verb, so that 'we avoid implausible verb senses such as "to cause to receive by kicking".' The idea seems to be that the lexical approach, in contrast, must countenance such implausible verb senses since a lexical rule adds a third argument.

But the lexical and constructional approaches are actually indistinguishable on this point. The lexical rule does not produce a verb with the 'implausible sense' in (9a). Instead it produces the sense in (9b):

[9] Goldberg's critique cites a 2001 presentation by Borer with the same title as Borer (2003). See Section 23.3.4 for more discussion of this issue. As far as we know, the *dine* / *eat* / *devour* minimal triplet originally came from Dowty (1989: 89–90).

(9) a. cause-to-receive-by-kicking(x, y, z)
b. cause(kick(x, y),receive(z,y))

The same sort of 'composite fused structure' is assumed under either view. With respect to the semantic structure, the number and plausibility of senses, and the polyadicity of the semantic relations, the two theories are identical. They mainly differ in the way this representation fits into the larger theory of syntax. They also differ in another respect: on the lexical theory, the derived three-argument valence structure is associated with the phonological string *kicked*. Next we present evidence for that claim.

23.2 Evidence for lexical approaches

23.2.1 Valence and coordination

On the lexical account, the verb *paint* in (4b), for example, is lexically a 2-argument verb, while the unary branching node immediately dominating it is effectively a 3-argument verb. On the constructional view there is no such predicate seeking three arguments that dominates only the verb. Coordination provides evidence for the lexical account.

A generalization about coordination is that two constituents have compatible syntactic properties can be coordinated and that the result of the coordination is an object that has the syntactic properties of each of the conjuncts. This is reflected by the Categorial Grammar analysis which assumes the category (X\X)/X for the conjunction: The conjunction takes an X to the right, an X to the left and the result is an X.

For example, in (10a) we have a case of the coordination of two lexical verbs. The coordination *know and like* behaves like the coordinated simplex verbs: It takes a subject and an object. Similarly, two sentences with a missing object are coordinated in (10b) and the result is a sentence with a missing object.

(10) a. I know and like this record.
b. Bagels, I like and Ellison hates.

The German examples in (11) show that the case requirement of the involved verbs has to be observed. In (11b,c) the coordinated verbs require accusative and dative respectively and since the case requirements are incompatible with unambiguously case marked nouns both of these examples are out.

(11) a. Ich kenne und unterstütze diesen Mann.
I know and support this man.ACC
b. * Ich kenne und helfe diesen Mann.
I know and help this man.ACC
c. * Ich kenne und helfe diesem Mann.
I know and help this man.DAT

Interestingly, it is possible to coordinate basic ditransitive verbs with verbs that have additional arguments licensed by the lexical rule. (12) provides examples in English and German ((12b) is quoted from Müller (2013b: 420)):

(12) a. She then offered and made me a wonderful espresso — nice.[10]

b. ich hab ihr jetzt diese Ladung Muffins mit den Herzchen drauf
I have her now this load Muffins with the little.heart there.on
gebacken und gegeben.[11]
baked and given
'I have now baked and given her this load of Muffins with the little heart on top.'

These sentences show that both verbs are 3-argument verbs at the V^0 level, since they involve V^0 coordination:

(13) [$_{V^0}$ offered and made] [$_{NP}$ me] [$_{NP}$ a wonderful espresso]

This is expected under the lexical rule analysis but not the non-lexical constructional one.[12]

Summarizing the coordination argument: Coordinated verbs generally must have compatible syntactic properties like valence properties. This means that in (12b), for example, *gebacken* 'baked' and *gegeben* 'given' have the same valence properties. In the lexical approach the creation verb *gebacken*, together with a lexical rule, licenses a ditransitive verb. So it can be coordinated with *gegeben*. In the phrasal approach however, the verb *gebacken* has two argument roles and is not compatible with the verb *gegeben*, which has three argument roles. In the phrasal model, *gebacken* can only realize three arguments when it enters the ditransitive phrasal construction or argument structure construction. But in sentences like (12) it is not *gebacken* alone that enters the phrasal syntax, but rather the combination of *gebacken* and *gegeben*. On that view the verbs are incompatible as far as the semantic roles are concerned.

To fix this under the phrasal approach, one could posit a mechanism such that the semantic roles that are required for the coordinate phrase *baked and given* are shared by each of its conjunct verbs and that they are therefore compatible. But this would amount to saying that there are several verb senses for *baked*, something that the anti-lexicalists claim to avoid, as discussed in the next section.

[10] http://www.thespinroom.com.au/?p=102 07.07.2012

[11] http://www.musiker-board.de/diverses-ot/35977-die-liebe-637-print.html. 08.06.2012

[12] One might wonder whether these sentences could be instances of Right Node Raising (RNR) out of coordinated VPs (Bresnan 1974; Abbott 1976):

(i) She [offered ___] and [made me ___] a wonderful espresso.

But this cannot be right. Under such an analysis the first verb has been used without a benefactive or recipient object. But *me* is interpreted as the recipient of both the offering and making. Secondly, the second object can be an unstressed pronoun (*She offered and made me it*, which is not possible in RNR. Note that *offered and made* cannot be a pseudo-coordination meaning 'offered to make'. This is possible only with stem forms of certain verbs such as *try*.

A reviewer correctly observes that a version of the ASC approach could work in the exactly same way as our lexical analysis. Our ditransitive lexical rule would simply be rechristened as a 'ditransitive ASC'. This construction would combine with *baked*, thus adding the third argument, prior to its coordination with *gave*. As long as the ASC approach is a non-distinct notational variant of the lexical rule approach then of course it works in exactly the same way. But the literature on the ASC approach represents it as a radical alternative to lexical rules, in which constructions are combined through inheritance hierarchies, instead of allowing lexical rules to alter the argument structure of a verb prior to its syntactic combination with the other words and phrases.

The reviewer also remarked that examples like (14) show that the benefactive argument has to be introduced on the phrasal level.

(14) I designed and built him a house.

Both *designed* and *built* are bivalent verbs and *him* is the benefactive that extends both *designed* and *built*. However, we assume that sentences like (14) can be analyzed as coordination of two verbal items that are licensed by the lexical rule that introduces the benefactive argument. That is, the benefactive is introduced before the coordination.

The coordination facts illustrate a more general point. The output of a lexical rule such as the one that would apply in the analysis of *gebacken* in (12b) is just a word (an X^0), so it has the same syntactic distribution as an underived word with the same category and valence feature. This important generalization follows from the lexical account while on the phrasal view it is at best mysterious. The point can be shown with any of the lexical rules that the anti-lexicalists are so keen to eliminate in favor of phrasal constructions. For example, active and passive verbs can be coordinated, as long as they have the same valence properties, as in this Swedish example:

(15) Vi beställde och serverade-s en bra Cheers chowder till att börja med, och
we ordered and served-PASS a good Cheers chowder to INF start with and
sedan en stor hummer varje.[13]
then a big lobster each
'We ordered and were served a good Cheers chowder to start with, and then a big lobster each.'

(English works the same way, as shown by the grammatical translation line.) The passive of the ditransitive verb *servera* 'serve' retains one object, so it is effectively transitive and can be coordinated with the active transitive *beställa* 'order'.

Moreover, the English passive verb form, being a participle, can feed a second lexical rule deriving adjectives from verbs. All categories of English participles can be converted to adjectives (Bresnan, 1982b, 2001: Chapter 3):

(16) a. active present participles (cp. The leaf is falling): *the falling leaf*
b. active past participles (cp. The leaf has fallen): *the fallen leaf*

[13] http://www.tripadvisor.se/ShowUserReviews-g40990-d412862-r130540331-Fox_s_Lobster_House-York_Beach_York_Maine.html. 16.07.2012

c. passive participles (cp. The toy is being broken (by the child).): *the broken toy*

That the derived forms are adjectives, not verbs, is shown by a host of properties, including negative *un-* prefixation: *unbroken* means 'not broken', just as *unkind* means 'not kind', while the *un-* appearing on verbs indicates, not negation, but action reversal, as in *untie* (Bresnan, 1982b: 21, 2001: Chapter 3). Predicate adjectives preserve the subject of predication of the verb and for prenominal adjectives the rule is simply that the role that would be assigned to the subject goes to the modified noun instead (*The toy remained (un-)broken.*; *the broken toy*). Being an A^0, such a form can be coordinated with another A^0, as in the following:

(17) a. The suspect should be considered [armed and dangerous].
b. any [old, rotting, or broken] toys

In (17b), three adjectives are coordinated, one underived (*old*), one derived from a present participle (*rotting*), and one from a passive participle (*broken*). Such coordination is completely mundane on a lexical theory. Each A^0 conjunct has a valence feature (in HPSG it would be the SPR feature for predicates or the MOD feature for the prenominal modifiers), which is shared with the mother node of the coordinate structure. But the point of the phrasal (or ASC) theory is to deny that words have such valence features.

The claim that lexical derivation of valence structure is distinct from phrasal combination is further supported with evidence from deverbal nominalization (Wechsler 2008a). To derive nouns from verbs, *-ing* suffixation productively applies to all declinable verbs (*the shooting of the prisoner*), while morphological productivity is severely limited for various other suffixes such as *-(a)tion* (**the shootation of the prisoner*). So forms such as *destruction* and *distribution* must be retrieved from memory while *-ing* nouns such as *looting* or *growing* could be (and in the case of rare verbs or neologisms, must be) derived from the verb or the root through the application of a rule (Zucchi 1993). This difference explains why *ing*-nominals always retain the argument structure of the cognate verb, while other forms show some variation. A famous example is the lack of the agent argument for the noun *growth* versus its retention by the noun *growing*: **John's growth of tomatoes* versus *John's growing of tomatoes* (Chomsky 1970).[14]

But what sort of rule derives the *-ing* nouns, a lexical rule or a phrasal one? On Marantz's (1997) phrasal analysis, a phrasal construction (notated as *vP*) is responsible for assigning the agent role of *-ing* nouns such as *growing*. For him, none of the words directly selects an agent via its argument structure. The *-ing* forms are permitted to appear in the *vP* construction, which licenses the possessive agent. Non-*ing* nouns such as *destruction* and *growth* do not appear in *vP*. Whether they allow expression of the agent depends on semantic and pragmatic properties of the word: *destruction* involves external causation so it does allow an agent, while *growth* involves internal causation so it does not allow an agent.

However, a problem for Marantz is that these two types of nouns can coordinate and share dependents (example (18a) is from Wechsler (2008a: Section 7)):

[14] See Section 23.3.3 for further discussion.

(18) a. With nothing left after the soldier's [destruction and looting] of their home, they reboarded their coach and set out for the port of Calais.[15]

b. The [cultivation, growing or distribution] of medical marijuana within the County shall at all times occur within a secure, locked, and fully enclosed structure, including a ceiling, roof or top, and shall meet the following requirements.[16]

On the phrasal analysis, the nouns *looting* and *growing* occur in one type of syntactic environment (namely *vP*), while forms *destruction*, *cultivation*, and *distribution* occur in a different syntactic environment. This places contradictory demands on the structure of coordinations like those in (18). As far as we know, neither this problem nor the others raised by Wechsler (2008a) have even been addressed by advocates of the phrasal theory of argument structure.

Consider one last example. In an influential phrasal analysis, Hale and Keyser (1993) derived denominal verbs like *to saddle* through noun incorporation out of a structure akin to [PUT a saddle ON x]. Again, verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

(19) Realizing the dire results of such a capture and that he was the only one to prevent it, he quickly [saddled and mounted] his trusted horse and with a grim determination began a journey that would become legendary.[17]

As in all of these X^0 coordination cases, under the phrasal analysis the two verbs place contradictory demands on a single phrase structure.

A lexical valence structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that valence structure is simply to indicate what sort of phrases the verb must (or can) combine with, and the result of semantic composition; if that were the whole story then the phrasal theory would be viable. But it is not. As it turns out, this lexical valence structure, once abstracted, can alternatively be used in other ways: among other possibilities, the verb (crucially including its valence structure) can be coordinated with other verbs that have a similar valence structure; or it can serve as the input to lexical rules specifying a new word bearing a systematic relation to the input word. The coordination and lexical derivation facts follow from the lexical view, while the phrasal theory at best leaves these facts as mysterious and at worst leads to irreconcilable contradictions for the phrase structure.

23.2.2 Valence and derivational morphology

Goldberg & Jackendoff (2004), Alsina (1996), and Asudeh, Dalrymple & Toivonen (2008; 2013) suggest analyzing resultative constructions and/or caused motion constructions

[15] http://www.amazon.com/review/R3IG4M3Q6YYNFT, 21.07.2012

[16] http://www.scribd.com/doc/64013640/Tulare-County-medical-cannabis-cultivation-ordinance#page=1, 22.10.2012

[17] http://www.jouetthouse.org/index.php?option=com_content&view=article&id=56&Itemid=63, 21.07.2012

as phrasal constructions.[18] As was argued in Müller (2006) this is incompatible with the assumption of lexical integrity, that is, that word formation happens before syntax (Bresnan & Mchombo 1995).[19] Let us consider a concrete example, such as (20):

(20) a. Er tanzt die Schuhe blutig / in Stücke.
he dances the shoes bloody into pieces

b. die in Stücke / blutig getanzten Schuhe
the into pieces bloody danced shoes

c. * die getanzten Schuhe
the danced shoes

The shoes are not a semantic argument of *tanzt*. Nevertheless the referent of the NP that is realized as accusative NP in (20a) is the element the adjectival participle in (20b) predicates over. Adjectival participles like the one in (20b) are derived from a passive participle of a verb that governs an accusative object. If the accusative object is licensed phrasally by configurations like the one in (20a) it cannot be explained why the participle *getanzte* can be formed despite the absence of an accusative object. See Müller (2006: Section 5) for further examples of the interaction of resultatives and morphology. The conclusion, which was drawn in the late 70s and early 80s by Dowty (1978: 412) and Bresnan (1982b: 21), is that phenomena that feed morphology should be treated lexically. The natural analysis in frameworks like HPSG, CG, CxG, and LFG is therefore one that assumes a lexical rule for the licensing of resultative constructions. See Verspoor 1997, Wechsler 1997, Wechsler & Noh 2001, Wunderlich 1992: 45; 1997: 120–126, Kaufmann & Wunderlich (1998), Müller (2002a: Chapter 5), Kay (2005), and Simpson (1983) for lexical proposals in some of these frameworks.

This argument is similar to the one that was discussed in connection with the GPSG representation of valence in Section 5.5: Morphological processes have to be able to see the valence of the element they attach to. This is not the case if arguments are introduced by phrasal configurations after the morphology level.

Ashudeh, Dalrymple, and Toivonen's papers are about the concept of lexical integrity and about constructions. Asudeh & Toivonen (2014) replied to our target article and pointed out (again) that their template approach makes it possible to specify the functional structure of words and phrases alike. In the original paper they discussed the Swedish word *vägen*, which is the definite form of *väg* 'way'. They showed that the f-

[18] Asudeh & Toivonen (2014: Section 2.3) argue that their account is not constructional. If a construction is a form-meaning pair, their account is constructional, since a certain c-structure is paired with a semantic contribution. Asudeh and Toivonen compare their approach with approaches in Constructional HPSG (Sag 1997) and Sign-Based Construction Grammar (see Section 10.6.2), which they term constructional. The only difference between these approaches and the approach by Ashudeh, Dalrymple, and Toivonen is that the constructions in the HPSG-based theories are modeled using types and hence have a name.

[19] Asudeh et al. (2013: 14) claim that the Swedish Directed Motion Construction does not interact with derivational morphology. However, the parallel German construction does interact with derivational morphology. The absence of this interaction in Swedish can be explained by other factors of Swedish grammar and given this I believe it to be more appropriate to assume an analysis that captures both the German and the Swedish data in the same way.

structure is parallel to the f-structure for the English phrase *the way*. In our reply (2014b) we gave in too early, I believe. Since the point is not about being able to provide the f-structure of words, the point is about morphology, that is – in LFG terms – about deriving the f-structure by a morphological analysis. More generally speaking, one wants to derive all properties of the involved words, that is, their valency, their meaning, and the linking of this meaning to their dependents. What we used in our argument was parallel to what Bresnan used in her classical argument for a lexical treatment of passive. So either Bresnan's argument (and ours) is invalid or both arguments are valid and there is a problem for Ashudeh, Dalrymple, and Toivonen's approach and for phrasal approaches in general. I want to give another example, that was already discussed in Müller (2006) but was omitted in Müller & Wechsler (2014a) due to space limitations. I will first point out why this example is problematic for phrasal approaches and then explain why it is not sufficient to be able to assign certain f-structures to words: In (21a), we are dealing with a resultative construction. According to the plugging approach, the resultative meaning is contributed by a phrasal construction into which the verb *fischt* is inserted. There is no lexical item that requires a resultative predicate as its argument. If no such lexical item exists, then it is unclear how the relation between (21a) and (21b) can be established:

(21) a. [dass] jemand die Nordsee leer fischt
that somebody the North.Sea empty fishes

b. wegen der *Leerfischung* der Nordsee[20]
because of.the empty.fishing of.the North.Sea

As Figure 23.1 on the following page shows, both the arguments selected by the heads and the structures are completely different. In (21b), the element that is the subject of the related construction in (21a) is not realized. As is normally the case in nominalizations, it is possible to realize it in a PP with the preposition *durch* 'by':

(22) wegen der Leerfischung der Nordsee durch die Anrainerstaaten
because of.the empty.fishing of.the North.Sea by the neighboring.countries

If one assumes that the resultative meaning comes from a particular configuration in which a verb is realized, there would be no explanation for (21b) since no verb is involved in the analysis of this example. One could of course assume that a verb stem is inserted into a construction both in (21a) and (21b). The inflectional morpheme *-t* and the derivational morpheme *-ung* as well as an empty nominal inflectional morpheme would then be independent syntactic components of the analysis. However, since Goldberg (2003b: 119) and Asudeh et al. (2013) assume lexical integrity, only entire words can be inserted into syntactic constructions and hence the analysis of the nominalization of resultative constructions sketched here is not an option for them.

One might be tempted to try and account for the similarities between the phrases in (21) using inheritance. One would specify a general resultative construction standing in

[20] taz, 20.06.1996, p. 6.

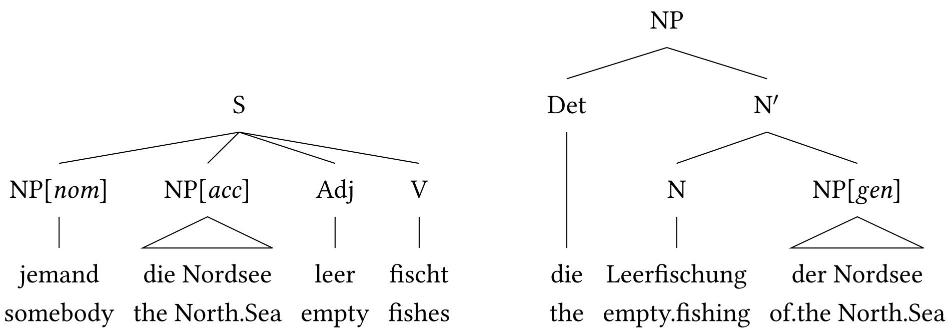


Figure 23.1: Resultative construction and nominalization

an inheritance relation to the resultative construction with a verbal head and the nominalization construction. I have discussed this proposal in more detail in Müller (2006: Section 5.3). It does not work as one requires embedding for derivational morphology and this cannot be modeled in inheritance hierarchies (Krieger & Nerbonne (1993), also see Müller (2006) for a detailed discussion).

It would also be possible to assume that both constructions in (23), for which structures such as those in Figure 23.1 would have to be assumed, are connected via metarules.[21,22]

(23) a. [Sbj Obj Obl V]
  b. [Det [[Adj V -ung]] NP[*gen*]]

The construction in (23b) corresponds to Figure 23.2 on the next page.[23] The genitive NP is an argument of the adjective. It has to be linked semantically to the subject slot of the adjective. Alternatively, one could assume that the construction only has the form [Adj V -*ung*], that is, that it does not include the genitive NP. But then one could also assume that the verbal variant of the resultative construction has the form [OBL V] and that Sbj and Obj are only represented in the valence lists. This would almost be a lexical analysis, however.

Turning to lexical integrity again, I want to point out that all that Ashudeh and Toivonen can do is assign some f-structure to the N in Figure 23.2. What is needed, however, is a principled account of how this f-structure comes about and how it is related to the resultative construction on the sentence level.

[21] Goldberg (p. c. 2007, 2009) suggests connecting certain constructions using GPSG-like metarules. Deppermann (2006: 51), who has a more Croftian view of CxG, rules this out. He argues for active/passive alternations that the passive construction has other information structural properties. Note also that GPSG metarules relate phrase structure rules, that is, local trees. The structure in Figure 23.2, however, is highly complex.

[22] The structure in (23b) violates a strict interpretation of lexical integrity as is commonly assumed in LFG. Booij (2005; 2009), working in Construction Grammar, subscribes to a somewhat weaker version, however.

[23] I do not assume zero affixes for inflection. The respective affix in Figure 23.2 is there to show that there is structure. Alternatively one could assume a unary branching rule/construction as is common in HPSG/Construction Morphology.

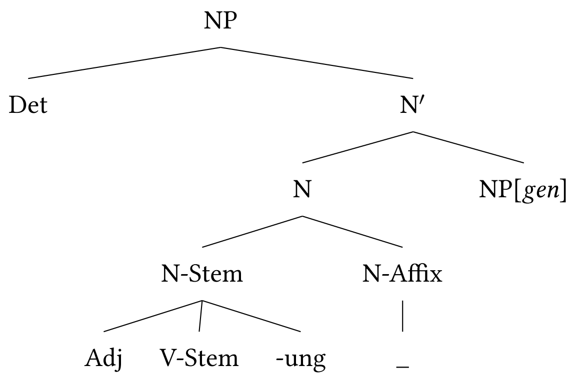


Figure 23.2: Resultative construction and nominalization

Before I turn to approaches with radical underspecification of argument structure in the next section, I want to comment on a more recent paper by Asudeh, Giorgolo & Toivonen (2014). The authors discuss the phrasal introduction of cognate objects and benefactives. (24a) is an example of the latter construction.

(24) a. The performer sang the children a song.
     b. The children were sung a song.

According to the authors the noun phrase *the children* is not an argument of *sing* but contributed by the c-structure rule that optionally licenses a benefactive.

(25) V′ → V DP DP
$\uparrow = \downarrow$ $(\uparrow \text{OBJ}) = \downarrow$ $(\uparrow \text{OBJ}_\theta) = \downarrow$
(@BENEFACTIVE)

Whenever this rule is called, the template BENEFACTIVE can add a benefactive role and the respective semantics if this is compatible with the verb that is inserted into the structure. The authors show how the mappings for the passive example in (24b) work, but they do not provide the c-structure that licenses such examples. In order to analyze these examples one would need a c-structure rule for passive VPs and this rule has to license a benefactive as well. So it would be:

(26) V′ → V[pass] DP
$\uparrow = \downarrow$ $(\uparrow \text{OBJ}_\theta) = \downarrow$
(@BENEFACTIVE)

Note that a benefactive cannot be added to any verb: Adding a benefactive to an intransitive verb as in (27a) is out and the passive that would correspond to (27a) is ungrammatical as well, as (27b) shows:

(27) a. * He laughed the children.

b. * The children were laughed.

So one could not just claim that all c-structure rule optionally introduce a benefactive argument. Therefore there is something special about the two rules in (25) and (26). The problem is that there is no relation between these rules. They are independent statements saying that there can be a benefactive in the active and that there can be one in the passive. This is what Chomsky (1957: 43) criticized in 1957 and this was the reason for the introduction of transformations (see Section 3.1.1 of this book). Bresnan-style LFG captured the generalizations by lexical rules and later by Lexical Mapping Theory. But if elements are added outside the lexical representations, the representations where these elements are added have to be related too. One could say that our knowledge about formal tools has changed since 1957. We know can use inheritance hierarchies to capture generalizations. So one can assume a type (or a template) that is the supertype of all those c-structure rules that introduce a benefactive. But since not all rules allow for the introduction of a benefactive element this basically amounts to saying: c-structure rule A, B, and C allow for the introduction of a benefactive. In comparison lexical rule-based approaches have one statement introducing the benefactive. The lexical rule states what verbs are appropriate for adding a benefactive and syntactic rules are not affected.

In Müller & Wechsler (2014a) we argued that the approach to Swedish caused motion constructions by Asudeh et al. (2008; 2013) would not carry over to German since German interacts with derivational morphology. Asudeh & Toivonen (2014) argued that Swedish is different from German and hence there would not be a problem. However, the situation is different with the benefactive constructions. Although English and German do differ in many respects both languages have similar dative constructions:

(28) a. He baked her a cake.

b. Er buk ihr einen Kuchen.
he baked her.DAT a cake

Now, the analysis of the free constituent order was explained by assuming binary branching structures in which a VP node is combined with one of its arguments or adjuncts (see Section 7.4). The c-structure rule is repeated in (29):

(29) $$\text{VP} \rightarrow \underset{(\uparrow \text{SUBJ} \mid \text{OBJ} \mid \text{OBJ}_\theta) = \downarrow}{\text{NP}} \quad \underset{\uparrow = \downarrow}{\text{VP}}$$

The dependent elements contributed to the f-structure of the verb and coherence/completeness ensured that all arguments of the verb where present. One could add the introduction of the benefactive argument to the VP node of the right-hand side of the rule. However, since the verb-final variant of (28b) would have the structure in (30), one would get spurious ambiguities, since the benefactive could be introduced at every node:

(30) weil [$_{VP}$ er [$_{VP}$ ihr [$_{VP}$ einen Kuchen [$_{VP}$ [$_{V}$ buk]]]]]
because he her a cake baked

So the only option seems to be to introduce the benefactive at the rule that got the recursion going, namely the rule that projected the lexical verb to the VP level. The rule (39) is repeated as (31) for convenience.

(31) VP → (V)
$\uparrow = \downarrow$

Note also that benefactive datives appear in adjectival environments as in (32):

(32) a. der seiner Frau einen Kuchen backende Mann
the his.DAT wife a.ACC cake backing man
'the man who is baking a cake for her'

b. der einen Kuchen seiner Frau backende Mann
the a.ACC cake his.DAT wife backing man
'the man who is baking a cake for her'

In order to account for these datives one would have to assume that the adjective to AP rule that would be parallel to (31) introduces the dative. The semantics of the benefactive template would have to somehow make sure that the benefactive argument is not added to intransitive verbs like *lachen* 'to laugh' or participles like *lachende* 'laughing'. While this may be possible, I find the overall approach unattractive. First it does not have anything to do with the original constructional proposal but just states that the benefactive may be introduced at several places in the syntax, secondly the unary branching syntactic rule is applying to a lexical item and hence is very similar to a lexical rule and thirdly the analysis does not capture cross-linguistic commonalities of the construction. In a lexical rule-based approach as the one that was suggested by Briscoe & Copestake (1999: Section 5), a benefactive argument is added to certain verbs and the lexical rule is parallel in all languages that have this phenomenon. The respective languages just differ in the way the arguments are realized in respect to their heads. In languages that have adjectival participles, these are derived from the respective verbal stems. The morphological rule is the same independent of benefactive arguments and the syntactic rules for adjectival phrases do not have to mention benefactive arguments.

23.3 Radical underspecification: The end of argument structure?

23.3.1 Neo-Davidsonianism

In the last section we examined proposals that assume that verbs come with certain argument roles and are inserted into prespecified structures that may contribute additional arguments. While we showed that this is not without problems, there are even more radical proposals that the construction adds all agent arguments, or even all arguments. The notion that the agent argument should be severed from its verbs is put forth by Marantz (1984; 1997), Kratzer (1996), Embick (2004) and others. Others suggest that no arguments are selected by the verb. Borer (2003) calls such proposals *exoskeletal* since the structure of the clause is not determined by the predicate, that is, the verb does not project an inner 'skeleton' of the clause. Counter to such proposals are *endoskeletal* approaches, in which the structure of the clause is determined by the predicate, that is, lexical proposals. The radical exoskeletal proposals are mainly proposed in Mainstream Generative Grammar (Borer 1994; 2003; 2005; Schein 1993; Hale & Keyser 1997; Lohndal 2012) but can also be found in HPSG (Haugereid 2009). We will not discuss these proposals in detail here, but we review the main issues insofar as they relate to the question of lexical argument structure.[24] We conclude that the available empirical evidence favors the lexical argument structure approach over such alternatives.

Davidson (1967) argued for an event variable in the logical form of action sentences (33a). Dowty (1989) coined the term *neo-Davidsonian* for the variant in (33b), in which the verb translates to a property of events, and the subject and complement dependents are translated as arguments of secondary predicates such as *agent* and *theme*. (Dowty (1989) called the system in (33a) an *ordered argument system*.) Kratzer (1996) further noted the possibility of mixed accounts such as (33c), in which the agent (subject) argument is severed from the *kill′* relation, but the theme (object) remains an argument of the *kill′* relation.[25]

(33) a. *kill*: $\lambda y \lambda x \exists e[kill(e, x, y)]$ (Davidsonian)
b. *kill*: $\lambda y \lambda x \exists e[kill(e) \wedge agent(e, x) \wedge theme(e, y)]$ (neo-Davidsonian)
c. *kill*: $\lambda y \lambda x \exists e[kill(e, y) \wedge agent(e, x)]$ (mixed)

Kratzer (1996) observed that a distinction between Davidsonian, neo-Davidsonian and mixed can be made either ‘in the syntax’ or ‘in the conceptual structure’ (Kratzer 1996: 110–111). For example, on a lexical approach of the sort we advocate here, any of the three alternatives in (33) could be posited as the semantic content of the verb *kill*. A lexical entry for *kill* on the mixed model appears in (34).

(34) $\begin{bmatrix} \textsc{phon} & \langle\ \textit{kill}\ \rangle \\ \textsc{arg-st} & \langle \text{NP}_x, \text{NP}_y \rangle \\ \textsc{content} & \textit{kill(e, y)} \wedge \textit{agent(e, x)} \end{bmatrix}$

In other words, the lexical approach is neutral on this question of the ‘conceptual structure’ of eventualities, as noted already in a different connection in Section 23.1.4. For that reason, certain semantic arguments for the neo-Davidsonian approach, such as those put forth by Schein (1993: Chapter 4) and Lohndal (2012), do not directly bear upon the issue of lexicalism, as far as we can tell.

But Kratzer (1996), among others, has gone further and argued for an account that is neo-Davidsonian (or rather, mixed) ‘in the syntax’. Kratzer’s claim is that the verb specifies only the internal argument(s), as in (35a) or (35b), while the agent (external argument) role is assigned by the phrasal structure. On the ‘neo-Davidsonian in the syntax’ view, the lexical representation of the verb has no arguments at all, except the event variable, as shown in (35c).

[24] See Müller (2010a: Section 11.11.3) for a detailed discussion of Haugereid’s approach.
[25] The event variable is shown as existentially bound, as in Davidson’s original account. As discussed below, in Kratzer’s version it must be bound by a lambda operator instead.

(35) a. *kill*: $\lambda y \lambda e[kill(e, y)]$ (agent is severed)
b. *kill*: $\lambda y \lambda e[kill(e) \wedge theme(e, y)]$ (agent is severed)
c. *kill*: $\lambda e[kill(e))]$ (all arguments severed)

On such accounts, the remaining dependents of the verb receive their semantic roles from silent secondary predicates, which are usually assumed to occupy the positions of functional heads in the phrase structure. An Event Identification rule identifies the event variables of the verb and the silent light verb Kratzer (1996: 22); this is why the existential quantifiers in (33) have been replaced with lambda operators in (35). A standard term for the agent-assigning silent predicate is 'little *v*'. These extra-lexical dependents are the analogues of the ones contributed by the constructions in Construction Grammar.

In the following subsections we address arguments that have been put forth in favor of the 'little v' hypothesis, from idiom asymmetries (Section 23.3.2) and deverbal nominals (Section 23.3.3). We argue that the evidence actually favors the lexical view. Then we turn to problems for exoskeletal approaches, from idiosyncratic syntactic selection (Section 23.3.4) and expletives (Section 23.3.5). We conclude with a look at the treatment of idiosyncratic syntactic selection under Borer's exoskeletal theory (Section 23.3.7), and a summary (Section 23.3.8).

23.3.2 Little *v* and idiom asymmetries

Marantz (1984) and Kratzer (1996) argued for severing the agent from the argument structure as in (35a), on the basis of putative idiom asymmetries. Marantz (1984) observed that while English has many idioms and specialized meanings for verbs in which the internal argument is the fixed part of the idiom and the external argument is free, the reverse situation is considerably rarer. To put it differently, the nature of the role played by the subject argument often depends on the filler of the object position, but not vice versa. To take Kratzer's examples (Kratzer 1996: 114):

(36) a. kill a cockroach
b. kill a conversation
c. kill an evening watching TV
d. kill a bottle (i.e. empty it)
e. kill an audience (i.e., wow them)

On the other hand, one does not often find special meanings of a verb associated with the choice of subject, leaving the object position open (examples from Marantz (1984: 26)):

(37) a. Harry killed NP.
b. Everyone is always killing NP.
c. The drunk refused to kill NP.
d. Silence certainly can kill NP.

Kratzer observes that a mixed representation of *kill* as in (38a) allows us to specify varying meanings that depend upon its sole NP argument.

(38) a. *kill*: $\lambda y \lambda e[kill(e, y)]$
  b. If *a* is a time interval, then kill(e, a) = truth if e is an event of wasting *a*
  If *a* is animate, then kill(e, a) = truth if e is an event in which *a* dies
  … etc.

On the polyadic (Davidsonian) theory, the meaning could similarly be made to depend upon the filler of the agent role. On the polyadic view, 'there is no technical obstacle' (Kratzer 1996: 116) to conditions like those in (38b), except reversed, so that it is the filler of the agent role instead of the theme role that affects the meaning. But, she writes, this could not be done if the agent is not an argument of the verb. According to Kratzer, the agent-severed representation (such as (38a)) disallows similar constraints on the meaning that depend upon the agent, thereby capturing the idiom asymmetry.

But as noted by Wechsler (2005), 'there is no technical obstacle' to specifying agent-dependent meanings even if the Agent has been severed from the verb as Kratzer proposes. It is true that there is no variable for the agent in (38a). But there is an event variable *e*, and the language user must be able to identify the agent of *e* in order to interpret the sentence. So one could replace the variable *a* with 'the agent of *e*' in the expressions in (38b), and thereby create verbs that violate the idiom asymmetry.

While this may seem to be a narrow technical or even pedantic point, it is nonetheless crucial. Suppose we try to repair Kratzer's argument with an additional assumption: that modulations in the meaning of a polysemous verb can only depend upon arguments of the *relation* denoted by that verb, and not on other participants in the event. Under that additional assumption, it makes no difference whether the agent is severed from the lexical entry or not. For example, consider the following (mixed) neo-Davidsonian representation of the semantic content in the lexical entry of *kill*:

(39) *kill*: $\lambda y \lambda x \lambda e[kill(e, y) \wedge agent(e, x)]$

Assuming that sense modulations can only be affected by arguments of the *kill(e,y)* relation, we derive the idiom asymmetry, even if (39) is the lexical entry for *kill*. So suppose that we try to fix Kratzer's argument with a different assumption: that modulations in the meaning of a polysemous verb can only depend upon an argument of the lexically denoted function. Kratzer's 'neo-Davidsonian in the syntax' lexical entry in (35a) lacks the agent argument, while the lexical entry in (39) clearly has one. But Kratzer's entry still fails to predict the asymmetry because, as noted above, it has the *e* argument and so the sense modulation can be conditioned on the 'agent of *e*'. As noted above, that event argument cannot be eliminated (for example through existential quantification) because it is needed in order to undergo event identification with the event argument of the silent light verb that introduces the agent Kratzer (1996: 22).

Moreover, recasting Kratzer's account in lexicalist terms allows for verbs to vary. This is an important advantage, because the putative asymmetry is only a tendency. Following are examples in which the subject is a fixed part of the idiom and there are open slots for non-subjects:

(40) a. A little bird told X that S.
   'X heard the rumor that S' (Nunberg et al. 1994: 526)
  b. The cat's got x's tongue.
   'X cannot speak.' (Bresnan 1982a: 349–350)
  c. What's eating x?
   'Why is X so galled?' (Bresnan 1982a: 349–350)

Further data and discussion of subject idioms in English and German can be found in Müller (2007b: Section 3.2.1).

The tendency towards a subject-object asymmetry plausibly has an independent explanation. Nunberg, Sag & Wasow (1994) argue that the subject-object asymmetry is a side-effect of an animacy asymmetry. The open positions of idioms tend to be animate while the fixed positions tend to be inanimate. Nunberg et al. (1994) derive these animacy generalizations from the figurative and proverbial nature of the metaphorical transfers that give rise to idioms. If there is an independent explanation for this tendency, then a lexicalist grammar successfully encodes those patterns, perhaps with a mixed neo-Davidsonian lexical decomposition, as explained above (see Wechsler (2005) for such a lexical account of the verbs *buy* and *sell*). But the 'little v' hypothesis rigidly predicts this asymmetry for all agentive verbs, and that prediction is not borne out.

23.3.3 Deverbal nominals

An influential argument against lexical argument structure involves English deverbal nominals and the causative alternation. It originates from a mention in Chomsky (1970), and is developed in detail by Marantz (1997); see also Pesetsky (1996) and Harley & Noyer (2000). The argument is often repeated, but it turns out that the empirical basis of the argument is incorrect, and the actual facts point in the opposite direction, in favor of lexical argument structure (Wechsler 2008b;a).

Certain English causative alternation verbs allow optional omission of the agent argument (41), while the cognate nominal disallows expression of the agent (42):

(41) a. that John grows tomatoes
  b. that tomatoes grow

(42) a. * John's growth of tomatoes
  b. the tomatoes' growth, the growth of the tomatoes

In contrast, nominals derived from obligatorily transitive verbs such as *destroy* allow expression of the agent, as shown in (44a):

(43) a. that the army destroyed the city
  b. * that the city destroyed

(44) a. the army's destruction of the city

b. the city's destruction

Following a suggestion by Chomsky (1970), Marantz (1997) argued on the basis of these data that the agent role is lacking from lexical entries. In verbal projections like (41) and (43) the agent role is assigned in the syntax by little *v*. Nominal projections like (42) and (44) lack little *v*. Instead, pragmatics takes over to determine which agents can be expressed by the possessive phrase: the possessive can express 'the sort of agent implied by an event with an external rather than an internal cause' because only the former can 'easily be reconstructed' (quoted from Marantz (1997: 218)). The destruction of a city has a cause external to the city, while the growth of tomatoes is internally caused by the tomatoes themselves (Smith 1970). Marantz points out that this explanation is unavailable if the noun is derived from a verb with an argument structure specifying its agent, since the deverbal nominal would inherit the agent of a causative alternation verb.

The empirical basis for this argument is the putative mismatch between the allowability of agent arguments, across some verb-noun cognate pairs: e.g. *grow* allows the agent but *growth* does not. But it turns out that the *grow/growth* pattern is rare. Most deverbal nominals precisely parallel the cognate verb: if the verb has an agent, so does the noun. Moreover, there is a ready explanation for the exceptional cases that exhibit the *grow/growth* pattern (Wechsler 2008a). First consider non-alternating theme-only intransitives ('unaccusatives'), as in (45) and non-alternating transitives as in (46). The pattern is clear: if the verb is agentless, then so is the noun:

(45) *arriv(al), disappear(ance), fall* etc.:

a. A letter arrived.

b. the arrival of the letter

c. * The mailman arrived a letter.

d. * the mailman's arrival of the letter

(46) *destroy/destruction, construct(ion), creat(ion), assign(ment)* etc.:

a. The army is destroying the city.

b. the army's destruction of the city

This favors the view that the noun inherits the lexical argument structure of the verb. For the anti-lexicalist, the badness of (45c) and (45d), respectively, would have to receive independent explanations. For example, on Harley and Noyer's 2000 proposal, (45c) is disallowed because a feature of the root ARRIVE prevents it from appearing in the context of *v*, but (45d) is instead ruled out because the cause of an event of arrival cannot be easily reconstructed from world knowledge. This exact duplication in two separate components of the linguistic system would have to be replicated across all non-alternating intransitive and transitive verbs, a situation that is highly implausible.

Turning to causative alternation verbs, Marantz's argument is based on the implicit generalization that noun cognates of causative alternation verbs (typically) lack the agent argument. But apart from the one example of *grow/growth*, there do not seem to be any clear cases of this pattern. Besides *grow(th)*, Chomsky 1970: examples (7c) and (8c) cited two experiencer predicates, *amuse* and *interest*: *John amused (interested) the children with his stories* versus **John's amusement (interest) of the children with his stories*. But this was later shown by Rappaport (1983) and Dowty (1989) to have an independent aspectual explanation. Deverbal experiencer nouns like *amusement* and *interest* typically denote a mental state, where the corresponding verb denotes an event in which such a mental state comes about or is caused. These result nominals lack not only the agent but all the eventive arguments of the verb, because they do not refer to events. Exactly to the extent that such nouns can be construed as representing events, expression of the agent becomes acceptable.

In a response to Chomsky 1970, Carlota Smith (1972) surveyed Webster's dictionary and found no support for Chomsky's claim that deverbal nominals do not inherit agent arguments from causative alternation verbs. She listed many counterexamples, including '*explode, divide, accelerate, expand, repeat, neutralize, conclude, unify*, and so on at length.' (Smith 1972: 137). Harley and Noyer (2000) also noted many so-called 'exceptions': *explode, accumulate, separate, unify, disperse, transform, dissolve/dissolution, detach(ment), disengage-(ment)*, and so on. The simple fact is that these are not exceptions because there is no generalization to which they can be exceptions. These long lists of verbs represent the norm, especially for suffix-derived nominals (in *-tion*, *-ment*, etc.). Many zero-derived nominals from alternating verbs also allow the agent, such as *change, release*, and *use*: *My constant change of mentors from 1992–1997. The frequent release of the prisoners by the governor. The frequent use of sharp tools by underage children.* (examples from Borer (2003: fn. 13)).[26]

Like the experiencer nouns mentioned above, many zero-derived nominals lack event readings. Some reject all the arguments of the corresponding eventive verb, not just the agent: **the freeze of the water, *the break of the window*, and so on. In the judgment of the second author, *his drop of the ball* is slightly odd, but *the drop of the ball* has exactly the same degree of oddness. The locution *a drop in temperature* matches the verbal one *The temperature dropped*, and both verbal and nominal forms disallow the agent: **The storm dropped the temperature. *the storm's drop of the temperature*. In short, the facts seem to point in exactly the opposite direction from what has been assumed in this oft-repeated argument against lexical valence. Apart from the one isolated case of *grow/growth*, event-denoting deverbal nominals match their cognate verbs in their argument patterns.

Turning to *grow/growth* itself, we find a simple explanation for its unusual behavior (Wechsler 2008a). When the noun *growth* entered the English language, causative (transitive) *grow* did not exist. The OED provides these dates of the earliest attestations of *grow* and *growth*:

[26] Pesetsky (1996: 79, ex. (231)) assigns a star to *the thief's return of the money*, but it is acceptable to many speakers, the *Oxford English Dictionary* lists a transitive sense for the noun *return* (definition 11a), and corpus examples like *her return of the spoils* are not hard to find.

(47) a. intransitive *grow*: c725 'be verdant' ... 'increase' (intransitive)
b. the noun *growth*: 1587 'increase' (intransitive)
c. transitive *grow*: 1774 'cultivate (crops)'

Thus *growth* entered the language at a time when transitive *grow* did not exist. The argument structure and meaning were inherited by the noun from its source verb, and then preserved into present-day English. This makes perfect sense if, as we claim, words have predicate argument structures. Nominalization by *-th* suffixation is not productive in English, so *growth* is listed in the lexicon. To explain why *growth* lacks the agent we need only assume that a lexical entry's predicate argument structure dictates whether it takes an agent argument or not. So even this one word provides evidence for lexical argument structure.

23.3.4 Idiosyncratic syntactic selections

The notion of lexical valence structure immediately explains why the argument realization patterns are strongly correlated with the particular lexical heads selecting those arguments. It is not sufficient to have general lexical items without valency information and let the syntax and world knowledge decide about argument realizations, because not all realizational patterns are determined by the meaning. The form of the preposition of a prepositional object is sometimes loosely semantically motivated but in other cases arbitrary. For example, the valence structure of the English verb *depend* captures the fact that it selects an *on*-PP to express one of its semantic arguments:

(48) a. John depends on Mary. (*counts, relies,* etc.)
b. John trusts (*on) Mary.
c. $\begin{bmatrix} \text{PHON} & \langle \textit{ depend } \rangle \\ \text{ARG-ST} & \langle \text{NP}_x \text{ , PP}[\textit{on}]_y \rangle \\ \text{CONTENT} & \textit{depend(x,y)} \end{bmatrix}$

Such idiosyncratic lexical selection is utterly pervasive in human language. The verb or other predicator often determines the choice between direct and oblique morphology, and for obliques, it determines the choice of adposition or oblique case. In some languages such as Icelandic even the subject case can be selected by the verb (Zaenen, Maling & Thráinsson 1985).

Selection is language-specific. English *wait* selects *for* (German *für*) while German *warten* selects *auf* 'on' with an accusative object:

(49) a. I am waiting for my man.
b. Ich warte auf meinen Mann.
I wait on my man.ACC

It is often impossible to find semantic motivation for case. In German there is a tendency to replace genitive (50a) with dative (50b) with no apparent semantic motivation:

(50) a. dass der Opfer gedacht werde
that the victims.GEN remembered was
'that the victims would be remembered'

b. daß auch hier den Opfern des Faschismus gedacht werde [...][27]
that also here the victims.DAT of.the fascism remembered was
'that the victims of fascism would be remembered here too'

The synonyms *treffen* and *begegnen* 'to meet' govern different cases (example from Pollard & Sag (1987: 126)).

(51) a. Er traf den Mann.
he.NOM met the man.ACC

b. Er begegnete dem Mann.
he.NOM met the man.DAT

One has to specify the case that the respective verbs require in the lexical items of the verbs.[28]

A radical variant of the plugging approach is suggested by Haugereid (2009). Haugereid (pages 12–13) assumes that the syntax combines a verb with an arbitrary combination of a subset of five different argument roles. Which arguments can be combined with a verb is not restricted by the lexical item of the verb.[29] A problem for such views is that the meaning of an ambiguous verb sometimes depends on which of its arguments are expressed. The German verb *borgen* has the two translations *borrow* and *lend*, which basically are two different perspectives on the same event (see Kunze (1991; 1993) for an extensive discussion of verbs of exchange of possession). Interestingly, the dative object is obligatory only with the *lend′* reading (Müller 2010a: 403):

(52) a. Ich borge ihm das Eichhörnchen.
I lend him the squirrel
'I lend the squirrel to him.'

b. Ich borge (mir) das Eichhörnchen.
I borrow me the squirrel
'I borrow the squirrel.'

If we omit it, we get only the *borrow′* reading. So the grammar must specify for specific verbs that certain arguments are necessary for a certain verb meaning or a certain perspective on an event.

Synonyms with differing valence specifications include the minimal triplet mentioned earlier: *dine* is obligatorily intransitive (or takes an *on*-PP), *devour* is transitive, and *eat*

[27] Frankfurter Rundschau, 07.11.1997, p. 6.

[28] Or at least mark the fact that *treffen* takes an object with the default case for objects and *begegnen* takes a dative object in German. See Haider (1985a), Heinz & Matiasek (1994), and Müller (2001) on structural and lexical case.

[29] Haugereid has the possibility to impose valence restrictions on verbs, but he claims that he uses this possibility just in order to get a more efficient processing of his computer implementation (p. 13).

can be used either intransitively or transitively (Dowty 1989: 89–90). Many other examples are given in Levin (1993) and Levin & Rappaport Hovav (2005).

In a phrasal constructionist approach one would have to assume phrasal patterns with the preposition or case, into which the verb is inserted. For (49b), the pattern includes a prepositional object with *auf* and an accusative NP, plus an entry for *warten* specifying that it can be inserted into such a structure (see Kroch & Joshi (1985: Section 5.2) for such a proposal in the framework of TAG). Since there are generalizations regarding verbs with such valence representations, one would be forced to have two inheritance hierarchies: one for lexical entries with their valency properties and another one for specific phrasal patterns that are needed for the specific constructions in which these lexical items can be used.

More often, proponents of neo-constructionist approaches either make proposals that are difficult to distinguish from lexical valence structures (see Section 23.3.7 below) or simply decline to address the problem. For instance Lohndal (2012) writes:

> An unanswered question on this story is how we ensure that the functional heads occur together with the relevant lexical items or roots. This is a general problem for the view that Case is assigned by functional heads, and I do not have anything to say about this issue here. (Lohndal 2012)

We think that getting case assignment right in simple sentences, without vast overgeneration of ill-formed word sequences, is a minimal requirement for a linguistic theory.

23.3.5 Expletives

A final example for the irreducibility of valence to semantics are verbs that select for expletives and reflexive arguments of inherently reflexive verbs in German:

(53) a. weil es regnet
because it rains

b. weil (es) mir (vor der Prüfung) graut
because EXPL me.DAT before the exam dreads

c. weil er es bis zum Professor bringt
because he EXPL until to.the professor brings
'because he made it to professor'

d. weil es sich um den Montag handelt
because EXPL REFL around the Monday trades
'It is about the Monday.'

e. weil ich mich (jetzt) erhole
because I myself now recreate
'because I am relaxing'

The lexical heads in (53) need to contain information about the expletive subjects/objects and/or reflexive pronouns that do not fill semantic roles. Note that German allows for subjectless predicates and hence the presence of expletive subjects cannot be claimed to follow from general principles. (53c) is an example with an expletive object. Explanations referring to the obligatory presence of a subject would fail on such examples in any case. Furthermore it has to be ensured that *erholen* is not realized in the [Sbj IntrVerb] construction for intransitive verbs or respective functional categories in a Minimalist setting although the relation *erholen′* (*relax′*) is a one-place predicate and hence *erholen* is semantically compatible with the construction.

23.3.6 An exoskeletal approach

In what follows I want to discuss Haugereid's proposal in more detail. His analysis has all the high-level problems that were mentioned in the previous subsections, but since it is worked out in detail it is interesting to see its predictions.

Haugereid (2007), working in the framework of HPSG, suggests an analysis along the lines of Borer (2005) where the meaning of an expression is defined as depending on the arguments that are present. He assumes that there are five argument slots that are assigned to semantic roles as follows:

- Arg1: agent or source
- Arg2: patient
- Arg3: benefactive or recipient
- Arg4: goal
- Arg5: antecedent

Here, antecedent is a more general role that stands for instrument, comitative, manner and source. The roles Arg1–Arg3 correspond to subject and objects. Arg4 is a resultative predicate of the end of a path. Arg4 can be realized by a PP, an AP or an NP. (54) gives examples for the realization of Arg4:

(54) a. John smashed the ball *out of the room.*
  b. John hammered the metal *flat.*
  c. He painted the car *a brilliant red.*

Whereas Arg4 follows the other participants in the causal chain of events, the antecedent precedes the patient in the order of events. It is realized as a PP. (55) is an example of the realization of Arg5:

(55) John punctured the balloon *with a needle.*

Haugereid now assumes that argument frames consist of these roles. He provides the examples in (56):

(56) a. John smiles. (arg1-frame)

b. John smashed the ball. (arg12-frame)
c. The boat arrived. (arg2-frame)
d. John gave Mary a book. (arg123-frame)
e. John gave a book to Mary. (arg124-frame)
f. John punctured the ball with a needle. (arg125-frame)

Haugereid points out that multiple verbs can occur in multiple argument frames. He provides the variants in (57) for the verb *drip*:

(57) a. The roof drips. (arg1-frame)
b. The doctor drips into the eyes. (arg14-frame)
c. The doctor drips with water. (arg15-frame)
d. The doctor drips into the eyes with water. (arg145-frame)
e. The roof drips water. (arg12-frame)
f. The roof drips water into the bucket. (arg124-frame)
g. The doctor dripped the eyes with water. (arg125-frame)
h. The doctor dripped into the eyes with water. (arg145-frame)
i. John dripped himself two drops of water. (arg123-frame)
j. John dripped himself two drops of water into his eyes. (arg1234-frame)
k. John dripped himself two drops of water into his eyes with a drop counter. (arg12345-frame)
l. Water dripped. (arg2-frame)
m. It drips. (arg0-frame)

He proposes the inheritance hierarchy in Figure 23.3 in order to represent all possible argument combinations, whereby the Arg5 role is omitted due to space considerations.

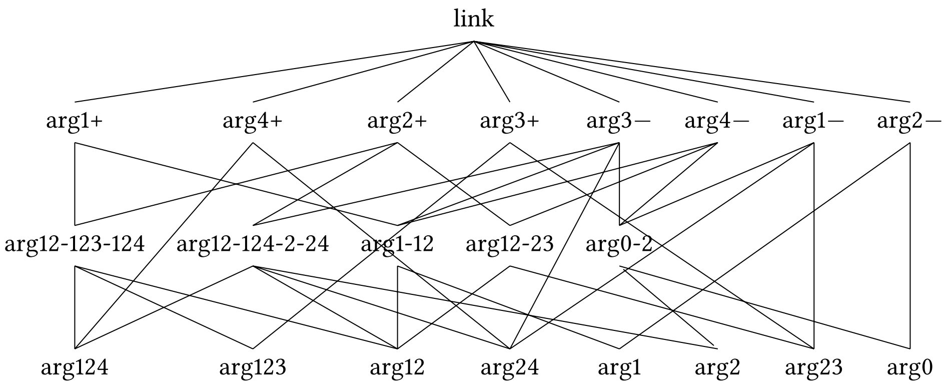


Figure 23.3: Hierarchy of argument frames following Haugereid (2007)

Haugereid assumes binary-branching structures where arguments can be combined with a head in any order. There is a dominance schema for each argument role. The schema realizing the argument role 3 provides a link value *arg3+*. If the argument role 2 is provided by another schema, we arrive at the frame *arg23*. For unergative intransitive verbs, it is possible to determine that it has an argument frame of *arg1*. This frame is only compatible with the types *arg1+*, *arg2−*, *arg3−* and *arg4−*. Verbs that have an optional object are assigned to *arg1-12* according to Haugereid. This type allows for the following combinations: *arg1+*, *arg2−*, *arg3−* and *arg4−* such as *arg1+*, *arg2+*, *arg3−* and *arg4−*.

This approach comes very close to an idea by Goldberg: Verbs are underspecified with regard to the sentence structures in which they occur and it is only the actual realization of arguments in the sentence that decides which combinations of arguments are realized. One should bear in mind that the hierarchy in Figure 23.3 corresponds to a considerable disjunction: She lists all possible realizations of arguments. If we say that *essen* 'to eat' has the type *arg1-12*, then this corresponds to the disjunction typearg1 ∨ *arg12*. In addition to the information in the hierarchy above, one also requires information about the syntactic properties of the arguments (case, the form of prepositions, verb forms in verbal complements). Since this information is in part specific to each verb, (see Section 23.1), it cannot be present in the dominance schemata and must instead be listed in each individual lexical entry. The lexical entry for *warten auf* 'to wait for', there must be information about the fact that the subject has to be an NP and that the prepositional object is an *auf*-PP with accusative. The use of a type hierarchy then allows one to elegantly encode the fact that the prepositional object is optional. The difference to a disjunctively specified SUBCAT list with the form of (58) is more formal in nature.

(58) SUBCAT ⟨ NP[*str*] ⟩ ∨ ⟨ NP[*str*], PP[*auf*, *acc*] ⟩

Since Haugereid's structures are binary-branching, it is possible to derive all permutations of arguments (59a–b), and adjuncts can be attached to every branching node (59c–d).

(59) a. dass [$_{\text{arg1}}$ keiner [$_{\text{arg2}}$ Pizza isst]]
that nobody pizza eats

b. dass [$_{\text{arg2}}$ Pizza [$_{\text{arg1}}$ keiner isst]]
that pizza nobody eats

c. dass [$_{\text{arg1}}$ keiner [gerne [$_{\text{arg2}}$ Pizza isst]]]
that nobody gladly pizza eats

d. dass [$_{\text{arg1}}$ [hier [keiner [$_{\text{arg2}}$ Pizza isst]]]
that here nobody pizza eats

Haugereid has therefore found solutions for some of the problems in Goldberg's analysis that are criticized in Müller (2006). Nevertheless, there are a number of other problems, which I will discuss in what follows. In Haugereid's approach, nothing is said about the composition of meaning. He follows the so-called Neo-Davidsonian approach. In these kind of semantic representations, arguments of the verb are not directly represented on the verb. Instead, the verb normally has an event argument and the argument roles belonging to the event in question is determined in a separate predication. (60) shows two alternative representations, where *e* stands for the event variable.

(60) a. Der Mann isst eine Pizza.
the man eats a pizza
'The man is eating a pizza'

b. eat'(e, x, y) ∧ man'(x) ∧ $pizza'$(y)

c. eat'(e) ∧ agent(e,x) ∧ theme(e,y) ∧ man'(x) ∧ $pizza'$(y)

Haugereid adopts Minimal Recursion Semantics (MRS) as his semantic formalism (also, see Section 9.1.6 und 21.3). The fact that arguments belong to a particular predicate is represented by the fact that the relevant predicates have the same handle. The representation in (60c) corresponds to (61):

(61) h1:$essen'$(e), h1:arg1(x), h1:arg2(y), h2:$mann'$(x), h3:$pizza'$(y)

This analysis captures Goldberg's main idea: Meaning arises from particular constituents being realized together with a head.

For the sentence in (62a), Haugereid (2007, p. c.) assumes the semantic representation in (62b):[30]

(62) a. der Mann den Teich leer fischt
the man the pond empty fishes

b. h1:$mann'$(x), h2:$teich'$(y), h3:$leer'$(e),
h4:$fischen'$(e2), h4:arg1(x), h4:arg2(y), h4:arg4(h3)

In (62b), the arg1, arg2 and arg4 relations have the same handle as $fischen'$. Following Haugereid's definitions, this means that arg2 is the patient of the event. In the case of (62a), this makes incorrect predictions since the accusative element is not a semantic argument of the main verb. It is a semantic argument of the secondary predicate *leer* 'empty' and has been raised to the object of the resultative construction. Depending on the exact analysis one assumes, the accusative object is either a syntactic argument of the verb or the adjective, however, it is never a semantic argument of the verb. In addition to this problem, the representation in (62b) does not capture the fact that *leer* predicates over the object. Haugereid (2007, p.c.) suggests that this is implicit in the representation and follows from the fact that all arg4s predicate over all arg2s. Unlike Haugereid's analysis, analyses using lexical rules that relate a lexical item of a verb to another verbal item with a resultative meaning allow for a precise specification of the semantic representation that then captures the semantic relation between the predicates involved. In addition, the lexical rule-based analysis makes it possible to license lexical

[30] See Haugereid (2009: 165) for an analysis of the Norwegian examples in (i).

(i) Jon maler veggen rød.
Jon paints wall.DEF red
'Jon paints the wall red.'

items that do not establish a semantic relation between the accusative object and the verb (Wechsler 1997; Wechsler & Noh 2001; Müller 2002a: Chapter 5).

Haugereid sketches an analysis of the syntax of the German clause and tackles active/passive alternations. However, certain aspects of the grammar are not elaborated on. In particular, it remains unclear how complex clauses containing AcI verbs such as *sehen* 'to see' and *lassen* 'to let' should be analyzed. Arguments of embedded and embedding verbs can also be permuted in these constructions. Haugereid (2007, p.c.) assumes special rules that allow one to saturate arguments of more deeply embedded verbs, for example, a special role that combines an arg2 argument of an argument with a verb. In order to combine *das Nilpferd* and *nicht füttern helfen lässt* in sentences such as (63), he is forced to assume a special grammatical rule that combines an argument of a doubly embedded verb with another verb:

(63) weil Hans Cecilia John das Nilpferd nicht füttern helfen lässt
because Hans Cecilia John the hippo not feed help let
'because Hans is not letting Cecilia help John feed the hippo.'

In Müller (2004c: 220), I have argued that embedding under complex-forming predicates is only constrained by performance factors (see also Section 12.6.3). In German, verbal complexes with more than four verbs are barely acceptable. Evers (1975: 58–59) has pointed out, however, that the situation in Dutch is different since Dutch verbal complexes have a different branching: In Dutch, verbal complexes with up to five verbs are possible. Evers attributes this difference to a greater processing load for German verbal complexes (also see Gibson (1998: Section 3.7)). Haugereid would have to assume that there are more rules for Dutch than for German. In this way, he would give up the distinction between competence and performance and incorporate performance restrictions directly into the grammar. If he wanted to maintain a distinction between the two, then Haugereid would be forced to assume an infinite number of schemata or a schema with functional uncertainty since depth of embedding is only constrained by performance factors. Existing HPSG approaches to the analysis of verbal complexes do without functional uncertainty. Since these kind of raising analyses are required for object raising anyway (as discussed above), they should be given preference.

Summing up, it must be said that Haugereid's exoskeletal approach does account for different orderings of arguments, but it neither gets the correct semantic representations nor does it offer a solution for the problem of idiosyncratic selection of arguments and the selection of expletives.

23.3.7 Is there an alternative to lexical valence structure?

The question for theories denying the existence of valence structure is what replaces it to explain idiosyncratic lexical selection. In her exoskeletal approach, Borer (2005) explicitly rejects lexical valence structures. But she posits post-syntactic interpretive rules that are difficult to distinguish from them. To explain the correlation of *depend* with an *on*-PP, she posits the following interpretive rule Borer (2005: Vol. II, p. 29):

(64) MEANING $\Leftrightarrow \pi_9 + [\langle e^{on} \rangle]$

Borer refers to all such cases of idiosyncratic selection as idioms. In a rule such as (64), 'MEANING is whatever the relevant idiom means.' (Borer 2005: Vol. II, p. 27) In (64), π_9 is the 'phonological index' of the verb *depend* and e^{on} 'corresponds to an open value that must be assigned range by the f-morph *on*' (Borer 2005: Vol. II, p. 29), where f-morphs are function words or morphemes. Hence this rule brings together much the same information as the lexical valence structure in (48c). Discussing such 'idiom' rules, Borer writes

> Although by assumption a listeme cannot be associated with any grammatical properties, one device used in this work has allowed us to get around the formidable restrictions placed on the grammar by such a constraint—the formation of idioms. [...] Such idiomatic specification could be utilized, potentially, not just for *arrive* and *depend on*, but also for obligatorily transitive verbs [...], for verbs such as *put*, with their obligatory locative, and for verbs which require a sentential complement.
>
> The reader may object that subcategorization, of sorts, is introduced here through the back door, with the introduction, in lieu of lexical syntactic annotation, of an articulated listed structure, called an *idiom*, which accomplishes, de facto, the same task. The objection of course has some validity, and at the present state of the art, the introduction of idioms may represent somewhat of a concession.
> (Borer 2005: Vol. II, p. 354–355)

Borer goes on to pose various questions for future research, related to constraining the class of possible idioms. With regard to that research program it should be noted that a major focus of lexicalist research has been narrowing the class of subcategorization and extricating derivable properties from idiosyncratic subcategorization. Those are the functions of HPSG lexical hierarchies, for example.

23.3.8 Summary

In Sections 23.3.2–23.3.5 we showed that the question of which arguments must be realized in a sentence cannot be reduced to semantics and world knowledge or to general facts about subjects. The consequence is that valence information has to be connected to lexical items. One therefore must either assume a connection between a lexical item and a certain phrasal configuration as in Croft's approach (2003) and in LTAG or assume our lexical variant. In a Minimalist setting the right set of features must be specified lexically to ensure the presence of the right case assigning functional heads. This is basically similar to the lexical valence structures we are proposing here, except that it needlessly introduces various problems discussed above, such as the problem of coordination raised in Section 23.2.1.

23.4 Relations between constructions

On the lexical rules approach, word forms are related by lexical rules: a verb stem can be related to a verb with finite inflection and to a passive verb form; verbs can be converted to adjectives or nouns; and so on. The lexical argument structure accompanies the word and can be manipulated by the lexical rule. In this section we consider what can replace such rules within a phrasal or ASC approach.

23.4.1 Inheritance hierarchies for constructions

For each valence structure that the lexicalist associates with a root lexeme (transitive, ditransitive, etc.), the phrasal approach requires multiple phrasal constructions, one to replace each lexical rule or combination of lexical rules that can apply to the word. Taking ditransitives, for example, the phrasal approach requires an active-ditransitive construction, a passive-ditransitive construction, and so on, to replace the output of every lexical rule or combination of lexical rules applied to a ditransitive verb. (Thus Bergen & Chang (2005: 169–170) assume an active-ditransitive and a passive-ditransitive construction and Kallmeyer & Osswald (2012: 171–172) assume active and passive variants of the transitive construction.) On that view some of the active voice constructions for German would be:

(65) a. Nom V
  b. Nom Acc V
  c. Nom Dat V
  d. Nom Dat Acc V

The passive voice constructions corresponding to (65) would be:

(66) a. V V-Aux
  b. Nom V V-Aux
  c. Dat V V-Aux
  d. Dat Nom V V-Aux

Merely listing all these constructions is not only uneconomical but fails to capture the obvious systematic relation between active and passive constructions. Since phrasalists reject both lexical rules and transformations, they need an alternative way to relate phrasal configurations and thereby explain the regular relation between active and passive. The only proposals to date involve the use of inheritance hierarchies, so let us examine them.

Researchers working in various frameworks, both with lexical and phrasal orientation, have tried to develop inheritance-based analyses that could capture the relation between valency patterns such as those in (65) and (66) (see for instance Kay & Fillmore (1999: 12); Michaelis & Ruppenhofer (2001: Chapter 4); Candito 1996; Clément & Kinyon 2003: 188;

Kallmeyer & Osswald 2012: 171–172; Koenig 1999: Chapter 3; Davis & Koenig 2000; Kordoni 2001 for proposals in CxG, TAG, and HPSG). The idea is that a single representation (lexical or phrasal, depending on the theory) can inherit properties from multiple constructions. In a phrasal approach the description of the pattern in (65b) inherits from the transitive and the active construction and the description of (66b) inherits from both the transitive and the passive constructions. Figure 23.4 illustrates the inheritance-based lexical approach: a lexical entry for a verb such as *read* or *eat* is combined with either an active or passive representation. The respective representations for the active and passive are responsible for the expression of the arguments.

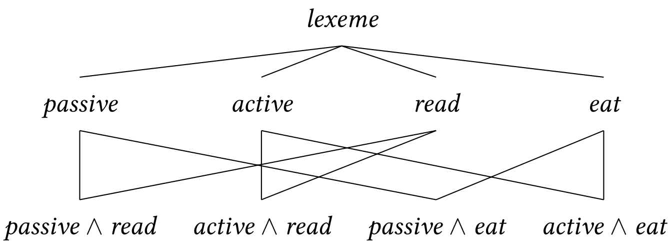


Figure 23.4: Inheritance Hierarchy for active and passive

As was already discussed in Section 10.2, inheritance-based analyses cannot account for multiple changes in valence as for instance the combination of passive and impersonal construction that can be observed in languages like Lithuanian (Timberlake 1982: Section 5), Irish (Noonan 1994), and Turkish (Özkaragöz 1986). Özkaragöz's Turkish examples are repeated here with the original glossing as (67) for convenience:

(67) a. Bu şato-da boğ-ul-un-ur. (Turkish)
this château-LOC strangle-PASS-PASS-AOR
'One is strangled (by one) in this château.'

b. Bu oda-da döv-ül-ün-ür.
this room-LOC hit-PASS-PASS-AOR
'One is beaten (by one) in this room.'

c. Harp-te vur-ul-un-ur.
war-LOC shoot-PASS-PASS-AOR
'One is shot (by one) in war.'

Another example from Section 10.2 that cannot be handled with inheritance is multiple causativization in Turkish. Turkish allows double and even triple causativization (Lewis 1967: 146):

(68) Öl-dür-t-tür-t- (Turkish)
'to cause somebody to cause somebody to kill somebody'

An inheritance-based analysis would not work, since inheriting the same information several times does not add anything new. Krieger & Nerbonne (1993) make the same point with respect to derivational morphology in cases like *preprepreversion*: Inheriting information about the prefix *pre-* twice or more often, does not add anything.

So assuming phrasal models, the only way to capture the generalization with regard to (65) and (66) seems to be to assume GPSG-like metarules that relate the constructions in (65) to the ones in (66). If the constructions are lexically linked as in LTAG, the respective mapping rules would be lexical rules. For approaches that combine LTAG with the Goldbergian plugging idea such as the one by Kallmeyer & Osswald (2012) one would have to have extended families of trees that reflect the possibility of having additional arguments and would have to make sure that the right morphological form is inserted into the respective trees. The morphological rules would be independent of the syntactic structures in which the derived verbal lexemes could be used. One would have to assume two independent types of rules: GPSG-like metarules that operate on trees and morphological rules that operate on stems and words. We believe that this is an unnecessary complication and apart from being complicated the morphological rules would not be acceptable as form-meaning pairs in the CxG sense since the aspect of the form that additional arguments are required is not captured in these morphological rules. If such morphological rules were accepted as proper constructions then there would not be any reason left to require that the arguments have to be present in a construction in order for it to be recognizable, and hence, the lexical approach would be accepted. Compare the discussion of *Totschießen* 'shoot dead' in example (80) below.

Inheritance hierarchies are the main explanatory device in Croft's Radical Construction Grammar (Croft 2001). He also assumes phrasal constructions and suggests representing these in a taxonomic network (an inheritance hierarchy). He assumes that every idiosyncrasy of a linguistic expression is represented on its own node in this kind of network. Figure 23.5 shows part of the hierarchy he assumes for sentences. There are

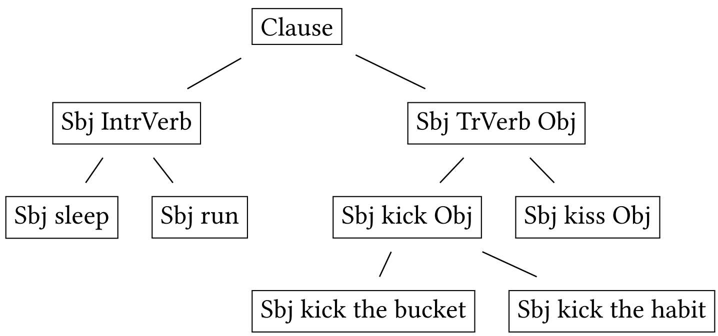


Figure 23.5: Classification of phrasal patterns in Croft (2001: 26)

sentences with intransitive verbs and sentences with transitive verbs. Sentences with the form [Sbj kiss Obj] are special instances of the construction [Sbj TrVerb Obj]. The [Sbj kick Obj] construction also has further sub-constructions, namely the constructions

[Sbj kick the bucket] and [Subj kick the habit]. Since constructions are always pairs of form and meaning, this gives rise to a problem: In a normal sentence with *kick*, there is a kicking relation between the subject and the object of *kick*. This is not the case for the idiomatic use of *kick* in (69):

(69) He kicked the bucket.

This means that there cannot be a normal inheritance relation between the [Sbj kick Obj] and the [Sbj kick the bucket] construction. Instead, only parts of the information may be inherited from the [Sbj kick Obj] construction. The other parts are redefined by the sub-construction. This kind of inheritance is referred to as *default inheritance*.

kick the bucket is a rather fixed expression, that is, it is not possible to passivize it or front parts of it without losing the idiomatic reading (Nunberg, Sag & Wasow 1994: 508). However, this is not true for all idioms. As Nunberg, Sag & Wasow (1994: 510) have shown, there are idioms that can be passivized (70a) as well as realizations of idioms where parts of idioms occur outside of the clause (70b).

(70) a. The beans were spilled by Pat.
b. The strings [that Pat pulled] got Chris the job.

The problem is now that one would have to assume two nodes in the inheritance hierarchy for idioms that can undergo passivization as the realization of constituents is different with the idiosyncratic meaning. The relation between the active and passive form would be not be captured. Kay (2002) has proposed a process where one can calculate objects (Construction-like objects = CLOs) from hierarchies that then license active and passive variants. As I have shown in Müller (2006: Section 3), this process does not deliver the desired results and it is far from straightforward to improve the procedure to the point that it actually works. Even if one were to adopt the changes I proposed, there are still phenomena that cannot be described using inheritance hierarchies (see Section 10.2 in this book).

A further interesting point is that the verbs have to be explicitly listed in the constructions. This begs the question of how constructions should be represented where the verbs are used differently. If a new node in the taxonomic network is assumed for cases like (71), then Goldberg's criticism of lexical analyses that assume several lexical entries for a verb that can appear in various constructions[31] will be applicable here: one would have to assume constructions for every verb and every possible usage of that verb.

(71) He kicked the bucket into the corner.

For sentences with negation, Croft assumes the hierarchy with multiple inheritance given in Figure 23.6 on the facing page. The problem with this kind of representation is that it remains unclear as to how the semantic embedding of the verb meaning

[31] Note the terminology: I used the word *lexical entry* rather than *lexical item*. The HPSG analysis uses lexical rules that correspond to Goldberg's templates. What Goldberg criticizes is lexical rules that relate lexical enbtries not lexial rules that licence new lexical items, which may be stored or not. HPSG takes the latter approach to lexical rules.

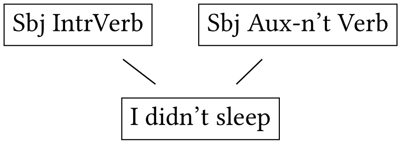


Figure 23.6: Interaction of phrasal patterns following Croft (2001: 26)

under negation can be represented. If all constructions are pairs of form and meaning, then there would have to be a semantic representation for [Sbj IntrVerb] (cont-Wert or sem value). Similarly, there would have to be a meaning for [Sbj Aux-n’t Verb]. The problem now arises that the meaning of [Sbj IntrVerb] has to be embedded under the meaning of the negation and this cannot be achieved directly using inheritance since X and not(X) are incompatible. There is a technical solution to this problem using auxiliary features. Since there are a number of interactions in grammars of natural languages, this kind of analysis is highly implausible if one claims that features are a direct reflection of observable properties of linguistic objects. For a more detailed discussion of approaches with classifications of phrasal patterns, see Müller (2010b) as well as Müller (2007b: Section 18.3.2.2) and for the use of auxiliary features in inheritance-based analyses of the lexicon, see Müller (2007b: Section 7.5.2.2).

23.4.2 Mappings between different levels of representations

Culicover & Jackendoff (2005: Chapter 6.3) suggest that passive should be analyzed as one of several possible mappings from the Grammatical Function tier to the surface realization of arguments. Surface realizations of referential arguments can be NPs in a certain case, with certain agreement properties, or in a certain position. While such analyses that work by mapping elements with different properties onto different representations are common in theories like LFG and HPSG (Koenig 1999; Bouma, Malouf & Sag 2001a), a general property of these analyses is that one needs one level of representation per interaction of phenomena (arg-st, sem-arg, add-arg in Koenig’s proposal, arg-st, deps, spr, comps in Bouma, Malouf, and Sag’s proposal). This was discussed extensively in Müller (2007b: Section 7.5.2.2) with respect to extensions that would be needed for Koenig’s analysis.

Since Culicover and Jackendoff argue for a phrasal model, we will discuss their proposal here. Culicover and Jackendoff assume a multilayered model in which semantic representations are linked to grammatical functions, which are linked to tree positions. Figure 23.7 shows an example for an active sentence. GF stands for Grammatical Function. Culicover & Jackendoff (2005: 204) explicitly avoid names like Subject and Object since this is crucial for their analysis of the passive to work. They assume that the first GF following a bracket is the subject of the clause the bracket corresponds to (p. 195–196) and hence has to be mapped to an appropriate tree position in English. Note that this view on grammatical functions and obliqueness does not account for subjectless

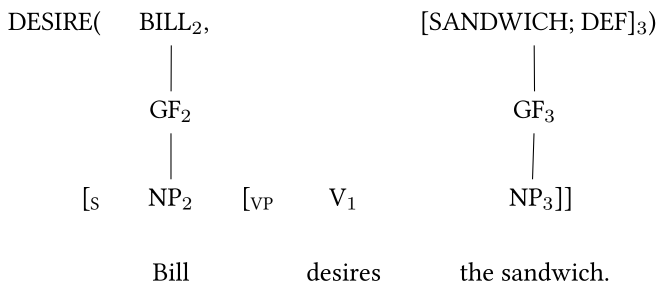


Figure 23.7: Linking grammatical functions to tree positions: active

sentences that are possible in some languages, for instance in German.[32]

Regarding the passive, the authors write:

> we wish to formulate the passive not as an operation that deletes or alters part of the argument structure, but rather as a piece of structure in its own right that can be unified with the other independent pieces of the sentence. The result of the unification is an alternative licensing relation between syntax and semantics. (Culicover & Jackendoff 2005: 203)

They suggest the following representation of the passive:

(72) $[GF_i > [GF\ ...]]_k \Leftrightarrow [\ ...\mathrm{V}_k + \mathrm{pass}\ ...(\mathrm{by\ NP}_i)\ ...]_k$

The italicized parts are the normal structure of the sentence and the non-italicized parts are an overlay on the normal structure, that is, additional constraints that have to hold in passive sentences. Figure 23.8 shows the mapping of the example discussed above that corresponds to the passive.

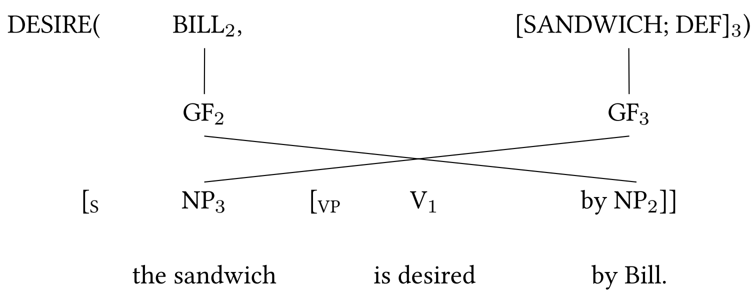


Figure 23.8: Linking grammatical functions to tree positions: passive

[32] Of course one could assume empty expletive subjects, as was suggested by Grewendorf (1993: 1311), but empty elements and especially those without meaning are generally avoided in the constructionist literature. See Müller (2010a: Section 3.4, Section 11.1.1.3) for further discussion.

Although Culicover and Jackendoff emphasize the similarity between their approach and Relational Grammar (Perlmutter 1983), there is an important difference: In Relational Grammar additional levels (strata) can be stipulated if additional remappings are needed. In Culicover and Jackendoff's proposal there is no additional level. This causes problems for the analysis of languages which allow for double passivization. Examples for such languages were already given in (67) in the previous subsection and specific examples from Turkish were provided in (67). Approaches that assume that the personal passive is the unification of a general structure with a passive-specific structure will not be able to capture this, since they committed to a certain structure too early. The problem for approaches that state syntactic structure for the passive is that such a structure, once stated, cannot be modified. Culicover and Jackendoff's proposal works in this respect since there are no strong constraints in the right-hand side of their constraint in (72). But there is a different problem: When passivization is applied the second time, it has to apply to the innermost bracket, that is, the result of applying (72) should be:

(73) $[GF_i > [GF_j \ldots]]_k \Leftrightarrow [\ \ldots \mathrm{V}_k + \mathrm{pass} \ldots (\mathrm{by\ NP}_i) \ldots (\mathrm{by\ NP}_j) \ldots]_k$

This cannot be done with unification, since unification checks for compatibility and since the first application of passive was possible it would be possible for the second time as well. Dots in representations are always dangerous and in the example at hand one would have to make sure that NP_i and NP_j are distinct, since the statement in (72) just says there has to be a *by* PP somewhere. What is needed instead of unification would be something that takes a GF representation and searches for the outermost bracket and then places a bracket to the left of the next GF. But this is basically a rule that maps one representation onto another one, just like lexical rules do.

If Culicover and Jackendoff want to stick to a mapping analysis, the only option to analyze the data seems to be to assume an additional level for impersonal passives from which the mapping to phrase structure is done. In the case of Turkish sentences like (74), which is a personal passive, the mapping to this level would be the identity function.

(74) Arkada-şım bu oda-da döv-ül-dü.
friend-my this room-LOC hit-PASS-AOR
'My friend is beaten (by one) in this room.'

In the case of double passivization the correct mappings would be implemented by two mappings between the three levels that finally result in a mapping as the one that is seen in (67b). Note that the double passivization is also problematic for purely inheritance based approaches. What all these approaches can suggest though is that they just stipulate three different relations between argument structure and phrase structure: active, passive, double passive. But this misses the fact that (67b) is a further passivization of (74).

In contrast, the lexical rule-based approach suggested by Müller (2003b) does not have any problems with double passivization: The first application of the passivization lexical rule suppresses the least oblique argument and provides a lexical item with the argument structure of a personal passive. The second application suppresses the now least oblique argument (the object of the active clause) and results in an impersonal passive.

23.4.3 Is there an alternative to lexical rules?

In this section we have reviewed the attempts to replace lexical rules with methods of relating constructions. Those attempts have not been successful, in our assessment. We believe that the essential problem with them is that they fail to capture the derivational character of the relationship between certain word forms. Alternations signaled by passive voice and causative morphology are relatively simple and regular when formulated as operations on lexical valence structures that have been abstracted from their phrasal context. But non-transformational rules or systems formulated on the phrasal structures encounter serious problems that have not yet been solved.

23.5 Further problems for phrasal approaches

Müller (2006) discussed the problems that proposals have that assume phrasal constructions to be a fixed configuration of adjacent material as for instance (Goldberg & Jackendoff 2004). I showed that many argument structure constructions allow great flexibility as far as the order of their parts is concerned. Back then I discussed resultative constructions in their interaction with free datives, passive and other valency changing phenomena and showed that for all these constructions that are licensed by such interactions the construction parts can be scrambled, the verb can appear in different positions, arguments can be extracted and so on. The following subsection discusses particle verbs, which pose similar problems for theories that assume a phrasal construction with fixed order of verb and particle.

23.5.1 Particle verbs and commitment to phrase structure configurations

A general problem of approaches that assume phrase structure configurations paired with meaning is that the construction may appear in different contexts: the construction parts may be involved in derivational morphology (as discussed in the previous subsection) or the construction parts may be involved in dislocations. A clear example of the latter type is the phrasal analysis of particle verbs that was suggested by Booij (2002: Section 2; To appear) and Blom (2005), working in the frameworks of Construction Grammar and LFG, respectively. The authors working on Dutch and German assume that particle verbs are licensed by phrasal constructions (pieces of phrase structure) in which the first slot is occupied by the particle.

(75) $[\ \mathrm{X}\ [\ ]_{\mathrm{V}}\ ]_{\mathrm{V'}}$ where X = P, Adv, A, or N

Examples for specific Dutch constructions are:

(76) a. $[\ \text{af}\ [\ ]_{\mathrm{V}}\ ]_{\mathrm{V'}}$
b. $[\ \text{door}\ [\ ]_{\mathrm{V}}\ ]_{\mathrm{V'}}$
c. $[\ \text{op}\ [\ ]_{\mathrm{V}}\ ]_{\mathrm{V'}}$

This suggestion comes with the claim that particles cannot be fronted. This claim is made frequently in the literature, but it is based on introspection and wrong for languages like Dutch and German. On Dutch see Hoeksema (1991: 19), on German Müller (2002a;c; 2003c; 2007d).[33] A German example is given in (77); several pages of attested examples can be found in the cited references and some more complex examples will also be discussed in Section 23.7.3 on page 623.

(77) *Los* damit *geht* es schon am 15. April.[34]
PART there.with went it already at.the 15 April
'It already started on April the 15th.'

Particle verbs are mini-idioms. So the conclusion is that idiomatic expressions that allow for a certain flexibility in order, so they should not be represented as phrasal configurations describing adjacent elements. For some idioms, a lexical analysis along the lines of Sag (2007) seems to be required.[35] The issue of particle verbs will be taken up in Section 23.7.3 again, where we discuss evidence for/against phrasal analyses from neuro science.

23.6 Arguments from language acquisition

The question whether language acquisition is pattern-based and hence can be seen as evidence for the phrasal approach has already been touched in the Sections 16.3 and 16.4. It was argued that constructions can be realized discontinuously in coordinations and hence it is the notion of dependency that has to be acquired, acquiring simple continuous patterns is not sufficient.

Since this section about phrasal and lexical approaches deals with specific proposals, I would like to add two more special subsections: Section 23.6.1 deals with the recognizability of constructions and Section 23.6.2 discusses specific approaches to coordination in order to demonstrate how frameworks deal with the discontinuous realization of constructions.

23.6.1 Recognizability of constructions

I think that a purely pattern-based approach is weakened by the existence of examples like (78):

(78) a. John tried to sleep.
b. John tried to be loved.

[33] Some more fundamental remarks on introspection and corpus data with relation to particle verbs can also be found in Müller (2007d); Meurers & Müller (2009).

[34] taz, 01.03.2002, p. 8, see also Müller (2005c: 313).

[35] Note also that the German example is best described as a clause with a complex internally structured constituent in front of the finite verb and it is doubtful whether linearization-based proposals like the ones in Kathol (1995: 244–248) or Wetta (2011) can capture this. See also the discussion of multiple frontings in connection to Dependency Grammar in Section 11.7.1.

Although no argument of *sleep* is present in the phrase *to sleep* and neither a subject nor an object is realized in the phrase *to be loved*, both phrases are recognized as phrases containing an intransitive and a transitive verb, respectively.[36]

The same applies to arguments that are supposed to be introduced/licensed by a phrasal construction: in (79) the resultative construction is passivized and then embedded under a control verb, resulting in a situation in which only the result predicate (*tot* 'dead' and the matrix verb (*geschossen* 'shot' are realized overtly within the local clause, bracketed here:

(79) Der kranke Mann wünschte sich, [tot geschossen zu werden].[37]
the sick man wished SELF dead shot to be
'The sick man wanted to be shot dead.'

Of course passivization and control are responsible for these occurrences, but the important point here is that arguments can remain unexpressed or implicit and nevertheless a meaning that is usually connected to some overt realization of arguments is present (Müller 2007c: Section 4). So, what has to be acquired by the language learner is that when a result predicate and a main verb are realized together, they contribute the resultative meaning. To take another example, NP arguments that are usually realized in active resultative constructions may remain implicit in nominalizations like the ones in (80):

(80) a. dann scheint uns das Totschießen mindestens ebensoviel Spaß zu machen[38]
then seems us the dead-shooting at.least as.much fun to make
'then the shooting dead seems to us to be as least as much fun'

b. Wir lassen heut das Totgeschieße, Weil man sowas heut nicht tut. Und wer einen Tag sich ausruht, Der schießt morgen doppelt gut.[39]
we let today the annoying.repeated.shooting.dead since one such.thing today not does and who a day SELF rests this shoots tomorrow twice good
'We do not shoot anybody today, since one does not do this, and those who rest a day shoot twice as well tomorrow.'

The argument corresponding to the patient of the verb (the one who is shot) can remain unrealized, because of the syntax of nominalizations. The resultative meaning is still understood, which shows that it does not depend upon the presence of a resultative construction involving Subj V Obj and Obl.

[36] Constructionist theories do not assume empty elements. Of course, in the GB framework the subject would be realized by an empty element. So it would be in the structure, although inaudible.

[37] Müller (2007c: 387).

[38] https://www.elitepartner.de/forum/wie-gehen-die-maenner-mit-den-veraenderten- anspruechen-der-frauen-um-26421-6.html. 26.03.0212.

[39] http://home.arcor.de/finishlast/indexset.html?dontgetmestarted/091201-1.html. 26.03.2012.

23.6.2 Coordination and discontinousness

The following subsection deals with analyses of coordination in some of the frameworks that were introduced in this book. The purpose of the section is to show that simple phrasal patterns have to be broken up in coordination structures. This was already mentioned in Section 16.3, but I think it is illuminative to have a look at concrete proposals.

In Categorial Grammar, there is a very elegant treatment of coordination (see Steedman (1991)). A generalization with regard to so-called symmetric coordination is that two objects with the same syntactic properties are combined to an object with those properties. We have already encountered the relevant data in the discussion of the motivation for feature geometry in HPSG on page 267. They are repeated below as (81):

(81) a. [der Mann] und [die Frau]
 the man and the woman

b. Er [kennt] und [liebt] diese Schallplatte.
 he knows and loves this record

c. Er ist [dumm] und [arrogant].
 he is dumb and arrogant

Steedman (1991) analyzes examples such as those in (81) with a single rule:

(82) X conj X ⇒ X

This rule combines two categories of the same kind with a conjunction in between to form a category that has the same category as the conjuncts.[40] Figure 23.9 shows the analysis (81a) and Figure 23.10 on the following page gives an analysis of the corresponding English example of (81b).[41]

$$\dfrac{\dfrac{\dfrac{der}{np/n} \quad \dfrac{Mann}{n}}{np}> \quad \dfrac{und}{conj} \quad \dfrac{\dfrac{die}{np/n} \quad \dfrac{Frau}{n}}{np}>}{np}\&$$

Figure 23.9: Coordination of two NPs in Categorial Grammar

If we compare this analysis to the one that would have to be assumed in traditional phrase structure grammars, it becomes apparent what the advantages are: One rule was required for the analysis of NP coordination where two NPs are coordinated to form

[40] Alternatively, one could analyze all three examples using a single lexical entry for the conjunction *und*: *und* is a functor that takes a phrase of any category to its right and after this combination then needs to be combined with an element of the same category to its left in order to form the relevant category after combining with this second element. This means that the category for *und* would have the form (X\X)/X. This analyses does not require any coordination rules. If one wants to assume, as is common in GB/MP, that every structure has a head, then this head-less analysis would be ruled out by using a special rule.

[41] Here, I am providing an analysis of an English example since the German example would require an analysis with verb-second order.

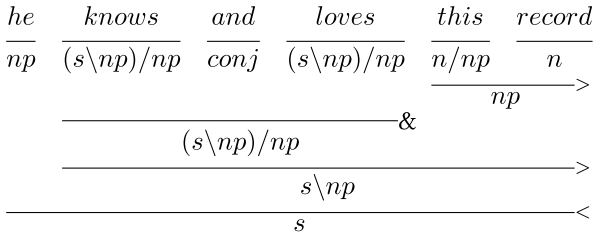


Figure 23.10: Coordination of two transitive verbs in Categorial Grammar

an NP and another was required for the analysis of V coordination. This is not only undesirable from a technical point of view, neither does it capture the basic property of symmetric coordination: Two symbols with the same syntactic category are combined with each other.

It is interesting to note that it is possible to analyze sentences such as (83) in this way

(83) give George a book and Martha a record

In Section 1.3.2.4, we have seen that these kind of sentences are problematic for constituent tests. However, in Categorial Grammar, it is possible to analyze them without any problems if one adopts rules for type raising and composition as Dowty (1988) and Steedman (1991) do. In Section 8.5, we have already seen forward type raising as well as forward and backward composition. In order to analyze (83), one would require backward type raising in (84) an backward composition repeated in (85):

(84) Backward type raising (< T)
X ⇒ T\(T/X)

(85) Backward composition (< B)
Y\Z * X\Y = X\Z

Dowty's analysis of (83) is given in Figure 23.11. VP stands for s\np.

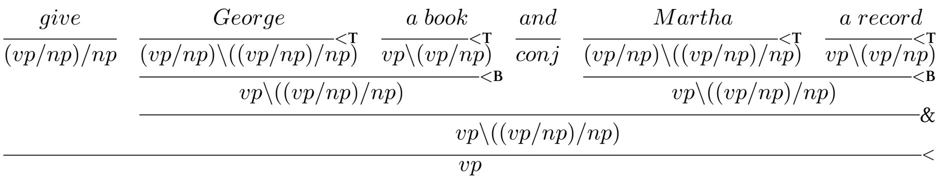


Figure 23.11: Gapping in Categorial Grammar

This kind of type-raising analysis was often criticized because raising categories leads to many different analytical possibilities for simple sentences. For example, one could first combine a type-raised subject with the verb and then combine the resulting constituent with the object. This would mean that we would have a [[S V] O] in addition to the standard [S [V O]] analysis. Steedman (1991) argues that both analyses differ in terms of information structure and it is therefore valid to assume different structures for the sentences in question.

I will not go into these points further here. However, I would like to compare Steedman's lexical approach to phrasal analyses: All approaches that assume that the ditransitive construction represents a continuous pattern encounter a serious problem with the examples discussed above. This can be best understood by considering the TAG analysis of coordination proposed by Sarkar & Joshi (1996). If one assumes that [Sbj TransVerb Obj] or [S [V O]] constitutes a fixed unit, then the trees in Figure 23.12 form the starting point for the analysis of coordination.

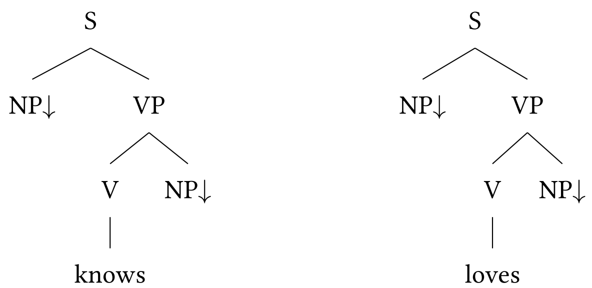


Figure 23.12: Elementary trees for *knows* and *loves*

If one wants to use these trees/constructions for the analysis of (86), there are in principle two possibilities: One assumes that two complete sentences are coordinated or alternatively, one assumes that some nodes are shared in a coordinated structure.

(86) He knows and loves this record.

Abeillé (2006) has shown that it is not possible to capture all the data if one assumes that cases of coordination such as those in (86) always involve the coordination of two complete clauses. It is also necessary to allow for lexical coordination of the kind we saw in Steedman's analysis. Sarkar & Joshi (1996) develop a TAG analysis in which nodes are shared in coordinate structures. The analysis of (86) can be seen in Figure 23.13 on the following page. The subject and object nodes are only present once in this figure. The S node of both elementary trees equally dominate the *he* NP. In the same way, the object NP node belongs to both VPs. The conjunction connects two verbs indicated by the thick lines. Sarkar and Joshi provide an algorithm that determines which nodes are to be shared. The structure may look strange at first, but for TAG purposes, it is not the derived tree but rather the derivation tree that is important, since this is the one that is used to calculate the semantic interpretation. The authors show that the derivation trees for the example under discussion and even more complex examples can be constructed correctly.

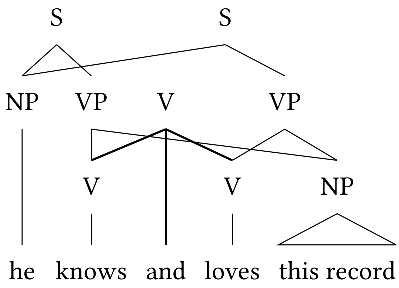


Figure 23.13: TAG analysis of *He knows and loves this record.*

In theories such as HPSG and LFG where structure building is, as in Categorial Grammar, driven by valency, the above sentence is unproblematic: Both verbs are conjoined and then the combination behaves like a simple verb. The analysis of this is given in Figure 23.14. This analysis is similar to the Categorial Grammar analysis in Figure 23.10.[42] With Goldberg's plugging analysis one could also adopt this approach to coordination:

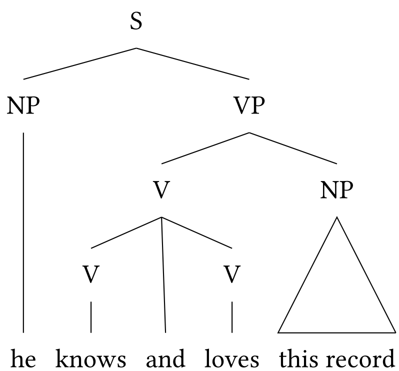


Figure 23.14: Selection-based analysis of *He knows and loves this record.* in tree notation

Here, *knows* and *loves* would first be plugged into a coordination construction and the result would then be plugged into the transitive construction. Exactly how the semantics of *knows and loves* is combined with that of the transitive construction is unclear since the meaning of this phrase is something like *and′*(*know′*(x, y), *love′*(x, y)), that is, a

[42] A parallel analysis in Dependency Grammar is possible as well. Tesnière's original analysis was different though. See Section 11.6.2.1 for discussion.

complex event with at least two open argument slots x and y (and possibly additionally an event and a world variable depending on the semantic theory that is used). Goldberg would probably have to adopt an analysis such as the one in Figure 23.13 in order to maintain the plugging analysis.

Croft would definitely have to adopt the TAG analysis since the verb is already present in his constructions. For the example in (83), both Goldberg and Croft would have to draw from the TAG analysis in Figure 23.15.

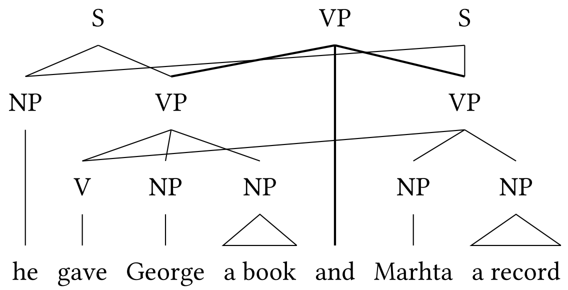


Figure 23.15: TAG analysis of *He gave George a book and Martha a record.*

The consequence of this is that one requires discontinuous constituents. Since coordination allows a considerable number of variants, there can be gaps between all arguments of constructions. An example with a ditransitive verb is given in (87):

(87) He gave George and sent Martha a record.

See Crysmann (2008) and Beavers & Sag (2004) for HPSG analyses that assume discontinuous constituents for particular coordination structures.

The result of these considerations is that the argument that particular elements occur next to each other and this occurrence is associated with a particular meaning is considerably weakened. What competent speakers acquire is actually the knowledge that heads must occur with their arguments somewhere in the utterance and the requirements of the heads involved have to somehow be satisfied altogether (θ Criterion, coherence/completeness, empty SUBCAT list). The heads themselves must not necessarily occur directly adjacent to their arguments. See the discussion in Section 16.3 about pattern-based models of language acquisition.

The semantics construction for complex structure such as those in Figure 23.15 is by no means trivial. In TAG, there is the derivation tree in addition to the derived tree that can then be used to compute the semantic contribution of a linguistic object. Construction Grammar does not have this separate level of representation. The question of how the meaning of the sentences discussed here is derived from their component parts still remains an open question for phrasal approaches.

Concluding the section on language acquisition, we assume that a valence representation is the result of language acquisition, since this is necessary for establishing the dependency relations in various possible configurations in an utterance. See also Behrens (2009: 439) for a similar conclusion.

23.7 Arguments from psycho- and neurolinguistics

This section has two parts: in the first part we compare approaches that assume that valence alternations are modeled by lexical rules, underspecification, or disjunctions with phrasal approaches. In the second part we discuss approaches to light verb constructions.

23.7.1 Lexical rules vs. phrasal constructions

Goldberg (1995: Section 1.4.5) uses evidence from psycholinguistic experiments to argue against lexical approaches that use lexical rules to account for argument structure alternations: Carlson & Tanenhaus (1988) showed that sentences with true lexical ambiguity like those in (88) and sentences with two verbs with the same core meaning have different processing times.

(88) a. Bill set the alarm clock onto the shelf.
     b. Bill set the alarm clock for six.

(89) a. Bill loaded the truck onto the ship.
     b. Bill loaded the truck with bricks.

Errors due to lexical ambiguity cause a bigger increase in processing time than errors in the use of the same verb. Experiments showed that there was a bigger difference in processing times for the sentences in (88) than for the sentences in (89). The difference in processing times between (89a) and (89b) would be explained by different preferences for phrasal constructions. In a lexicon-based approach one could explain the difference by assuming that one lexical item is more basic, that is, stored in the mental dictionary and the other is derived from the stored one. The application of lexical rules would be time consuming, but since the lexical items are related, the overall time consumption is smaller than the time that is needed to process two unrelated items (Müller 2002a: 405).

Alternatively one could assume that the lexical items for both valency patterns are the result of lexical rule applications. As with the phrasal constructions, the lexical rules would have different preferences. This shows that the lexical approach can explain the experimental results as well, so that they do not force us to prefer phrasal approaches.

Goldberg (1995: 18) claims that lexical approaches have to assume two variants of *load* with different meaning and that this would predict that *load* alternations would behave like two verbs that really have absolutely different meanings. The experiments discussed above show that such predictions are wrong and hence lexical analyses would be falsified. However, as was shown in Müller (2010a: Section 11.11.8.2), the argumentation contains two flaws: Let's assume that the construction meaning of the construction that licenses (89a) is C_1 and the construction meaning of the construction that licenses (89b) is C_2. Under such assumptions the semantic contribution of the two lexical items in the lexical analysis would be (90). load(…) is the contribution of the verb that would be assumed in phrasal analyses.

(90) a. load (onto): $C_1 \wedge$ load(…)
b. load (with): $C_2 \wedge$ load(…)

(90) shows that the lexical items partly share their semantic contribution. We hence predict that the processing of the dispreferred argument realization of *load* is simpler than the dispreferred meaning of *set*: in the latter case a completely new verb has to be activated while in the first case parts of the meaning are activated already. (See also Croft (2003: 64–65) for a brief rejection of Goldberg's interpretation of the experiment that corresponds to what is said here)

Goldberg (1995: 107) argues against lexical rule-based approaches for locative alternations like (91), since according to her such approaches have to assume that one of the verb forms has to be the more basic form.

(91) a. He loaded hay onto the wagon.
b. He loaded the wagon with hay.

She remarks that this is problematic since we do not have clear intuitions on what is the basic and what the derived form. She argues that the advantage of phrasal approaches is that various constructions can be related to each other without necessitating the assumption that one of the constructions is more basic than the other. There are two phrasal patterns and the verb is used in one of the two patterns. This criticism can be addressed in two ways: First one could introduce two lexical types (for instance *onto-verb* and *with-verb*) into a type hierarchy. The two types correspond to two valency frames that are needed for the analysis of (91a) and (91b). These types can have a common supertype (*onto-with-verb*) which is relevant for all *spray*/*load* verbs. One of the subtypes or the respective lexical item of the verb is the preferred one. This corresponds to a disjunction in the lexicon, while the phrasal approach assumes a disjunction in the set of constructions.

A variant of this approach is to assume that the lexical description of *load* just contains the supertyp, that describes all *spray*/*load* verbs. Since Model Theoretic Approaches assume that all structures that are models of utterances contain only maximally specific types (see for instance King (1999) and Pollard & Sag (1994: 21)), it is sufficient to say about verbs like *load* that they are of type *onto-with-verb*. Since this type has exactly two subtypes, *load* has to be either *onto-verb* or *with-verb* in an actual model.[43]

A second option is to stick with lexical rules and to assume a single representation for the root of a verb that is listed in the lexicon. In addition one assumes two lexical rules that map this basic lexical item onto other items that can be used in syntax after being inflected. The two lexical rules can be described by types that are part of a type hierarchy

[43] This analysis does not allow one to specify verb specific preferences for one of the realization patterns since the lexicon contains the general type only.

and that have a common supertype. This would capture commonalities between the lexical rules. We therefore have the same situation as with phrasal constructions (two lexical rules vs. two phrasal constructions). The only difference is that the action is one level deeper in the lexical approach, namely in the lexicon (Müller 2002a: 405–406).

The argumentation with regard to the processing of resultative constructions like (92c) is parallel:

(92) a. He drinks.
b. He drinks the milk.
c. He drinks the pub empty.

When humans parse a sentence they build up structure incrementally. If one hears a word that is incompatible with the current hypothesis, the parsing process breaks down or the current hypothesis is revised. In (92c) *the pub* does not correspond to the normal transitive use of *drink*, so the respective hypothesis has to be revised. In the phrasal approach the resultative construction would have to be used instead of the transitive construction. In the lexical analysis the lexical item that is licensed by the resultative lexical rule would have to be used rather than the bi-valent one. The building of syntactic structure and lexicon access in general place different demands on our processing capacities. However, when (92c) is parsed, the lexical items for *drink* are active already, we only have to use a different one. It is currently unclear to us whether psycholinguistic experiments can differentiate between the two approaches, but it seems to be unlikely.

23.7.2 Light verbs

Wittenberg, Jackendoff, Kuperberg, Paczynski, Snedeker & Wiese (2014) report on a number of experiments that test predictions that are made by various approaches to light verb constructions. (93a) shows a typical light verb construction: *take* is a light verb that is combined with the nominal that provides the main predication.

(93) a. take a walk to the park
b. walk to the park

Wittenberg & Piñango (2011) examined two psychologically plausible theories of light verb constructions. The phrasal approach assumes that light verb constructions are stored objects associated with semantics (Goldberg 2003b). The alternative compositional view assumes that the semantics is computed as a fusion of the semantics of the event noun and the semantics of the light verb (Grimshaw 1997; Butt 2003; Jackendoff 2002; Culicover & Jackendoff 2005; Müller 2010b; Beavers et al. 2008). Since light verb constructions are extremely frequent (Piñango, Mack & Jackendoff 2006; Wittenberg & Piñango 2011: 399), the phrasal approaches that assume that light verb constructions are stored items with the object and verb fixed predict that light verb constructions should be retrievable faster than non-light verb constructions like (94) (Wittenberg & Piñango 2011: 396).

(94) take a frisbee to the park

This is not the case. As Wittenberg and Piñango found, there is no difference in processing at the licensing condition (the noun in VO languages like English and the verb in OV languages like German).

However, Wittenberg & Piñango (2011) found an increased processing load 300ms *after* the light verb construction is processed. The authors explain this by assuming that semantic integration of the noun with the verbal meaning takes place after the syntactic combination. While the syntactic combination is rather fast, the semantic computation takes additional resources and this is measurable at 300ms. The verb contributes aspectual information and integrates the meaning of the nominal element. The semantic roles are fused. The resource consumption effect would not be expected if the complete light verb construction were a stored item that is retrieved together with the complete meaning (p. 404). We can conclude that Wittenberg and Piñango's results are compatible with the lexical proposal, but are incompatible with the phrasal view.

23.7.3 Arguments from Neurolinguistics

Pulvermüller, Cappelle & Shtyrov (2013) discuss neurolinguistic facts and relate them to the CxG view on grammar theory. One important finding is that deviant words (lexical items) cause brain responses that differ in polarity from brain responses on incorrect strings of words, that is, syntactic combinations. This suggests that there is indeed an empirical basis for deciding the issue.

Concerning the standard example of the caused motion construction in (95) the authors write the following:

(95) She sneezed the foam off the cappuccino.[44]

> this constellation of brain activities may initially lead to the co-activation of the verb *sneeze* with the DCNAs for *blow* and thus to the sentence mentioned. Ultimately, such co-activation of a one-place verb and DCNAs associated with other verbs may result in the former one-place verb being subsumed into a three-place verb category and DCNA set, a process which arguably has been accomplished for the verb *laugh* as used in the sequence *laugh NP off the stage.* (Pulvermüller, Cappelle & Shtyrov 2013)

A DCNA is a discrete combinatorial neuronal assembly. Regarding the specifics of DCNAs the authors write that

> Apart from linking categories together, typical DCNAs establish a temporal order between the category members they bind to. DCNAs that do not impose temporal order (thus acting, in principle, as AND units for two constituents) are thought to join together constituents whose sequential order is free or allow for scrambling. (Pulvermüller, Cappelle & Shtyrov 2013)

[44] Goldberg (2006: 42)

We believe that this view is entirely compatible with the lexical view outlined above: the lexical item or DCNA requires certain arguments to be present. A lexical rule that relates an intransitive verb to one that can be used in the caused motion construction is an explicit representation of what it means to activate the valence frame of *blow*.

The authors cite earlier work (Cappelle, Shtyrov & Pulvermüller 2010) and argue that particle verbs are lexical objects, admitting for a discontinuous realization of particle verbs despite their lexical status (p. 21). They restrict their claim to frequently occurring particle verbs. This claim is of course compatible with our assumptions here, but the differences in brain behavior are interesting when it comes to fully productive uses of particle verbs. For instance any semantically appropriate mono-valent verb in German can be combined with the aspectual particle *los*: *lostanzen* 'start to dance', *loslachen* 'start to laugh', *lossingen* 'start to sing', …. Similarly, the combination of mono-valent verbs with the particle *an* with the reading *directed-towards* is also productive. *anfahren* 'drive towards', *anlachen* 'laugh in the direction of', *ansegeln* 'sail towards', …(see Stiebels (1996) on various productive patterns). The interesting question is how particle verbs behave that follow these patterns but occur with low frequency. This is still an open question as far as the experimental evidence is concerned, but as we argue below lexical proposals to particle verbs as the one suggested by Müller (2003c) are compatible with both possible outcomes.

Summarizing the discussion so far, lexical approaches are compatible with the accumulated neurobiological evidence and as far as particle verbs are concerned they seem to be better suited than the phrasal proposals by Booij (2002: Section 2) and Blom (2005) (See Section 23.5.1 for discussion). However, in general it remains an open question what it means to be a discontinuous lexical item. The idea of discontinuous words is pretty old (Wells 1947), but there have not been many formal accounts of this idea. Nunberg, Sag & Wasow (1994) suggest a representation in a linearization-based framework of the kind that was proposed by Reape (1994) and Kathol (1995: 244–248) and Crysmann (2002) worked out such analyses in detail. Kathol's lexical item for *aufwachen* 'to wake up' is given in (96):

(96) *aufwachen* (following Kathol 1995: 246):

$$\left[\begin{array}{l} \ldots|\textsc{head}\ \boxed{1}\ \textit{verb} \\ \ldots|\textsc{vcomp}\ \langle\rangle \\ \textsc{dom}\ \left\langle \left[\begin{array}{l} \langle\ \textit{wachen}\ \rangle \\ \ldots|\textsc{head}\ \boxed{1} \\ \ldots|\textsc{vcomp}\ \langle\ \boxed{2}\ \rangle \end{array}\right] \right\rangle \bigcirc \left\langle \left[\begin{array}{l} \langle\ \textit{auf}\ \rangle \\ \textsc{synsem}\ \boxed{2} \left[\ldots|\textsc{head} \left[\begin{array}{l} \textsc{flip}\ - \\ \textit{sepref} \end{array}\right] \right] \\ \textit{vc} \end{array}\right] \right\rangle \end{array}\right]$$

The lexical representation contains the list-valued feature DOM that contains a description of the main verb and the particle (see Section 11.7.2.2 for details). The DOM list is a list that contains the dependents of a head. The dependents can be ordered in any order provided no linearization rule is violated (Reape 1994). The dependency between the particle and the main verb was characterized by the value of the VCOMP feature, which is a valence feature for the selection of arguments that form a complex predicate with their head. The shuffle operator ◯ concatenates two lists without specifying an order between the elements of the two lists, that is, both *wachen*, *auf* and *auf*, *wachen* are possible. The little marking *vc* is an assignment to a topological field in the clause.

Müller (2007c) criticized such linearization-based proposals since it is unclear how analyses that claim that the particle is just linearized in the domain of its verb can account for sentence like (97), in which complex syntactic structures are involved. German is a V2 language and the fronting of a constituent into the position before the finite verb is usually described as some sort of nonlocal dependency, that is, even authors who assume linearization-based analyses do not assume that the initial position is filled by simple reordering of material (Kathol 2000; Müller 1999a; 2002a; Bjerre 2006).[45]

(97) a. [$_{vf}$ [$_{mf}$ Den Atem] [$_{vc}$ an]] hielt die ganze Judenheit.[46]
the breath PART held the whole Jewish.community
'The whole Jewish community held their breath.'

b. [$_{vf}$ [$_{mf}$ Wieder] [$_{vc}$ an]] treten auch die beiden Sozialdemokraten.[47]
again PART kick also the two social.democrats
'The two Social Democrates are also running for office again.'

c. [$_{vf}$ [$_{vc}$ Los] [$_{nf}$ damit]] geht es schon am 15. April.[48]
PART there.with went it already at.the 15 April
'It already started on April the 15th.'

The conclusion that has to be drawn from examples like (97) is that particles interact in complex ways with the syntax of sentences. This is captured by the lexical treatment that

[45] Kathol (1995: Section 6.3) suggested such an analysis, but later changed his view. Wetta (2011) also assumes a purely linearization-based approach. He assumes that sentences in which multiple constituents are fronted (Müller 2003c) are analyzed in such a way that more than one linearization object are inserted as one single object into the position before the finite verb. This fails to account for multiple frontings that cross a clause boundary as in the examples in (i) discussed by Fanselow (1993: 67):

(i) a. Der Maria das Buch wenn du denkst daß du geben darfst bist du schön blöd.
the Maria the book if you think that you give be.allowed.to are you pretty stupid
'You are pretty stupid if you think you are allowed to give Maria the book.'

b. Der Maria einen Ring glaub ich nicht daß er je schenken wird.
the Maria a ring believe I not that he ever give.as.apresent will
'I do not believe that he will ever give Maria a ring as a present.'

If such sentences are to be analyzed as verb second sentences involving a dislocation mechanism – as also assumed by Wetta for non-local extraction (p. 265), there has to be a connection between a single element in the embedded clause and the fronted constituents *der Maria das Buch* and *der Maria einen Ring*. However, no such single projection exists in linearization-based proposals, since *der Maria* is the dative object and *das Buch* is the accusative object and the two do not form a constituent on the level of structure that is relevant for extraction. In earlier work Müller (2000b) suggested two nonlocal dependencies for the analysis of multiple frontings, but this was revised later and superseded by an analysis that integrates nicely into the rest of German grammar (Müller 2005b;c; 2015b).

[46] Lion Feuchtwanger, *Jud Süß*, p. 276, quoted from Grubačić (1965: 56).

[47] taz, bremen, 24.05.2004, p. 21.

[48] taz, 01.03.2002, p. 8.

was suggested in Müller (2002a: Chapter 6) and Müller (2003c): The main verb selects for the verbal particle. By assuming that *wachen* selects for *auf* the tight connection between verb and particle is represented.[49] Such a lexical analysis provides an easy way to account for fully intransparent particle verbs like *an-fangen* 'to begin'. However, we also argued for a lexical treatment of transparent particle verbs like *losfahren* 'to start to drive' and *jemanden/etwas anfahren* 'drive directed towards somebody/something'. The analysis involves a lexical rule that licenses a verbal item that selects for an adjunct particle. The particles *an* and *los* can modify verbs and contribute arguments (in the case of *an*) and the particle semantics. This analysis can be shown to be compatible with the neuro-mechanical findings: if it is the case that even transparent particle verb combinations with low frequency are stored, then the rather general lexical rule that was suggested by Müller is the generalization of the relation between a large amount of lexical particle verb items and their respective main verb. The individual particle verbs would be special instantiations that have the form of the particle specified as it is also the case for non-transparent particle verbs like *anfangen*. If it should turn out that productive particle verb combinations with particle verbs of low frequency cause syntactic reflexes in the brain, this could be explained as well: The lexical rule licenses an item that selects for an adverbial element. This selection would then be seen as parallel to the relation between the determiner and the noun in the NP *der Mut* 'the courage', which Cappelle et al. (2010: 191) discuss as an example of a syntactic combination. Note that Müller's analysis is also compatible with another observation made by Shtyrov, Pihko & Pulvermüller (2005): Morphological affixes also cause the lexical reflexes. In Müller's analysis the stem of the main verb is related to another stem that selects for a particle. This stem can be combined with (derivational and inflectional) morphological affixes causing the lexical activation pattern in the brain. After this combination the verb is combined with the particle and the dependency can be either a lexical or a syntactic one, depending on the results of the experiments to be carried out. The analysis is compatible with both results.

Note that Müller's analysis allows the principle of lexical integrity to be maintained. We therefore do not follow (Cappelle et al. 2010: 198), who claim that they *provide proof that potentially separable multi-word items can nonetheless be word-like themselves, and thus against the validity of a once well-established linguistic principle, the lexical integrity Principle.* We agree that non-transparent particle verbs are multi-word lexemes, but the existence of multi-word lexemes does not show that syntax has access to the word internal morphological structure. The parallel between particle verbs and clearly phrasal idioms was discussed in Müller (2002a;c) and it was concluded that idiom-status is irrelevant for the question of wordhood. Since the interaction of clearly phrasal idioms with derivational morphology as evidenced by examples like (98) did not force grammarians to give up on lexical integrity, it can be argued that particle verbs are not convincing

[49] Cappelle et al. (2010: 197) write: *the results provide neurophysiological evidence that phrasal verbs are lexical items. Indeed, the increased activation that we found for existing phrasal verbs, as compared to infelicitous combinations, suggests that a verb and its particle together form one single lexical representation, i. e. a single lexeme, and that a unified cortical memory circuit exists for it, similar to that encoding a single word* We believe that Müller's analysis is compatible with this statement.

evidence for giving up the lexical integrity Principle either.[50]

(98) a. Er hat ins Gras gebissen.
he has in.the gras bit
'He bit the dust.'

b. „Heath Ledger“ kann ich nicht einmal schreiben, ohne dass mir sein
Heath Ledger can I not even write without that me his
ins Gras-Gebeiße wieder so wahnsinnig leid tut[51]
in.the grass.biting again so crazy sorrow does
'I cannot even write “Heath Ledger” without being sad again about his repeated biting the dust.'

The example in (98b) involves the discontinuous derivation with the circumfix *Ge- -e* (Lüdeling 2001: Section 3.4.3; Müller 2002a: 324–327, 372–377; Müller 2003c: Section 2.2.1, Section 5.2.1). Still the parts of the idiom *ins Gras beiß-* are present and with them the idiomatic reading. See Sag (2007) for a lexical analysis of idioms that can explain examples like (98).

So, while we think that it is impossible to distinguish phrasal and lexical approaches for phenomena where heads are used with different valence pattern (Section 23.7.1), there seem to be ways to test whether patterns with high frequency and strong collocations should be analyzed as one fixed chunk of material with a fixed form and a fixed meaning or whether this should be analyzed compositionally.

23.8 Arguments from statistical distribution

In this section, we want to look at arguments from statistics that have been claimed to support a phrasal view. We first look at data-oriented parsing, a technique that was successfully used by Bod (2009b) to model language acquisition and then we turn to the collostructional analysis by Stefanowitsch & Gries (2009). Lastly we argue that these distributional analyses cannot decide the question whether argument structure constructions are phrasal or lexical.

23.8.1 Unsupervised Data-Oriented Parsing

In Section 13.8.3, we saw Rens Bod's approach to the structuring of natural language utterances. If one assumes that language is acquired from the input without innate knowledge, the structure that Bod extracts from the distribution of words would have to be the ones that children also learn (parts of speech, meaning and context would also have to be included). These structures would then also have to the ones assumed in linguistic theories. Since Bod does not have enough data, he carried out experiments under the assumption of binary-branching trees and for this reason, it is not possible to draw any

[50] However, see Booij (2009) for some challenges to lexical integrity.
[51] http://www.coffee2watch.at/egala. 23.03.2012

conclusions from his work about whether rules license flat or binary-branching structures. There will almost certainly be interesting answers to this question in the future. What can certainly not be determined in a distribution-based analysis is the exact node in the tree where meaning is introduced. Bod (2009a: 132) claims that his approach constitutes „a testable realization of CxG“ in the Goldbergian sense, however the trees that he can construct do not help us to decide between phrasal or lexical analyses or analyses with empty heads. These alternative analyses are represented in Figure 23.16.[52] The first

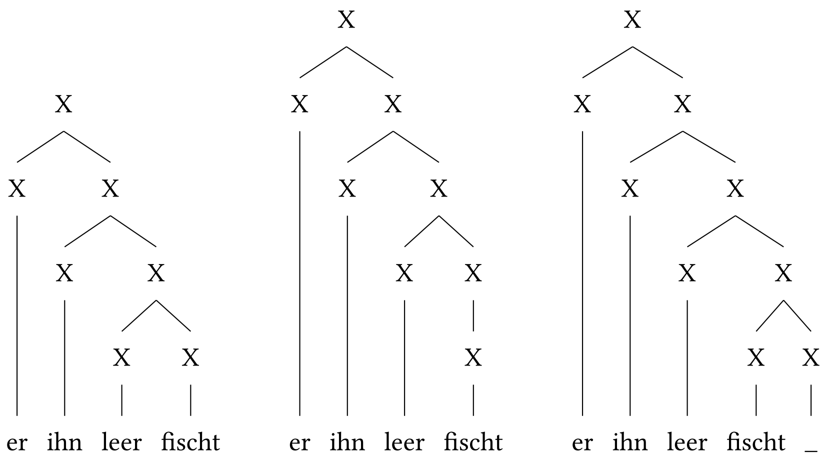
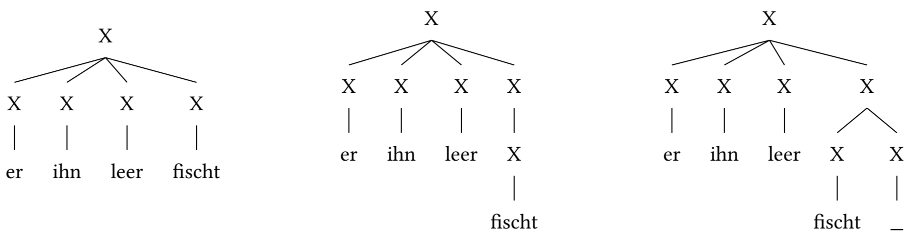


Figure 23.16: Three possible analyses for resultative construction: Holistic construction, lexical rule, empty head

figure stands for a complex construction that contributes the meaning as whole. The second figure corresponds to the analysis with a lexical rule and the third corresponds to the analysis with an empty head. A distributional analysis cannot decide between

[52] The discussion is perhaps easier to follow if one assumes flat structures rather than binary-branching ones.



The first figure corresponds to the Goldbergian view of phrasal constructions where the verb is inserted into the construction and the meaning is present at the top-most node. In the second figure, there is a lexical rule that provides the resultative semantics and the corresponding valency information. In the third analysis, there is an empty head that combines with the verb and has ultimately the same effect as the lexical rule.

these theoretical proposals. Distribution is calculated with references to words, what the words actually mean is not taken into account. As such, it is only possible to say that the word *fischt* 'fishes' occurs in a particular utterance, however it is not possible to see if this word contains resultative semantics or not. Similarly, a distribution analysis does not help one to distinguish between analyses with or without a lexical head. The empty head is not perceptible in the signal. It is a theoretical construct and, as we have seen in Section 21.5, it is possible to translate an analysis using an empty head into one with a lexical rule. For the present example, any argumentation for a particular analysis will be purely theory-internal.

Although U-DOP cannot help us to decide between analyses, there are areas of grammar for which these structures are of interest: Under the assumption of binary-branching structures, there are different branching possibilities depending on whether one assumes an analysis with verb movement or not. This means that although one does not see an empty element in the input, there is a reflex in statistically-derived trees. The left tree in Figure 23.17 shows a structure that one would expect from an analysis following Steedman (2000: 159) (see Section 8.3). The tree on the right shows a structure that would be expected from a GB-type verb movement analysis (see Section 3.2). But at present, there

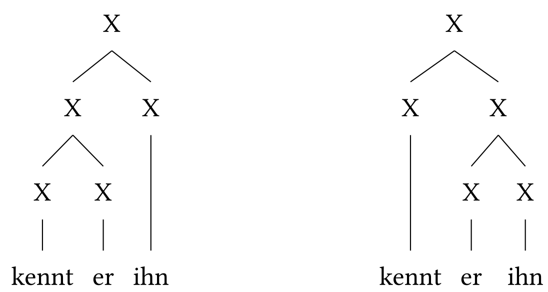


Figure 23.17: Structures corresponding to analysis with or without verb movement

is no clear finding in this regard (Bod, p. c. 2009). There is a great deal of variance in the U-DOP trees. The structure assigned to an utterance depends on the verb (Bod referring to the Wall Street Journal). Here, it would be interesting to see if this changes with a larger data sample. In any case, it would be interesting to look at how all verbs as well as particular verb classes behave. The U-DOP procedure applies to trees containing at least one word each. If one makes use of other parts of speech, the result is structures that correspond to the ones we have seen in the preceding chapters. Sub-trees would then not have two Xs as their daughters but rather NP and V, for example. It is still possible to do statistic work wit these kind of subtrees and as such one would have a generalization about all parts of speech and therefore a larger set of data to calculate structures for utterances such as those in Figure 23.17. I believe that there are further results in this area to be found in years to come.

Concluding this subsection, we contend that Bod's paper is a milestone in the Poverty of the Stimulus debate, but it does not and cannot show that a particular version of constructionist theories, namely the phrasal one, is correct.

23.8.2 Collostructions

Stefanowitsch & Gries (2009: Section 5) assume a plugging analysis: *words occur in (slots provided by) a given construction if their meaning matches that of the construction.* The authors claim that their *collostructional analysis has confirmed* [*the plugging analysis*] *from various perspectives.* Stefanowitsch and Gries are able to show that certain verbs occur more often than not in particular constructions, while other verbs never occur in the respective constructions. For instance, *give*, *tell*, *send*, *offer* and *show* are attracted by the Ditransitive Construction, while *make* and *do* are repelled by this construction, that is they occur significantly less often in this construction than what would be expected given the overall frequency of verbs in the corpus. Regarding this distribution the authors write:

> These results are typical for collexeme analysis in that they show two things. First, there are indeed significant associations between lexical items and grammatical structures. Second, these associations provide clear evidence for semantic coherence: the strongly attracted collexemes all involve a notion of 'transfer', either literally or metaphorically, which is the meaning typically posited for the ditransitive. This kind of result is typical enough to warrant a general claim that collostructional analysis can in fact be used to identify the meaning of a grammatical construction in the first place. (Stefanowitsch & Gries 2009: 943)

We hope that the preceding discussion made clear that the distribution of words in a corpus cannot be seen as evidence for a phrasal analysis. The corpus study shows that *give* usually is used with three arguments in a certain pattern that is typical for English (Subject Verb Object1 Object2) and that this verb forms a cluster with other verbs that have a transfer component in their meaning. The corpus data does not show whether this meaning is contributed by a phrasal pattern or by lexical entries that are used in a certain configuration.

23.9 Conclusion

The essence of the lexical view is that a verb is stored with a valence structure indicating how it combines semantically and syntactically with its dependents. Crucially, that structure is abstracted from the actual syntactic context of particular tokens of the verb. Once abstracted, that valence structure can meet other fates besides licensing the phrasal structure that it most directly encodes: it can undergo lexical rules that manipulate that structure in systematic ways; it can be composed with the valence structure of another predicate; it can be coordinated with similar verbs; and so on. Such an abstraction allows for simple explanations of a wide range of robust, complex linguistic phenomena. We have surveyed the arguments against the lexical valence approach, and in favor of a phrasal representation instead. We find the case for a phrasal representation of argument structure to be unconvincing: there are no compelling arguments in favor of such approaches, and they introduce a number of problems:

- They offer no account for the interaction of valence changing processes and derivational morphology.
- They offer no account for the interaction of valence changing processes and coordination of words.
- They offer no account for the iteration of valence changing processes.
- They overgenerate, unless a link between lexical items and phrasal constructions is assumed.
- They offer no account of distribution of arguments in partial fronting examples.

Assuming a lexical valence structure allows us to solve all the problems that arise for phrasal approaches.

23.10 Why (phrasal) constructions?

In previous sections, I have argued against assuming too much phrasality in grammatical descriptions. If one wishes to avoid transformations in order to derive alternative patterns from a single base structure, while still maintaining lexical integrity, then phrasal analyses become untenable for analyzing all those phenomena where changes in valency and derivational morphology interact. There are, however, some areas in which these two do not interact. In these cases, there is mostly a choice between analyses with silent heads and those with phrasal constructions. In this section, I will discuss some of these cases.

23.10.1 Verbless directives

Jacobs (2008) showed that there are linguistic phenomena, where it does not make sense to assume that there is a head in a particular group of words. These configurations are best described as phrasal constructions, in which the adjacency of particular constituents leads to a complete meaning that goes beyond the sum of its parts. Examples of the phenomena that are discussed by Jacobs are phrasal templates such as those in (99) and verbless directives as in (104):

(99) $\text{Pro}_{\text{+w,kaus/fin}}$ NP

a. Wozu Konstruktionen?
   why constructions
   'Why constructions?'

b. Warum ich?
   why I.NOM
   'Why me?'

(100) NP_{akk} $Y_{PP/A/Adv}$
Den Hut in der Hand (kam er ins Zimmer).
the hat in the hand came he into.the room
'(He came into the room) hat in hand.'

In (99), we are dealing with abbreviated questions:

(101) a. Wozu braucht man Konstruktionen? / Wozu sollte man Konstruktionen annehmen?
   to.what needs one constructions to.what should one constructions assume
   'Why do we need constructions?' / 'Why should we assume constructions?'

b.

c. Warum soll ich das machen? / Warum wurde ich ausgewählt? / Warum passiert mir sowas?
   why should I that do why was I chosen why happens me something.like.that
   'Why should I do that?' / 'Why was I chosen?' / 'Why do things like that happen to me?'

In (100b), a participle has been omitted:

(102) Den Hut in der Hand haltend kam er ins Zimmer.
the hat.ACC in the hand holding came he in.the room
'He came into the room hat in hand.'

Cases such as (100b) can be analyzed with an empty head that corresponds to *haltend* 'holding'. For (100a), on the other hand, one would require either a syntactic structure with multiple empty elements, or an empty head that selects both parts of the construction and contributes the components of meaning that are present in (101). If one adopts the first approach with multiple silent elements, then one would have to explain why these elements cannot occur in other constructions. For example, it would be necessary to assume an empty element corresponding to *man* 'one'/'you'. But such an empty element could never occur in embedded clauses since subjects cannot simply be omitted there:

(103) * weil dieses Buch gerne liest
because this book gladly reads
Intended: 'because he/she/it likes to read this book'

If one were to follow the second approach, one would be forced to assume an empty head with particularly odd semantics.

The directives in (104) and (105) are similarly problematic (see also Jackendoff & Pinker (2005: 220) for parallel examples in English):

(104) a. Her mit dem Geld / dem gestohlenen Geld!
here with the money the stolen money
'Hand over the (stolen) money!'

b. Weg mit dem Krempel / dem alten Krempel!
away with the junk the old junk
'Get rid off this (old) junk!'

c. Nieder mit den Studiengebühren / den sozialfeindlichen Studiengebühren!
down with the tuition.fees the antisocial tuition.fees
'Down with tuition fees!'

(105) a. In den Müll mit diesen Klamotten.
in the trash with these clothes
'Throw these clothes in the trash!'

b. Zur Hölle mit dieser Regierung.
to.the hell with this government
'To hell with this government!'

Here, it is also not possible to simply identify an elided verb. It is, of course, possible to assume an empty head that selects an adverb or a *mit* PP, but this would be *ad hoc*. Alternatively, it would be possible to assume that adverbs in (104) select the *mit*-PP. Here, one would have to disregard the fact that adverbs do not normally take any arguments. The same is true of Jacobs' examples in (105). For these, one would have to assume that *in* and *zur* 'to the' are the respective heads. Each of the prepositions would then have to select a noun phrase and a *mit* PP. While this is technically possible, it is as unattractive as the multiple lexical entries that Categorial Grammar has to assume for pied-piping constructions (see Section 8.6).

A considerably more complicated analysis has been proposed by Gereon Müller in 2009 at the lecture series *Algorithmen und Muster – Strukturen in der Sprache* at the FU Berlin (G. Müller 2011). He treats verbless directives as antipassive constructions. Antipassive constructions involve either the complete suppression of the direct object or its realization as an oblique element (PP). There can also be morphological marking on the verb. The subject is normally not affected by the antipassive but can, however, receive a different case in ergative case systems due to changes in the realization of the object. According to G. Müller, there is a relation between (106a) and (106b) that is similar to active-passive pairs:

(106) a. [dass] jemand diese Klamotten in den Müll schmeißt
that somebody these clothes in the trash throws

b. In den Müll mit diesen Klamotten!
in the rubbish with these clothes

An empty passive morpheme absorbs the capability of the verb to assign accusative (also see Section 3.4 on the analysis of the passive in GB theory). The object therefore has to be realized as a PP or not at all. It follows from Burzio's Generalization that as the accusative object has been suppressed, there cannot be an external argument. G. Müller assumes, like proponents of Distributed Morphology (e. g. Marantz (1997)), that lexical entries are inserted into complete trees post syntactically. The antipassive morpheme creates a feature bundle in the relevant tree node that is not compatible with German verbs such as *schmeißen* 'throw' and this is why only a null verb with the corresponding specifications can be inserted. Movement of the directional PP is triggered by mechanisms that cannot be discussed further here. The antipassive morpheme forces an obligatory reordering of the verb in initial position (to C, see Section 3.2). Filling the prefield is only possible in sentences where the C position is filled by a visible verb and this is why G. Müller's analysis can only derive verb-initial clauses. These are interpreted as imperatives or polar questions. Figure 23.18 gives the analysis of (106b). Budde (2010) and Maché (2010) note

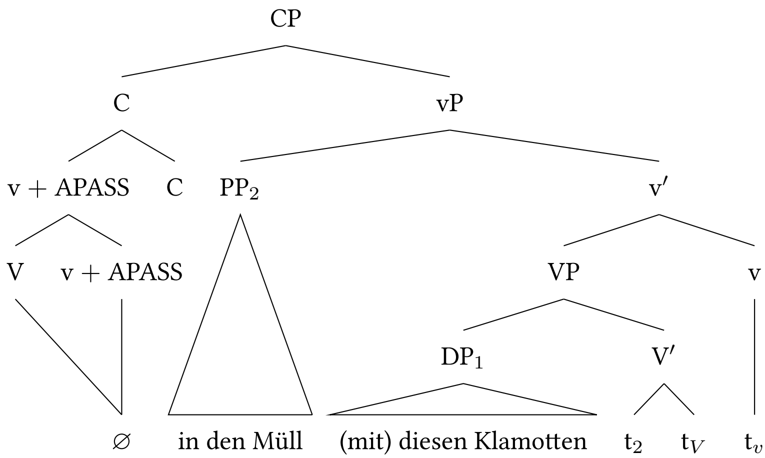


Figure 23.18: *In den Müll mit diesen Klamotten* 'in the trash with these clothes' as an antipassive following Gereon Müller (2009)

that the discussion of the data has neglected the fact that there are also interrogative variants of the construction:

(107) a. Wohin mit den Klamotten?
where.to with the clothes
'Where should the clothes go?'

b. Wohin mit dem ganzen Geld?
where.to with the entire money
'Where should all this money go?'

One therefore does not require the constraint that the prefield can only be filled if the C position is filled.

One major plus point of this analysis is that it derives the different sentence types that are possible with these kind of constructions: The V1-variants correspond to polar questions and imperatives, and the V2-variants with a question word correspond to *wh*-questions. A further consequence of the approach that was pointed out by Gereon Müller is that no further explanation is required for other interactions with the grammar. For example, the way in which the constructions interact with adverbs, follows from the analysis:

(108) a. Schmeiß den Krempel weg!
throw the junk away

b. Schmeiß den Krempel schnell weg!
throw the junk quickly away

c. ?* Schmeiß den Krempel sorgfältig weg!
throw the junk carefully away

(109) a. Weg mit dem Krempel!
away with the junk

b. Schnell weg mit dem Krempel!
quickly away with the junk

c. ?* Sorgfältig weg mit dem Krempel!
carefully away with the junk

Nevertheless one should still bear the price of this analysis in mind: It assumes an empty antipassive morpheme that is otherwise not needed in German. It would only be used in constructions of the kind discussed here. This morpheme is not compatible with any verb and it also triggers obligatory verb movement, which is something that is not known from any other morpheme that is used to form verb diatheses.

The costs of this analysis are of course less severe if one assumes that humans already have this antipassive morpheme anyway, that is, this morpheme is part of our innate Universal Grammar. But if one follows the argumentation from the earlier sections of this chapter, then one should only assume innate linguistic knowledge if there is no alternative explanation.

G. Müller's analysis can be translated into HPSG. The result is given in (110):

(110)
$$\begin{bmatrix} \textsc{rels} & \left\langle \begin{bmatrix} \textsc{event} & \boxed{2} \\ \multicolumn{2}{l}{\textit{imperative-or-interrogative}} \end{bmatrix} \right\rangle \oplus \Box \\ \textsc{lex-dtr} & \begin{bmatrix} \textsc{phon} & \langle\rangle \\ \textsc{ss}|\textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head}|\textsc{mod} & \textit{none} \\ \textsc{subcat} & \left\langle \text{XP}[\textsc{mod} \ldots \textsc{ind}\ \boxed{1}],\ (\text{PP}[\textit{mit}]_{\boxed{1}}) \right\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{ind} & \boxed{2} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textsc{event} & \boxed{2} \\ \textsc{patient} & \boxed{1} \\ \multicolumn{2}{l}{\textit{directive}} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \multicolumn{2}{l}{\textit{verb-initial-lr}} \end{bmatrix}$$

(110) contains a lexical entry for an empty verb in verb-initial position. *directive′* is a placeholder for a more general relation that should be viewed as supertype of all possible meanings of this construction. These subsumes both *schmeißen* and cases such as (111) that were pointed out to me by Monika Budde:

(111) Und mit dem Klavier ganz langsam durch die Tür!
and with the piano very slowly through the door
'Carry the piano very slowly through the door!'

Since only verb-initial and verb-second orders are possible in this construction, the application of the lexical rule for verb-initial position (see page 288) is obligatory. This can be achieved by writing the result of the application of this lexical rule into the lexicon, without having the object to which the rule should have applied actually being present in the lexicon itself. Koenig (1999: Section 3.4.2, 5.3) proposed something similar for English *rumored* 'it is rumored that …' and *aggressive*. There is no active variant of the verb *rumored*, a fact that can be captured by the assumption that only the result of applying a passive lexical rule is present in the lexicon. The actual verb or verb stem from which the participle form has been derived exists only as the daughter of a lexical rule but not as an independent linguistic object. Similarly, the verb * *agress* only exists as the daughter of a (non-productive) adjective rule that licenses *agressive* and a nominalization rule licensing *agression*.

The optionality of the *mit*-PP is signaled by the brackets in (110). If one adds the information inherited from the type *verb-initial-lr* under SYNSEM, then the result is (112).

(112)
$$
\begin{bmatrix}
\text{SYNSEM|LOC} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ \text{INITIAL} & + \\ \text{DSL} & \textit{none} \\ \textit{verb} & \end{bmatrix} \\ \text{SUBCAT} & \left\langle \begin{bmatrix} \text{LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{DSL} & \boxed{3} \\ \textit{verb} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\
\text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{2} \\ \textit{imperative-or-interrogative} & \end{bmatrix} \right\rangle \oplus \boxed{4} \\
\text{LEX-DTR} & \begin{bmatrix} \text{PHON} & \langle\rangle \\ \text{SS|LOC}\ \boxed{3} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD|MOD} & \textit{none} \\ \text{SUBCAT} & \left\langle \text{XP[MOD ... IND } \boxed{1}\text{]}, (\text{PP[}\textit{mit}\text{]}_{\boxed{1}}) \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{2} \\ \text{RELS}\ \boxed{4} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{2} \\ \text{PATIENT} & \boxed{1} \\ \textit{directive} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\textit{verb-initial-lr} &
\end{bmatrix}
$$

The valency properties of the empty verb in (112) are to a large extent determined by the lexical rule for verb-initial order: The V1-LR licenses a verbal head that requires a VP to its right that is missing a verb with the local properties of the LEX-DTR ($\boxed{3}$).

Semantic information dependent on sentence type (assertion, imperative or question) is determined inside the V1-LR depending on the morphological make-up of the verb and the SLASH value of the selected VP (see Müller (2007b: Section 10.3) and Müller (2014e)). Setting the semantics to *imperative-or-interrogative* rules out *assertion* as it occurs in V2-clauses. Whether this type is resolved in the direction of *imperative* or *interrogative* is ultimately decided by further properties of the utterance such as intonation or the use of interrogative pronouns.

The valency of the lexical daughters in (112) as well as the connection to the semantic role (the linking to the patient role) are simply stipulated. Every approach has to stipulate that an argument of the verb has to be expressed as a *mit*-PP. Since there is no antipassive in German, the effect that could be otherwise achieved by an antipassive lexical rule in (112) is simply written into the LEX-DTR of the verb movement rule.

The SUBCAT list of LEX-DTR contains a modifier (adverb, directional PP) and the *mit*-PP. This *mit*-PP is co-indexed with the patient of *directive′* and the modifier refers to the referent of the *mit*-PP. The agent of *directive′* is unspecified since it depends on the context (speaker, hearer, third person).

This analysis is shown in Figure 23.19 on the following page. Here, V[LOC $\boxed{2}$] corresponds to LEX-DTR in (112). The V1-LR licenses an element that requires a maximal

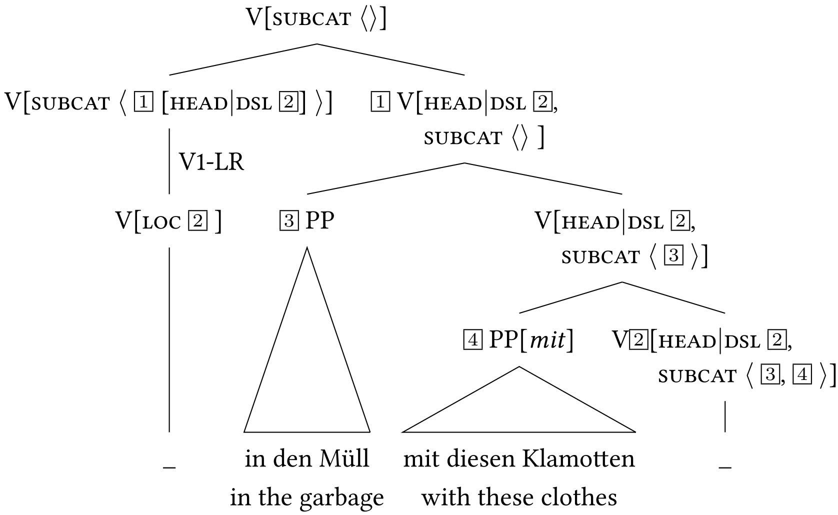


Figure 23.19: HPSG variant of the analysis of *In den Müll mit diesen Klamotten!/?*

verb projection with that exact DSL value $\boxed{2}$. Since DSL is a head feature, the information is present along the head path. The DSL value is identified with the LOCAL value ($\boxed{2}$ in Figure 23.19) in the verb movement trace (see page 289). This ensures that the empty element at the end of sentence has exactly the same local properties that the LEX-DTR in (112) has. Thus, both the correct syntactic and semantic information is present on the verb trace and structure building involving the verb trace follows the usual principles. The structures correspond to the structures that were assumed for German sentences in Chapter 9. Therefore, there are the usual possibilities for integrating adjuncts. The correct derivation of the semantics, in particular embedding under imperative or interrogative semantics, follows automatically (for the semantics of adjuncts in conjunction with verb position, see Müller (2007b: Section 9.4)). Also, the ordering variants with the *mit*-PP preceding the direction (111) and direction preceding the *mit*-PP (106b) follow from the usual mechanisms.

If one rejects the analyses discussed up to this point, then one is only really left with phrasal constructions or dominance schemata that connect parts of the construction and contribute the relevant semantics. Exactly how one can integrate adjuncts into the phrasal construction in a non-stipulative way remains an open question, however, there are already some initial results by Jakob Maché (2010) that suggest that directives can still be sensibly integrated into the entire grammar provided an appropriate phrasal schema is assumed.

23.10.2 Serial verbs

There are languages with so-called serial verbs. For example, it is possible to form sentences in Mandarin Chinese where there is only one subject and several verb phrases. There are multiple readings depending on the distribution of aspect marking inside the VP: If the first VP contains a perfect marker, then we have the meaning 'VP1 in order to do/achieve VP2' (113a). If the second VP contains a perfect marker, then the entire construction means 'VP2 because VP1' (113b) and if the first VP contains a durative marker and the verb and contains the verb *hold* or *use*, then the entire construction means 'VP2 using VP1' (113c).

(113) a. Ta1 qu3 le qian2 qu4 guang1jie1.
he withdraw PRF money go shop
'He withdrew money to go shopping.'

b. Ta1 zhu4 Zhong1guo2 xue2 le Han4yu3.
he live China learn PRF Chinese
'He learned Chinese because he lived in China.'

c. Ta1 na2 zhe kuai4zi chi1 fan4.
he take DUR chopsticks eat food
'He eats with chopsticks.'

If we consider the sentences, we only see two adjacent VPs. The meanings of the entire sentences, however, contain parts of meaning that go beyond the meaning of the verb phrases. Depending on the kind of aspect marking, we arrive at different interpretations with regard to the semantic combination of verb phrases. As can be seen in the translations, English sometimes uses conjunctions in order to express relations between clauses or verb phrases.

There are three possible ways to capture these data:

1. One could claim that speakers of Chinese simply deduce the relation between the VPs from the context,
2. one could assume that there are empty heads in Chinese corresponding to *because* or *to*, or
3. one could assume a phrasal construction for serial verbs that contributes the correct semantics for the complete meaning depending on the aspect marking inside the VPs.

The first approach is unsatisfactory because the meaning does not vary arbitrarily. There are grammaticalized conventions that should be captured by a theory. The second solution has a stipulative character and thus, if one wishes to assume empty elements, only the third solution remains. Müller & Lipenkova (2009) have presented a corresponding analysis.

23.10.3 Relative and interrogative clauses

Sag (1997) develops a phrasal analysis of relative clauses as have Ginzburg & Sag (2000) for interrogative clauses. These sentences consist of a fronted phrase and a clause or a verb phrase missing the fronted phrase. The fronted phrase contains a relative or interrogative pronoun.

(114)
a. the man [who] sleeps
b. the man [who] we know
c. the man [whose mother] visited Kim
d. a house [in which] to live

(115)
a. I wonder [who] you know.
b. I want to know [why] you did this.

The GB analysis of relative clauses is given in Figure 23.20. In this analysis, an empty head is in the C position and an element from the IP is moved to the specifier position.

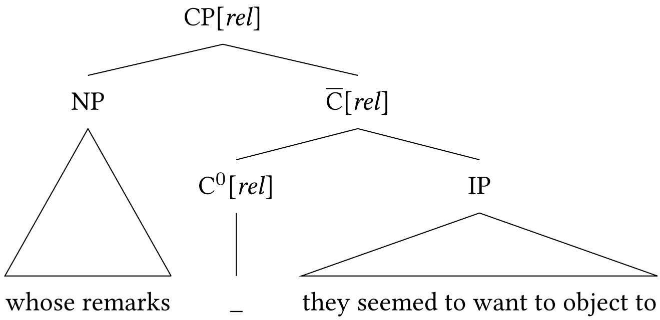


Figure 23.20: Analysis of relative clauses in GB theory

The alternative analysis shown in Figure 23.21 on the next page involves combining the subparts directly in order to form a relative clause. Borsley (2006) has shown that one would require six empty heads in order to capture the various relative clauses possible in English, if one would want to analyze them lexically. These heads can be avoided and replaced by corresponding schemata. A parallel argument can also be found in Webelhuth (2011) for German: Grammars of German would also have to assume six empty heads for the relevant types of relative clause.

Unlike the resultative constructions that we have already discussed, there is no variability among interrogative and relative clauses with regard to the order of their parts. There are no changes in valency and no interaction with derivational morphology. Thus, nothing speaks against a phrasal analysis. If one wishes to avoid the assumption of empty heads, then one should opt for the analysis of relative clauses by Sag, or the variant in Müller (1999a: Chapter 10; 2007b: Chapter 11). The latter analysis does without a

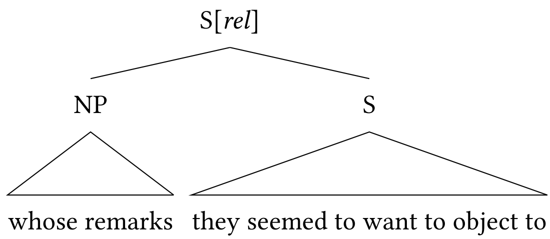


Figure 23.21: Analysis of relative clauses in HPSG following Sag (1997)

special schema for noun-relative clause combinations since the semantic content of the relative clause is provided by the relative clause schema.

Sag (2010) discusses long-distance dependencies in English that are subsumed under the term *wh*-movement in GB theory and the MP. He shows that this is by no means a uniform phenomenon. He investigates *wh*-questions (116), *wh*-exclamatives (117), Topicalization (118), *wh*-relative clauses (119) and *the*-clauses (120):

(116) a. How foolish is he?
  b. I wonder *how foolish he is*.

(117) a. What a fool he is!
  b. It's amazing *how odd it is*.

(118) The bagels, I like.

(119) a. I met the person *who they nominated*.
  b. I'm looking for a bank *in which to place my trust*.

(120) a. The more people I met, *the happier I became*.
  b. *The more people I met*, the happier I became.

These individual constructions vary in many respects. Sag lists the following questions that have to be answered for each construction:

- Is there a special *wh*-element in the filler daughter and, if so, what kind of element is it?
- Which syntactic categories can the filler daughters have?
- Can the head-daughter be inverted or finite? Is this obligatory?
- What is the semantic and/or syntactic category of the mother node?
- What is the semantic and/or syntactic category of the head-daughter?

- Is the sentence an island? Does it have to be an independent clause?

The variation that exists in this domain has to somehow be captured by a theory of grammar. Sag develops an analysis with multiple schemata that ensure that the category and semantic contribution of the mother node correspond to the properties of both daughters. The constraints for both classes of constructions and specific constructions are represented in an inheritance hierarchy so that the similarities between the constructions can be accounted for. The analysis can of course also be formulated in a GB-style using empty heads. One would then have to find some way of capturing the generalizations pertaining to the construction. This is possible if one represents the constraints on empty heads in an inheritance hierarchy. Then, the approaches would simply be notational variants of one another. If one wishes to avoid empty elements in the grammar, then the phrasal approach would be preferable.

23.10.4 The NPN construction

Jackendoff (2008) discusses the NPN constructions. Examples of this construction are given in (121):

(121) a. day by day, paragraph by paragraph, country by country
b. dollar for dollar, student for student, point for point
c. face to face, bumper to bumper
d. term paper after term paper, picture after picture
e. book upon book, argument upon argument

This construction is relatively restricted: articles and plural nouns are not allowed. The phonological content of the first noun has to correspond to that of the second. There are also similar constructions in German:

(122) a. Er hat Buch um Buch verschlungen.
he has book around book swallowed
'He swallowed (read) book after book.'
b. Zeile für Zeile[53]
line for line
'line by line'

Determining the meaning contribution of this kind of NPN construction is by no means trivial. Jackendoff suggests the meaning *many Xs in succession* as an approximation.

Jackendoff points out that this construction is problematic from a syntactic perspective since it is not straightforwardly possible to determine a head. It is also not clear what the structure of the remaining material is if one is working under assumptions of $\overline{\text{X}}$ Theory. If the preposition *um* were the head, then one would expect that it is combined with a NP, however this is not possible:

[53] *Zwölf Städte*. Einstürzende Neubauten. Fünf auf der nach oben offenen Richterskala, 1987.

(123) a. * Er hat dieses Buch um jenes Buch verschlungen.
he has this book around this book swallowed

b. * Er hat ein Buch um ein Buch verschlungen.
he has a book around a book swallowed

For these kind of structures, it would be necessary to assume that a preposition selects a noun to its right and, if it find this, it then requires a second noun of this exact form to its left. For N-*um*-N and N-*für*-N, it is not entirely clear what the entire construction has to do with the individual prepositions. It would be possible to pursue a lexical analysis at this point, however, the facts are different to those for resultative constructions: In resultative constructions, the semantics of simplex verbs plays a clear role. Furthermore, unlike with the resultative construction, the order of the component parts of the construction are fixed in the NPN construction. It is not possible to extract a noun or place the preposition in front of both nouns. Syntactically, the NPN combination with some prepositions behaves like an NP (Jackendoff 2008: 9):

(124) Student after/upon/*by student flunked.

This is also strange if one wishes to view the preposition as the head of the construction.

Instead of a lexical analysis, Jackendoff proposes the following phrasal construction for N-*after*-N combinations:

(125) Meaning: MANY X_is IN SUCCESSION [or however it is encoded]
Syntax: $[_{\text{NP}}\ N_i\ P_j\ N_i]$
Phonology: Wd_i after_j Wd_i

The entire meaning as well as the fact that the NPN has the syntactic properties of an NP would be captured on the construction level.

G. Müller (2011) proposed a lexical analysis of the NPN construction. He assumes that prepositions can have a feature REDUP. In the analysis of *Buch um Buch* 'book after book', the preposition is combined with the right noun *um Buch*. In the phonological component, reduplication of *Buch* is triggered by the REDUP feature, thereby yielding *Buch um Buch*. This analysis still suffers from the problems pointed out by Jackendoff: In order to derive the semantics of the construction, the semantics would have to be present in the lexical entry of the reduplicating preposition (or in a relevant subsequent component that interprets the syntax).

24 Universal Grammar and doing comparative linguistics without an a priori assumption of a (strong) UG

The following two sections deal with the tools that I believe to be necessary to capture generalizations and the way one can derive such generalizations.

24.1 Formal tools for capturing generalizations

In Chapter 13, it was shown that all the evidence that has previously been brought forward in favor of innate linguistic knowledge is in fact controversial. In some cases, the facts are irrelevant to the discussion and in other cases, they could be explained in other ways. Sometimes, the chains of argumentation are not logically sound or its axioms are not supported. In other cases, the argumentation is circular. As a result, the question of whether there is innate linguistic knowledge still remains unanswered. All theories that presuppose the existence of this kind of knowledge are making very strong assumptions. If one assumes, as Kayne (1994) for example, that all languages have the underlying structure [specifier [head complement]] and that movement is exclusively to the left, then, while it is possible to develop a very elegant system, the basic assumptions must be part of this innate linguistic knowledge since there are many languages for which there is no evidence for a specifier-head-complement structure.

As we have seen, there are number of alternative theories that are much more surface-oriented than some variants of Transformational Grammar. These theories often differ with regard to particular assumptions that have been discussed in the preceding sections. For example, there are differences in the treatment of long-distance dependencies that have led to a proliferation of lexical entries in Categorial Grammar. As has been shown by Jacobs (2008), Jackendoff (2008) and others, approaches such as Categorial Grammar that assume that every phrase must have a functor/head cannot explain certain constructions in a plausible way. Inheritance-based phrasal analyses that only list heads with a core meaning in the lexicon and have the constructions in which the heads occur determine the meaning of a complex expression, turn out to be incompatible with assumptions such as lexical integrity. We therefore need a theory that handles argument structure changing processes in the lexicon and still has some kind of phrase structure or relevant schemata. Some variants of GB/MP as well as LFG, HPSG, TAG and variants of CxG are examples of this kind of theory. Of these theories, only HPSG and some variants of CxG make use the same descriptive tools ((typed) feature descriptions) for roots, stems, words, lexical rules and phrases. By using a uniform description for all these objects, it is possible to formulate generalizations over the relevant objects. It is therefore possible to capture what particular words have in common with lexical rules or phrases. For example, the *-bar* derivation (1a) corresponds to a complex passive construction with a modal verb.

(1) a. Das Rätsel ist lösbar.
the puzzle is solveable

b. Das Rätsel kann gelöst werden.
the puzzle can solved be

'The puzzle can be solved'

It is also possible to capture cross-linguistic generalizations: Something that is inflection/derivation in one language, can be syntax in another.

It is possible to formulate principles that hold for both words and phrases and furthermore, it is possible to capture cross-linguistic generalizations or generalizations that hold for certain groups of languages. For example, languages can be divided into those with fixed constituent order and those with more free or completely free constituent order. The corresponding types can be represented with their constraints in a type hierarchy. Different languages can use a particular part of the hierarchy and also formulate different constraints for each of the types (seeAckerman & Webelhuth (1998: Section 9.2)). HPSG differs from theories such LFG and TAG in that phrases do not have a different status than words. This means that there are no special c-structures or tree structures. Descriptions of complex phrases simply have additional features that say something about their daughters. In this way, it is possible to formulate cross-linguistic generalizations about dominance schemata. In LFG, the number of c-structures per language is normally specified separately. Another advantage of consistent description is that one can capture similarities between words and lexical rules and also between words and phrases. For example, a complementizer such as *dass* shares a number of properties with simple or coordinated verb in initial position.

(2) a. [dass] Maria die Platte kennt und liebt
that Maria the record knows and loves

'that Maria knows and loves the record'

b. [Kennt und liebt] Maria die Platte?
knows and loves Maria the record

'Does Mary know and love the record?'

The differences between the linguistic objects mainly lies in the kind of phrase they select: The complementizer requires a sentence with a visible finite verb, whereas the verb in initial position requires a sentence without a visible finite verb.

In Section 9.1.5, a small part of an inheritance hierarchy was presented. The part presented contains types that probably play a role in the grammars of all natural languages:

There are head-argument combinations in every language. Without this kind of combinatorial operation, it would not be possible to establish a relation between two concepts. The ability to create relations, however, is one of the basic properties of language.

As well as more general types, the type hierarchy of a particular language contains language-specific types or those specific to a particular class of languages. All languages presumably have one and two-place predicates and for most languages (if not all), it makes sense to talk about verbs. It is then possible to talk about one and two-place verbs. Depending on the language, this can then be subdivided into intransitive and transitive. Constraints are formulated for the various types that can either hold generally or be language-specific.

In English, verbs have to occur before their complements and therefore have the INITIAL value +, whereas verbs in German have the INITIAL value − and it is the lexical rule for initial position that derives a verb with an INITIAL value +.

The differing setting of the INITIAL value for German and English is reminiscent of parameters from GB-Theory. There is one crucial difference, however: It is not assumed that a language learner sets the INITIAL value once and for all for all heads. The use of an INITIAL value is compatible with models of acquisition that assume that learners learn individual words with the positional properties. It is certainly possible for the respective words to exhibit different values for a particular feature. Generalizations about the position of entire word classes can then only be learned at a later point in the acquisition process.

A hierarchy analogous to those proposed by Croft (see Section 23.4.1) is given in Figure 24.1. For inflected words, there are the relevant roots in the lexicon. Examples of this

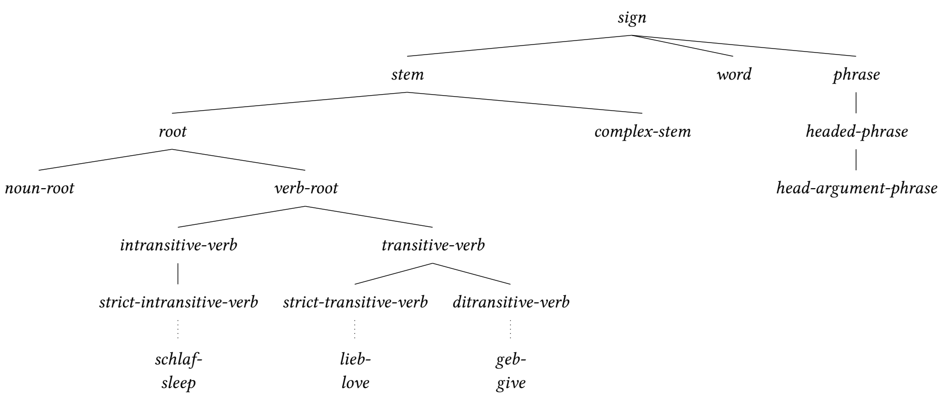


Figure 24.1: Section of an inheritance hierarchy with lexical entries and dominance schemata

are *schlaf-* 'sleep', *lieb-* 'love' and *geb-* 'give'. In Figure 24.1, there are different subtypes of *root*, the general type for roots: *intrans-verb* for intransitive verbs and *trans-verb* for transitive verbs. Transitive verbs can be further subdivided into strictly transitive verbs, those with nominative and accusative arguments, and ditransitive verbs, those with nominative and both accusative and dative arguments. The hierarchy above would have to of course be refined considerably as there are even further sub-classes for both transitive and intransitive verbs. For example, one can divide intransitive verbs into unaccusative and unergative verbs and even strictly transitive verbs would have to be divided into further sub-classes (see Welke (2009: Section 2)).

In addition to a type for roots, the above figure contains types for stems and words. Complex stems are complex objects that are derived from simple roots but still have to be inflected (*lesbar-* 'readable', *besing-* 'to praise'). Words are objects that do not inflect. Examples of these are the pronouns *er* 'he', *sie* 'she' etc. as well as prepositions. An inflected form can be formed from a verbal stem (*geliebt* 'loved', *besingt* 'praised'). Relations between inflected words and (complex) stems can be formed again using derivation rules. In this way, *geliebt* 'loved' can be recategorized as an adjective stem that must then be combined with adjectival endings (*geliebt-e*). The relevant descriptions of complex stems/words are subtypes of *complex-stem* or *word*. These subtypes describe the form that complex words such as *geliebte* must have. For a technical implementation of this, see Müller (2002a: Section 3.2.7). All words can be combined to phrases using dominance schemata. The hierarchy given here is of course by no means complete. There are a number of additional valency classes and one could also assume more general types that simply describe one, two and three-place predicates. These types are probably plausible for the description of other languages. Here, we are only dealing with a small section of the type hierarchy in order to have a comparison to the Croftian hierarchy: In Figure 24.1, there are no types sentence patterns with the form [Sbj IntrVerb], but rather types for lexical objects with a particular valency (V[SUBCAT ⟨ NP[*str*] ⟩]). Lexical rules can then be applied to the relevant lexical objects that license objects with another valency or introduce information about inflection. Complete words can be combined in the syntax with relatively general rules, for example in head-argument structures. The problems from which purely phrasal approaches suffer are thereby avoided. Generalizations about lexeme classes and the utterances that can be formed can still be captured in the hierarchy.

There are also principles in addition to inheritance hierarchies: The principles for the construction of meaning that were presented in Section 9.1.6 hold for all languages. The Case Principle that we also saw is a constraint that only applies to class of languages, namely nominative-accusative languages. Other languages have an ergative-absolutive system.

The assumption of innate linguistic knowledge is not necessary for the theory of language sketched here. As this entire discussion has shown, the question of whether this kind of knowledge exists has still not been answered conclusively. If it should turn out that this knowledge actually exists, the question arises of what exactly is innate. It would be a plausible assumption that the part of the inheritance hierarchy that is relevant for languages is innate together with the relevant principles. It could, however, also be the case that only a part of the more generally valid types and principles is innate since something being innate does not follow from the fact that it is present in all languages (also see Section 13.1.9).

In sum, one can say that theories that describe linguistic objects using a consistent descriptive inventory and then make use of inheritance hierarchies to capture generalizations, are the ones best suited to representing similarities between languages. Furthermore, this kind of theories are compatible with both a positive and negative answer to the question of innate linguistic knowledge.

24.2 How to arrive linguistic theories that capture cross-linguistic generalizations

In the previous subsection I argued for a uniform representation of linguistic knowledge on all descriptive levels and that type hierarchies are a good tool for representing generalizations. This subsection explores a way to develop grammars that are motivated by facts from several languages.

If one looks at the current practice in various linguistic schools one finds two extreme ways of approaching language. One the one hand we have the Mainstream Generative Grammar (MGG) camp and on the other hand we have the Construction Grammar/Cognitive Grammar camp. I hasten to say, that what I say here does not hold for all members of these groups, but for the extreme cases. The caricature of the MGG scientist is that he is looking for underlying structures. Since these have to be the same for all languages (poverty of the stimulus), it is sufficient to look at one language, say English. The result of this research strategy is that one ends up with models that were suggested by the most influential linguist for English and that one then tries to find ways to accommodate other languages. Since English has an NP VP structure, all languages to have it. Since English reorders constituent in passive sentences, passive is movement and all languages have to work this way. I discussed the respective analyses of German in more detail in the Section 3.4.2 and Chapter 22 and showed that the assumption that passive is movement makes unwanted predictions for German since the subject of passives stays in the object position. Furthermore, this analysis requires the assumption of invisible expletives, that is, entities that cannot be seen and do not have any meaning.

On the other extreme of the spectrum we find people working in Construction Grammar and who claim that all languages are so different that we not even can use the same vocabulary to analyze them and within languages we have so many different objects that it is impossible (or too early) to state any generalizations. Again, what I describe here is extreme positions and clichés.

In what follows, I sketch the procedure that we apply in the CoreGram project[1] (Müller 2013a; 2015a). In the CoreGram project we work on a set of typologically diverse languages in parallel:

- German (Müller 2007b; 2009c; 2012; Müller & Ørsnes 2011; 2013a; Müller 2014b; 2015b)

[1] http://hpsg.fu-berlin.de/Projects/CoreGram.html, May 20, 2015.

- Danish (Ørsnes 2009b; Müller 2009c; 2012; Müller & Ørsnes 2011; 2013a;b; 2015)
- Persian (Müller 2010b; Müller & Ghayoomi 2010; Müller, Samvelian & Bonami In Preparation)
- Maltese (Müller 2009b)
- Mandarin Chinese (Lipenkova 2009; Müller & Lipenkova 2009; 2013; 2015)
- Yiddish (Müller & Ørsnes 2011)
- English (Müller 2009c; 2012; Müller & Ørsnes 2013a)
- Hindi
- Spanish
- French

These languages belong to diverse language families (Indo-European, Afro-Asiatic, Sino-Tibetan) and among the Indo-European languages the languages belong to different groups (Germanic, Romance, Indo-Iranian). Figure 24.2 provides an overview. We work

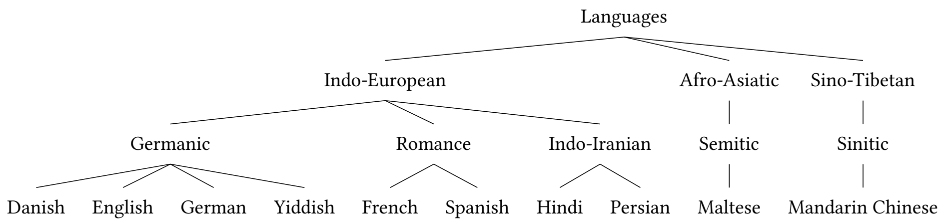


Figure 24.2: Language families and groups of the languages covered in the CoreGram project

out fully formalized, computer-processable grammar fragments in the framework of HPSG that all have a semantics component. The details will not be discussed here, but the interested reader is referred to (Müller 2015a).

As was argued in previous sections, the assumption of innate language-specific knowledge should be kept to a minimum. This is also what Chomsky suggested in his Minimalist Program. It is even possible that there is no language-specific innate knowledge at all, a view taken in Construction Grammar/Cognitive Grammar. So instead of imposing constraints from one language onto other languages, a bottom-up approach seems to be more appropriate: Grammars for individual languages should be motivated language-internally. Grammars that share certain properties can be grouped in classes. This makes it possible to capture generalizations about groups of languages and language as such. Let us consider some example languages: German, Dutch, Danish, English and French. If we start developing grammars for German and Dutch, we find that they share a lot of properties: both are SOV and V2 languages, both have a verbal complex. One main difference is the order of elements in the verbal complex. The situation can be depicted as in Figure 24.3. There are some properties that are shared between German and Dutch (Set

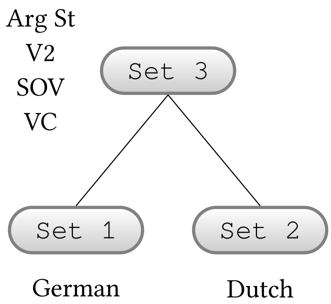


Figure 24.3: Shared properties of German and Dutch

3). For instance, the argument structure of lexical items, a list containing descriptions of syntactic and semantic properties of arguments, and the linking of these arguments to the meaning of the lexical items is contained in Set 3. In addition the constraints for SOV languages, the verb position in V2 clauses and the fronting of a constituent in V2 clauses are contained in Set 3. The respective constraints are shared between the two grammars. Although these sets are arranged in a hierarchy in the figures this has nothing to do with the type hierarchies that have been discussed in the previous subsection. These type hierarchies are part of our linguistic theories and parts of such hierarchies can be in different sets: Those parts of the type hierarchy that concern more general aspects can be in Set 3 in Figure 24.3 and those that are specific to Dutch or German are in the respective other sets. When we add another language, say Danish, we get further differences. While German and Dutch are SOV, Danish is an SVO language. Figure 24.4 on the following page shows the resulting situation: The top-most node represents constraints that hold for all the languages considered so far (for instance the argument structure constraints, linking and V2) and the node below it (Set 4) contains constraints that hold for German and Dutch only.[2] For instance Set 4 contains constraints regarding verbal complexes and SOV order. The union of Set 4 and Set 5 is Set 3 of Figure 24.3.

If we add further languages, further constraint sets will be distinguished. Figure 24.5 on the next page shows the situation that results when we add English and French. Again, the picture is not complete since there are constraints that are shared by Danish and English but not by French, but the general idea should be clear: By systematically working this way, we should arrive at constraint sets that directly correspond to those that have been established in the typological literature.

[2] In principle, there could be constraints that hold for Dutch and Danish but not for German and for German and Danish, but not for Dutch. These constraints would be removed from Set 1 and Set 2 respectively and put into another constraint set higher up in the hierarchy. These sets are not illustrated in the figure and I keep the names Set 1 and Set 2 from Figure 24.3 for the constraint sets for German and Dutch.

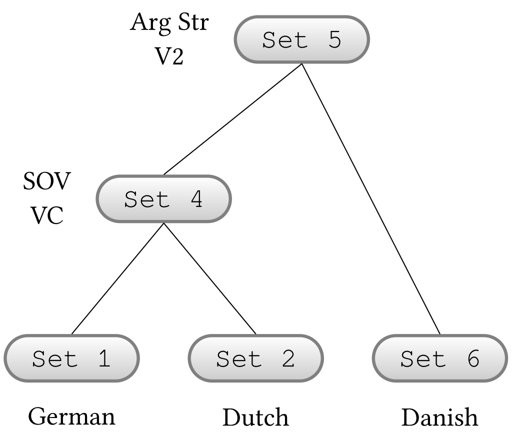


Figure 24.4: Shared Properties of German, Dutch, and Danish

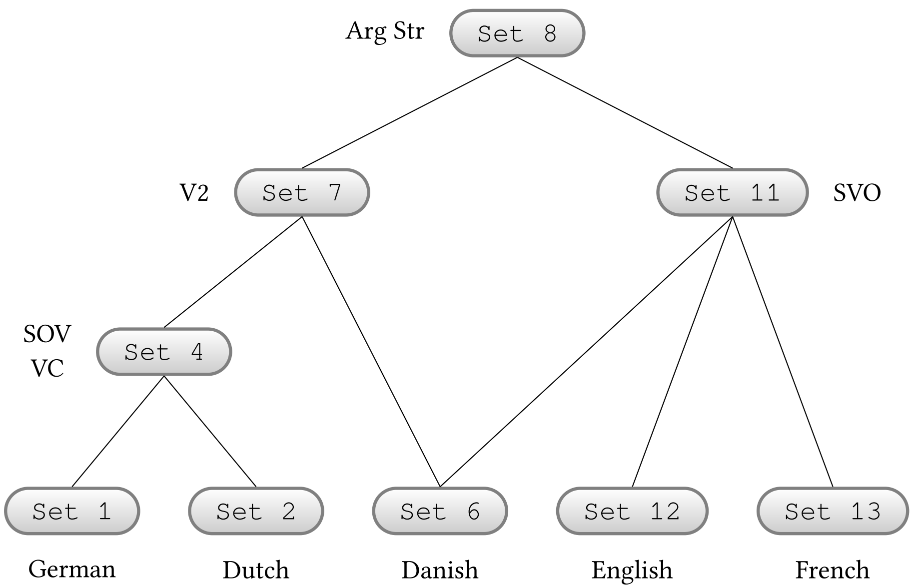


Figure 24.5: Languages and language classes

The interesting question is what will be the top-most set if we consider enough languages. On the first glance one would expect that all languages have valence representations and linkings between these and the semantics of lexical items (argument structure lists in the HPSG framework). However, Koenig & Michelson (2012) argue for an analysis of Oneida (a Northern Iroquoian language) that does not include a representation of syntactic valence. If this analysis is correct, syntactic argument structure would not be universal. It would of course be characteristic for a large number of languages, but it would not be part of the top-most set. So this leaves us with just one candidate for the top-most set from the area of syntax: the constraints that license the combination of two or more linguistic objects. This is basically Chomsky's External Merge without the restriction on binarity[3]. In addition the top-most set would of course contain the basic machinery for representing phonology and semantics.

It should be clear from what has been said so far that the goal of every scientist who works this way is to find generalizations and to describe a new language in a way that reuses theoretical constructs that have been found useful for a language that is already covered. However, as was explained above, the resulting grammars should be motivated by data of the respective languages and not by facts from other languages. In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X. I call this approach the *bottom-up approach with cheating*: Unless there is contradicting evidence we can reuse analyses that have been developed for other languages.

Note that this approach is compatible with the rather agnostic view advocated by Haspelmath (2010a), Dryer (1997), Croft (2001: Section 1.4.2–1.4.3), and others, who argue that descriptive categories should be language-specific, that is, the notion of *subject* for Tagalog is different from the one for English, the category *noun* in English is different from the category *noun* in Persian and so on. Even if one follows such extreme positions, one can still derive generalizations regarding constituent structure, head-argument relations and so on. However, I believe that some categories can fruitfully be used cross-linguistically, if not universally, then at least for language classes. As Newmeyer (2010: 692) notes with regard to the notion of *subject*: Calling two items *subject* in one language does not entail that they have identical properties. The same is true for two linguistic items from different languages: calling a Persian linguistic item *subject* does not entail that it has exactly the same properties as an English linguistic object that is called *subject*. The same is, of course, true for all other categories and relations, for instance parts of speech: Persian nouns do not share all properties with English nouns.[4] Haspelmath

[3] Note that binarity is more restrictive than flat structures: There is an additional constraint that there have to be exactly two daughters. As was argued in Section 23.10.4 one needs phrasal constructions with more than two constituents.

[4] Note that using labels like *Persian Noun* and *English Noun* (see for instance Haspelmath (2010a: Section 2) for such a suggestion regarding case, e. g. Russian Dative, Korean Dative, …) is somehow strange since it implies that both Persian nouns and English nouns are somehow nouns. Instead of using the category *Persian Noun* one could assign objects of the respective class to the class *noun* and add a feature LANGUAGE with the value *persian*. This simple trick allows one to assign both objects of the type *Persian Noun* and

(2010b: 697) writes: *Generative linguists try to use as many crosslinguistic categories in the description of individual languages as possible, and this often leads to insurmountable problems.* If the assumption of a category results in problems, they have to be solved. If this is not possible with the given set of categories/features, new ones have to be assumed. This is not a drawback of the methodology, quite the opposite is true: If we have found something that does not integrate nicely into what we already have, this is a sign that we have discovered something new and exciting. If we stick to language-particular categories and features, it is much harder to notice that a special phenomenon is involved since all categories and features are specific for one language anyway. Note also that not all speakers of a language community have exactly the same categories. If one were to take the idea of language-particular category symbols to an extreme, one would end up with person specific category symbols like *Klaus-English-noun.*

After my talk at the MIT in 2013 members of the linguistics department objected to the approach taken in the CoreGram project and claimed that it would not make any predictions as far as possible/impossible languages are concerned. Regarding predictions two things must be said: Firstly, predictions are being made on a language particular basis. As an example consider the following sentences from Netter (1991):

(3) a. [Versucht, zu lesen], hat er das Buch nicht.
tried to read has he the book not
'He did not try to read the book.'

b. [Versucht, einen Freund vorzustellen], hat er ihr noch nie.
tried a friend to.introduce has he her yet never
'He never tried to introduce her to a friend.'

When I first read them I had no idea about their structure. I switched on my computer and typed them in and within milliseconds I got an analysis of the sentences and by inspecting the result I realized that these sentences are combinations of partial verb phrase fronting and the so-called third construction (Müller 1999a: 439). I had previously implemented analyses of both phenomena but had never thought about the interaction of the two. The grammar predicted that examples like (3) are grammatical. Similarly the constraints of the grammar can interact to rule out certain structures. So predictions about ungrammaticality/impossible structures are in fact made as well.

Secondly, the top-most constraint set holds for all languages seen so far. It can be regarded as a hypothesis about properties that are shared by all languages. This constraint set contains constraints about the connection between syntax and information structure and such constraints allow for V2 languages but rule out languages with the verb in penultimate position (see Kayne (1994: 50) for the claim that such languages do

objects of the type *English Noun* to the class *noun* and still maintain the fact that there are differences. Of course, no theoretical linguist would introduce the LANGUAGE feature to differentiate between Persian and English nouns, but nouns in the respective languages have other features that make them differ. So the part of speech classification as noun is a generalization over nouns in various languages and the categories *Persian Noun* and *English Noun* are feature bundles that contain further, language-specific information.

not exist. Kayne develops a complicated syntactic system that predicts this). Of course if a language is found that puts the verb in penultimate position for the encoding of sentence types or some other communicative effect, a more general top-most set has to be defined, but this is parallel for Minimalist theories: If languages are found that are incompatible with basic assumptions, the basic assumptions have to be revised. As with the language particular constraints the constraints from the top-most set make certain predictions about what can be and what cannot be found in languages.

Frequently discussed examples such as those languages that form questions by reversing the order of the words in a string (Musso et al. 2003) need not be ruled out in the grammar since they are ruled out by language external constraints: We simply do not have enough working memory to do such complex computations.

Cinque (1999: 106) suggested a cascade of functional projections to account for reoccurring orderings in the languages of the world. He assumes elaborate tree structures to play a role in the analysis of all sentences in all languages even if there is no evidence for respective morphosyntactic distinctions in a particular language (see also Cinque & Rizzi (2010: 55)). In the latter case Cinque assumes that the respective tree nodes are empty. Cinque's results could be incorporated in the model advocated here. We would define part of speech categories and morpho-syntactic features in the top-most set and state linearization constraints that enforce the order that Cinque encoded directly in his tree structure. In languages in which such categories are not manifested by lexical material the constraints would never apply. Neither empty elements nor elaborate tree structures would be needed. So, we have shown that Cinque's data could be covered in a better way in an HPSG with a rich UG but we, nevertheless, refrain from introducing 400 categories (or features) into the theories of all languages and, again, we point out that this is implausible from a genetic point of view and wait for other, probably functional, explanations of the Cinque data.

Note also that implicational universals can be derived from such constraint sets. For instance, we can derive from Figure 24.5 the implicational statement that all SVO languages are V2 languages since there is no language that has constraints from Set 4 that does not also have the constraints of Set 7. Of course this implicational statement is wrong, since there are lots and lots of SOV languages and just exceptionally few V2 languages. So as soon as we add other languages as for instance Persian or Japanese, the picture will change.

The methodology suggested here differs from what is done in MGG in that the general constraints that are supposed to hold for all languages are just stipulated on the basis of general speculations about language. In the best case theses general assumptions are fed by a lot of experience with different languages and grammars, in the worst case they are derived from insights gathered from one or more Indo-European languages. Quite often impressionistic data is used to motivate rather far-reaching fundamental design decisions (Fanselow 2009; Sternefeld & Richter 2012; Haider 2014). It is interesting to note that this is exactly what members of the MGG camp accuse typologists of. Evans & Levinson (2009a) pointed out that counter examples can be found for many alleged universals. A frequently uttered reply is that unanalyzed data cannot refute grammatical hypotheses (see for instance Freidin (2009: 454)). In the very same way it has to be said that unanalyzed data should not be used to build theories on (Fanselow 2009). In the CoreGram project, we aim to develop broad-coverage grammars of several languages, so those constraints that make it to the top node are motivated and not stipulated on the basis of intuitive implicit knowledge about language.

This research strategy is compatible with work done in Construction Grammar (see Goldberg (2013b: 481) for an explicit statement to this end), since it is data-oriented and does not presuppose innate language-specific knowledge and should also be compatible with the Minimalist world.

25 Conclusion

The analyses discussed in this book show a number of similarities. All frameworks use complex categories to describe linguistic objects. This is most obvious for GPSG, LFG, CxG and FTAG, however, GB/Minimalism and Categorial Grammar also talk about NPs in third person singular and the relevant features for part of speech, person and number form part of a complex category. In GB, there is the feature N and V with binary values (Chomsky 1970: 199), Stabler (1992: 119) formalizes *Barriers* with feature-value pairs and Sauerland & Elbourne (2002: 290–291) argue for the use of feature-value pairs in a Minimalist theory. Also, see Veenstra (1998) for a constraint-based formalization of a Minimalist analysis using typed feature descriptions. Dependency Grammar dialects like Hellwig's Dependency Unification Grammar also use feature value pairs (Hellwig 2003: 612).

Furthermore, there is a consensus in all current frameworks (with the exception of Construction Grammar and Dependency Grammar) about how the sentence structure of German should be analyzed: German is an SOV and V2 language. Clauses with verb-initial order resemble verb-final ones in terms of their structure. The finite verb is either moved (GB) or stands in a relation to an element in verb-final position (HPSG). Verb-second clauses correspond to verb-initial clauses in which a constituent has been fronted. It is also possible to see some convergence with regard to the analysis of the passive: Some ideas originally formulated by Haider (1984; 1985a; 1986a) in the framework of GB have been adopted by HPSG. Some variants of Construction Grammar also make use of a specially marked 'designated argument' (Michaelis & Ruppenhofer 2001: 55–57).

If we consider new developments in the individual frameworks, it becomes clear that the nature of the proposed analysis can sometimes differ drastically. Whereas CG, LFG, HPSG and CxG are surface-oriented, sometimes very abstract structures are assumed in Minimalism and in some cases, one tries to trace all languages back to a common base structure (Universal Base Hypothesis).[1] This kind of approach only makes sense if one assumes that there is innate linguistic knowledge about all base structure common to all these language as well as the operations necessary to derive the surface structures. As was shown in Chapter 13, all arguments for the assumption of innate linguistic knowledge are either not tenable or controversial at the very least. The acquisition of linguistic abilities can to a large extent receive an input-based explanation (Section 13.8.3, Section 16.3 and Section 16.4). Not all questions about acquisition have been settled once and for all, however, input-based approaches are at least plausible enough for one to be very cautious about any assumption of innate linguistic knowledge.

[1] It should be noted that there are currently many subvariants and individual opinions in the Minimalist community so that it is only possible – as with CxG – to talk about tendencies.

Models such as LFG, CG, HPSG, CxG and TAG are compatible with performance data, something that is not true of certain transformation-based approaches. These approaches are viewed as theories of competence that do not make any claims about performance. It is assumed that there are other mechanisms for working with linguistic knowledge, for example, those that combine 'chunks' (fragments of linguistic material). If one wishes to make these assumptions, then it is necessary to explain how chunks and the processing of chunks are acquired and not how a complex system of transformations and transformation-comparing constraints is acquired. This means that the problem of language acquisition would be a very different one. If one assumes a chunk-based approach, then the innate knowledge about a universal transformational base would only be used to derive a surface-oriented grammar. This then poses the question as to what exactly the evidence for transformations in a competence grammar is and if it would not be preferable to simply assume that the competence grammar is of the kind assumed by LFG, CG, HPSG, CxG or TAG. One should therefore bear in mind that constraint-based analyses and the kind of transformational approaches that allow reformulation as constraint-based ones are the only approaches that are compatible with the current facts, whereas all other analyses require additional assumptions.

A number of works in Minimalism differ from those in other frameworks in that they assume structures (sometimes also invisible structures) that can only be motivated by evidence from other languages. This approach can streamline the entire apparatus for deriving different structures. However, the overall costs of the approach are not reduced: Some amount of the cost is just transferred to the UG component. The abstract grammars that result cannot be learned from the input.

One can take from this discussion that only constraint-based, surface-oriented models are adequate and explanatory: They are also compatible with psycholinguistic facts and plausible from the point of view of acquisition.

If we now compare these approaches, we see that a number of analyses can be translated into one another. LFG (and some variants of CxG) differ from all other theories in that grammatical functions such as subject and object are primitives of the theory. If one does not want this, then it is possible to replace these labels with Argument1, Argument2, etc. The numbering of arguments would correspond to their relative obliqueness. LFG would then move closer to HPSG. Alternatively, one could mark arguments in HPSG and CxG with regard to their grammatical function additionally. This is what is done for the analysis of the passive (DESIGNATED ARGUMENT).

LFG, HPSG, CxG and variants of Categorial Grammar (Moens et al. 1989; Briscoe 2000; Villavicencio 2002) possess means for the hierarchical organization of knowledge that is important for capturing generalizations. It is, of course, possible to expand any other framework in this way, but this has never been done explicitly apart from in computer implementations and does not play an active role in theorizing in the other frameworks.

In HPSG and CxG, roots, stems, words, morphological and syntactic rules are all objects that can be described with the same descriptive means. This then allows one to make generalizations that affect very different objects (see Chapter 24). In LFG, c-structures are viewed as something fundamentally different, which is why this kind of generalization is not possible. In cross-linguistic work, there is an attempt to capture similarities in the f-structure, the c-structure is less important and is not even discussed in a number of works. Furthermore, its implementation from language to language can differ enormously. For this reason, my personal preference is for frameworks that describe all linguistic objects using the same means, that is, HPSG and CxG. Formally, nothing stands in the way of a description of the c-structure of an LFG grammar using feature-value pairs so that in years to come there could be even more convergence between the theories. For hybrid forms of HPSG and LFG, see Ackerman & Webelhuth (1998) and Hellan & Haugereid (2003), for example.

If one compares CxG and HPSG, it becomes apparent that the degree of formalization in CxG works is relatively low and a number of questions remain unanswered. The more formal approaches in CxG are variants of HPSG. There are relatively few precisely worked-out analyses in Construction Grammar and no description of German that would be comparable to the other approaches presented in this book. In fairness, it must be said that Construction Grammar is the youngest of the theories discussed here. Its most important contributions to linguistic theory have been integrated into frameworks such as HPSG and LFG.

The theories of the future will be a fusion of surface-oriented, constraint-based and model-theoretic approaches CG, LFG, HPSG, Construction Grammar, equivalent variants of TAG and GB/Minimalist approaches that will be reformulated as constraint-based. (Variants of) Minimalism and (variants of) Construction Grammar are the most widely adopted approaches at present. I actually suspect the truth to lie somewhere in the middle. The linguistics of the future will be data-oriented. Introspection as the sole method of data collection has proven unreliable (Müller 2007d; Meurers & Müller 2009) and is being increasingly complemented by experimental and corpus-based analyses.

Statistical information and statistical processes play a very important role in machine translation and are becoming more important for linguistics in the narrow sense (Abney 1996). We have seen that statistical information is important in the acquisition process and Abney discusses cases of other areas of language such as language change, parsing preferences and gradability with grammaticality judgments. Following a heavy focus on statistical procedures, there is now a transition to hybrid forms in computational linguistics,[2] since it has been noticed that it is not possible to exceed certain levels of quality with statistical methods alone (Steedman 2011; Church 2011; Kay 2011). The same holds here as above: The reality is somewhere in the middle, that is, in combined systems. In order to have something to combine, the relevant linguistic theories first need to be developed. As Manfred Pinkal said: „It is not possible to build systems that understand language without understanding language.“

[2] See Kaufmann & Pfister (2007) and Kaufmann (2009) for the combination of a speech recognizer with a HPSG grammar.

A Solutions to the exercises

A.1 Introduction and basic terms

(1) a. $\underbrace{\text{Karl}}_{\text{VF}}\ \underbrace{\text{isst}}_{\text{LS}}$.

b. $\underbrace{\text{Der Mann}}_{\text{VF}}\ \underbrace{\text{liebt}}_{\text{LS}}\ \underbrace{\text{eine Frau,}}_{\text{MF}}\ \underbrace{\underbrace{\text{den}}_{\text{VF}}\ \underbrace{\text{Peter}}_{\text{MF}}\ \underbrace{\text{kennt}}_{\text{RS}}}_{\text{NF}}.$

c. $\underbrace{\text{Der Mann}}_{\text{VF}}\ \underbrace{\text{liebt}}_{\text{LS}}\ \underbrace{\text{eine Frau,}\ \underbrace{\text{die}}_{\text{VF}}\ \underbrace{\text{Peter}}_{\text{MF}}\ \underbrace{\text{kennt}}_{\text{RS}}}_{\text{MF}}.$

d. $\underbrace{\text{Die Studenten}}_{\text{VF}}\ \underbrace{\text{haben}}_{\text{LS}}\ \underbrace{\text{behauptet,}}_{\text{RS}}\ \underbrace{\underbrace{\text{nur wegen der Hitze}}_{\text{MF}}\ \underbrace{\text{einzuschlafen}}_{\text{RS}}}_{\text{NF}}.$

e. $\underbrace{\underbrace{\text{Dass}}_{\text{LS}}\ \underbrace{\text{Peter nicht}}_{\text{MF}}\ \underbrace{\text{kommt,}}_{\text{RS}}}_{\text{VF}}\ \underbrace{\text{ärgert}}_{\text{LS}}\ \underbrace{\text{Klaus}}_{\text{MF}}.$

f. $\underbrace{\underbrace{\text{Einen Mann}}_{\text{MF}}\ \underbrace{\text{küssen,}}_{\text{RS}}\ \underbrace{\underbrace{\text{der}}_{\text{VF}}\ \underbrace{\text{ihr nicht}}_{\text{MF}}\ \underbrace{\text{gefällt,}}_{\text{RS}}}_{\text{NF}}}_{\text{VF}}\ \underbrace{\text{würde}}_{\text{LS}}\ \underbrace{\text{sie nie}}_{\text{MF}}.$

On (1c): Theoretically, this could also be an extraposition of the relative clause to the postfield. Since *eine Frau, die Peter kennt* is a constituent, however, it is assumed that no reordering of the relative clause has taken place. Instead, we have a simpler structure with *eine Frau, die Peter kennt* as a complete NP in the middle field.

A.2 Phrase structure grammars

1. For any grammar, it is possible to assume additional symbols and rules that create unnecessary structure or are simply never used because there are no words or

phrases that could be used on the right-hand side of the rule. If we were to add the following rule to our grammar, for example, we would have a more complex grammar that can still analyze the same fragment of the language.

(2) Trallala → Trulla Trollolo

2. In general, it is assumed that the grammar with the fewest rules is the best one. Therefore, we can reject grammars that contain unnecessary rules such as (2).

   One should bear in mind what the aim of a theory of grammar is. If our goal is to describe the human capacity for language, then a grammar with more rules could be better than other grammars with less rules. This is because psycholinguistic research has shown that highly-frequent units are simply stored in our brains and not built up from their individual parts each time, although we would of course be able to do this.

3. The problems here is the fact that it is possible to derive a a completely empty noun phrase (see Figure A.1). This noun phrases could be inserted in all positions where

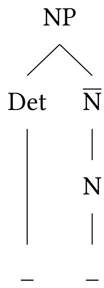


Figure A.1: Noun phrases without a visible determiner and noun

   an otherwise filled NP would have to stay. Then, we would be able to analyze sequences of words such as (3), where the subject of *schläft* 'sleeps' is realized by an empty NP:

   (3) * Ich glaube, dass schläft.
   I believe that sleeps

   This problem can be solved using a feature that ensures that determines whether the left periphery of the $\overline{\text{N}}$ is empty. Visible Ns and $\overline{\text{N}}$ with at least an adjective would have the value '−' and otherwise '+'. Empty determiners could then only be combined with $\overline{\text{N}}$s that have the value '−'. See Netter (1994).

4. If *Bücher* 'books' were an NP in the lexicon, then adjectives such as *interessant* 'interesting' would have to modify NPs in order for phrases such as (4) to be analyzed.

(4) interessante Bücher
interesting books

If adjectives are combined with NPs, however, it still has to be explained why (5) is ungrammatical.

(5) * interessante die Bücher
interesting the books

For a detailed discussions of this topic, see Müller (2007b: Section 6.6.2).

5. This kind of rule cannot analyze noun phrases such as those in (6):

   (6) a. interessante [Aufsätze und Bücher]
   interesting essays and books

   b. interessante [Aufsätze und Bücher aus Stuttgart]
   interesting essays and books from Stuttgart

   Since adjectives can only be combined directly with nouns, these phrases cannot be analyzed. *Bücher* 'books' or *Bücher aus Stuttgart* 'books from Stuttgart' would be complete NPs. Since it is assumed that coordinated elements always have the same syntactic category, then *Aufsätze* 'essays' would have to be an NP. *Aufsätze und Bücher* and *Aufsätze und Bücher aus Stuttgart* would then also be NPs and it remains unexplained how an adjective can be combined with this NP. Because of (5), we can rule out analyses that assume that full NPs combine with adjectives.

   See Chapter 21 for a general discussion of empty elements.

6. If a specific determiner or just any determiner would be combined with an adjective to form a complete NP, there would be no room for the integration of post-nominal modifiers like modifying genetives, PPs and relative clauses. For PPs and relative clauses, analyses have been suggested in which these postnominal modifiers attach to complete NPs, but modifying genetives usually attach to smaller units. But even if one admits postnominal modifiers to attach to complete NPs, one cannot account for the iteration of adjectives and for arguments that depend on the elided noun.

   So, the simplest way to cope with this data is the assumtion of an empty noun. Alternatively one could assume that an adjective is directly projected to an $\overline{\mathrm{N}}$. This $\overline{\mathrm{N}}$ then can be modified by further adjectives or postnominal modifiers. The $\overline{\mathrm{N}}$ is combined with a determiner to form a full NP. For phrases that involve elided relational nouns, one would have to assume the projection of an argument like *vom Gleimtunnel* 'of the Gleimtunnel' to $\overline{\mathrm{N}}$. The $\overline{\mathrm{N}}$ could be further modified or combined with a determiner directly.

7. Adjective phrases such as those in (7) cannot be analyzed since the degree modifier occurs between the complement and the adjective:

(7) der auf seinen Sohn sehr stolze Vater
the on his son very proud father
'the father very proud of his son'

One would either have to allow for specifiers to be combined with their heads before complements or allow crossing lines in trees. Another assumption could be that German is like English, however then adjectival complements would have to be obligatorily reordered before their specifier. For a description of this kind of reordering, see Chapter 3 as well as Section 13.1.2 for a discussion of $\overline{\text{X}}$-Theory.

8. Write a phrase structure grammar that can analyze the sentences in (8), but does not allow the strings of words in (9).

(8) a. Der Mann hilft der Frau.
the.NOM man helps the.DAT woman

b. Er gibt ihr das Buch.
he.NOM gives her.DAT the.ACC book

c. Er wartet auf ein Wunder.
he.NOM waits on a miracle.ACC
'He is waiting for a miracle.'

(9) a. * Der Mann hilft er.
the.NOM man helps he.NOM

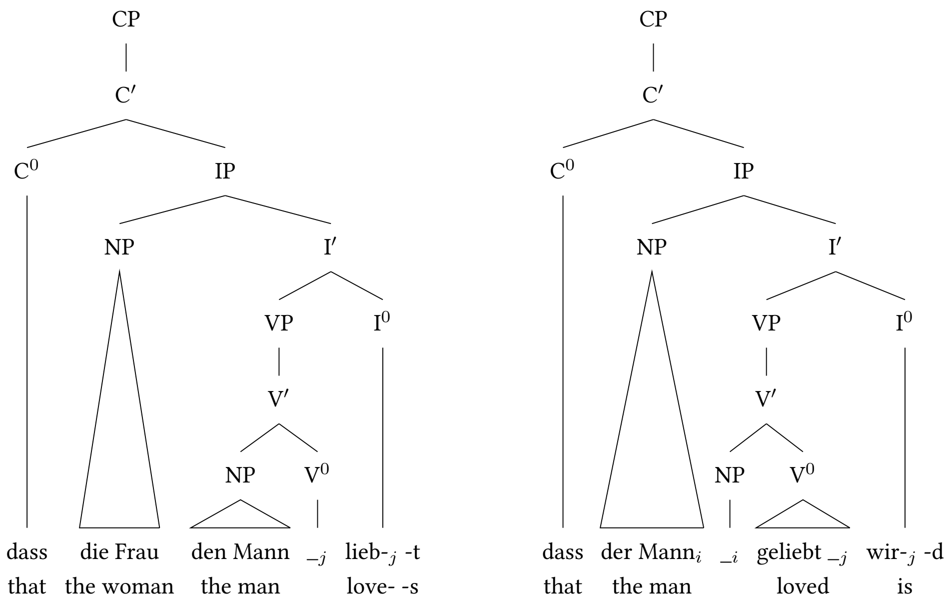
b. * Er gibt ihr den Buch.
he.NOM gives her.DAT the.ACC book

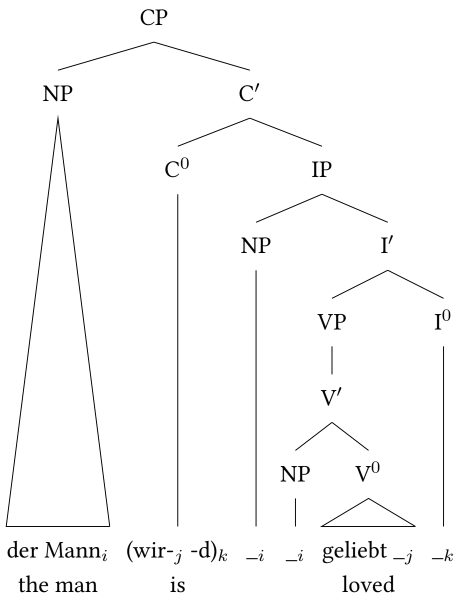
In order to rule out the last two sentences, the grammar has to contain information about case. The folloinwg grammar will do the job:

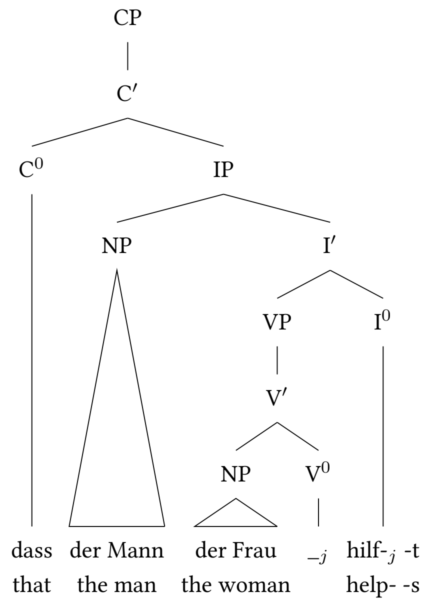
(10) a. s → np(nom) v(nom_dat), np(dat)
b. s → np(nom), v(nom_dat_akk), np(dat), np(akk)
c. s → np(nom), v(nom_pp_auf), pp(auf,akk)
d. pp(Pform,Kas) → p(Pform,Kas), np(Kas)
e. np(Kas) → d(Kas), n(Kas)
f. v(nom_dat) → hilft
g. v(nom_dat_akk) → gibt
h. v(nom_pp_auf) → wartet
i. np(nom) → er
j. np(dat) → ihr
k. d(nom) → der
l. d(dat) → der

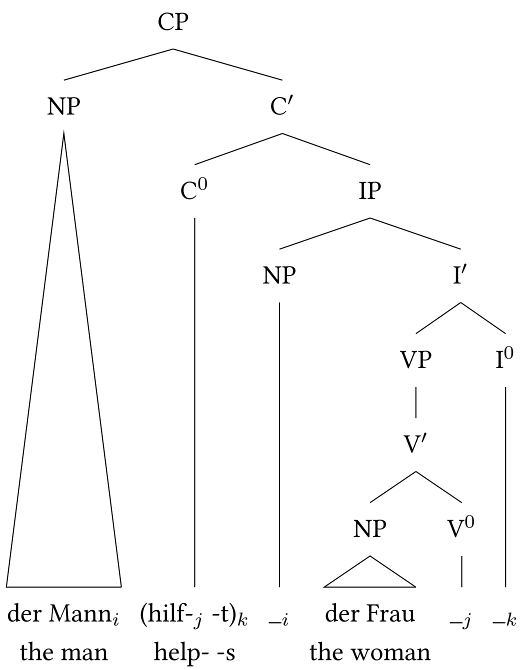
m. d(akk) → das
n. d(akk) → ein
o. n(nom) → Mann
p. n(dat) → Frau
q. n(akk) → Buch
r. n(akk) → Wunder
s. p(auf,akk) → auf

A.3 Transformational Grammar – Government & Binding




CP
NP
C′
C0
IP
NP
I′
VP
I0
V′
NP
V0
der Manni (wir-j -d)k _i _i geliebt _j _k
the man is loved


CP
C′
C0
IP
NP
I′
VP
I0
V′
NP
V0
dass der Mann der Frau _j hilf-j -t
that the man the woman help- -s


CP
NP
C′
C0
IP
NP
I′
VP
I0
V′
NP
V0
der Manni (hilf-j -t)k _i der Frau _j _k
the man help- -s the woman

A.4 Generalized Phrase Structure Grammar

In order to analyze the sentences in (11), one requires a rule for transitive verbs and a metarule for the extraction of an element. Furthermore, rules for the combination of elements in the noun phrase are required.

(11) a. [dass] der Mann ihn liest
   that the man it reads
   'that the man reads it'
  b. [dass] ihn der Mann liest
   that it the man reads
   'that the man reads it'
  c. Der Mann liest ihn.
   the man reads it
   'The man reads it.'

It is possible to analyze the sentences in (11a,b) using the rules in (12) and the lexical entries in (13).

(12) a. V3 → H[6], N2[CASE nom], N2[CASE acc]
  b. N2 → Det[CASE CAS], H1[CASE CAS]
  c. N1 → H[27]

(13) a. Det[CASE nom] → der
  b. N[27] → Mann
  c. V[6, +FIN] → liest
  d. N2[CASE acc] → ihn

The rules (12b,c) correspond to $\overline{\text{X}}$-rules that we encountered in Section 2.4.1. They only differ from these rules in that the part of speech of the head is not given on the right-hand side of the rule. The part of speech is determined by the Head Feature Convention. The part of speech of the head is identical to that on the left-hand side of the rule, that is, it must be N in (12b,c). It also follows from the Head Feature Convention that the whole NP has the same case as the head and therefore does not have to be mentioned additionally in the rule above. 27 is the SUBCAT value. This is number is arbitrary.

In order for the verb to appear in the correction position, we need linearization rules:

(14) V[+MC] < X
  X < V[−MC]

The fact that the determiner precedes the noun is ensured by the following LP-rule:

(15) Det < X

The Extraction Meta Rule in (16) is required in order to analyze (11c):

(16) V3 → W, X ↦
V3/X → W

Among other, this metarule licenses the rule in (17) for (12a):

(17) V3/N2[CASE nom] → H[6], N2[CASE acc]

The rule in (18) is used to bind off long-distance dependencies.

(18) V3[+FIN] → X[+TOP], V3[+MC]/X

The following linearization rule ensures that the +TOP-constituent precedes the sentence in which it is missing:

(19) [+TOP] < X

Figure A.2 shows the structure licensed by the grammar. In sum, one can say that the

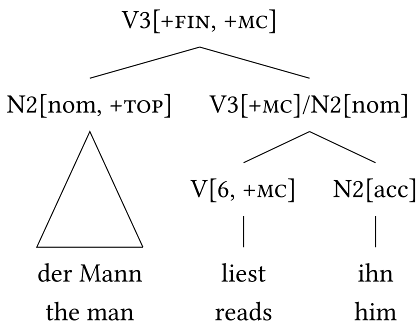


Figure A.2: Analysis of *Der Mann liest ihn.* 'The man reads it.'

grammar that licenses the sentences in (11) should have (at least) the following parts:

1. ID rules:

   (20) a. V3 → H[6], N2[CASE nom], N2[CASE acc]
   b. N2 → Det[CASE CAS], H1[CASE CAS]
   c. N1 → H[27]

2. LP rules:

   (21) V[+MC] < X
   X < V[−MC]
   Det < X
   [+TOP] < X

3. metarules:

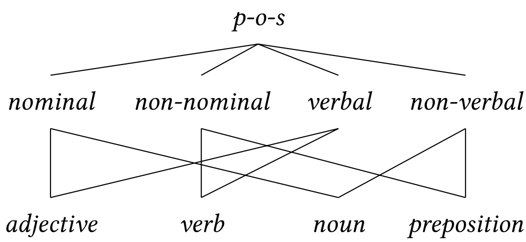
(22) V3 → W, X ↦
V3/X → W

4. Lexical entries

(23) a. Det[CASE nom] → der
b. N[27] → Mann
c. V[6, +FIN] → liest
d. N2[CASE acc] → ihn

A.5 Feature descriptions

1. For the class [+V], the type *verbal* is assumed with the subtypes *adj* and *verb*. For the class [−V] there is the type *non-verbal* and its subtypes *noun* and *preposition*. This is analogous for the N values. The corresponding hierarchy is given in the following figure:



2. Lists can be described using recursive structures that consist of both a list beginning and the rest. The rest can either be a non-empty list (*ne_list*) or an empty list (*e_list*). The lists ⟨ *a*, *b*, *c* ⟩ can be represented as follows:

(24) $$\begin{bmatrix} \textsc{first} & a \\ \textsc{rest} & \begin{bmatrix} \textsc{first} & b \\ \textsc{rest} & \begin{bmatrix} \textsc{first} & c \\ \textsc{rest} & e\_list \\ \multicolumn{2}{l}{ne\_list} \end{bmatrix} \\ \multicolumn{2}{l}{ne\_list} \end{bmatrix} \\ \multicolumn{2}{l}{ne\_list} \end{bmatrix}$$

3. If we extend the data structure in (24) by two additional features, it is possible to do without *append*. The keyword is *difference list*. A difference list consists of a list and a pointer to the last element of the list.

(25)
$$\begin{bmatrix} \textsc{list} & \begin{bmatrix} \textsc{first} & a \\ \textsc{rest} & \begin{bmatrix} \textsc{first} & b \\ \textsc{rest} & \boxed{1}\ list \\ \textit{ne\_list} & \end{bmatrix} \\ \textit{ne\_list} & \end{bmatrix} \\ \textsc{last} & \boxed{1} \\ \textit{diff-list} & \end{bmatrix}$$

Unlike the list representation in (24), the REST value of the end of the list is not *e_list*, but rather simply *list*. It is then possible to extend a list by adding another list to the point where it ends. The concatenation of (25) and (26a) is (26b).

(26) a.
$$\begin{bmatrix} \textsc{list} & \begin{bmatrix} \textsc{first} & c \\ \textsc{rest} & \boxed{2}\ list \\ \textit{ne\_list} & \end{bmatrix} \\ \textsc{last} & \boxed{2} \\ \textit{diff-list} & \end{bmatrix}$$

b.
$$\begin{bmatrix} \textsc{list} & \begin{bmatrix} \textsc{first} & a \\ \textsc{rest} & \begin{bmatrix} \textsc{first} & b \\ \textsc{rest} & \begin{bmatrix} \textsc{first} & c \\ \textsc{rest} & \boxed{2}\ list \\ \textit{ne\_list} & \end{bmatrix} \\ \textit{ne\_list} & \end{bmatrix} \\ \textit{ne\_list} & \end{bmatrix} \\ \textsc{last} & \boxed{2} \\ \textit{diff-list} & \end{bmatrix}$$

In order to combine the lists, the LIST value of the second list has to be identified with the LAST value of the first list. The LAST value of the resulting list then corresponds to the LAST value of the second list ($\boxed{2}$ in the example.)

Information about the encoding of difference lists can be found by searching for the keywords *list*, *append*, and *feature structure*. In the search results, one can find pages on developing grammars that explain difference lists.

A.6 Lexical Functional Grammar

1. *kannte* 'knew' is a transitive verb:

(27) *kannte* V (↑ PRED) = 'KENNEN⟨SUBJ, OBJ ⟩'
(↑ SUBJ AGR CAS = NOM)
(↑ OBJ AGR CAS = ACC)
(↑ TENSE) = PAST

2. In the sentence (28), the object of *verschlingen* is in the prefield.

(28) Den Apfel verschlingt David.
the.ACC apple devours David.NOM
'David is devouring the apple.'

The analysis is a combination of the analysis in Figure 7.2 on page 231 and the analysis of long-distance dependencies that was presented in Section 7.5. The object is not realized inside the VP, but rather in the prefield.

The necessary c-structure rules are given in (29):

(29) a. VP → NP (↑SUBJ |OBJ) = ↓ VP ↑= ↓
b. VP → (V) ↑= ↓
c. C′ → C ↑= ↓ VP ↑= ↓
d. CP → XP (↑DF) = ↓ (↑DF) = (↑COMP* GF) C′ ↑=↓

These rules allow two f-structures for the example in question: one in which the NP *den Apfel* 'the apple' is the topic and another in which this NP is the focus. Figure A.3 shows the analysis with a topicalized constituent in the prefield.

A.7 Categorial Grammar

1. The analysis of *The children in the room laugh loudly.* is given in Figure A.4.

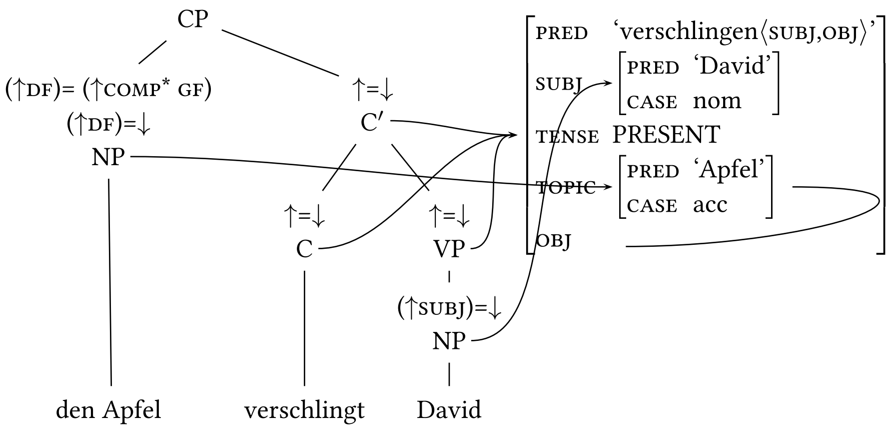


Figure A.3: Analysis of verb second

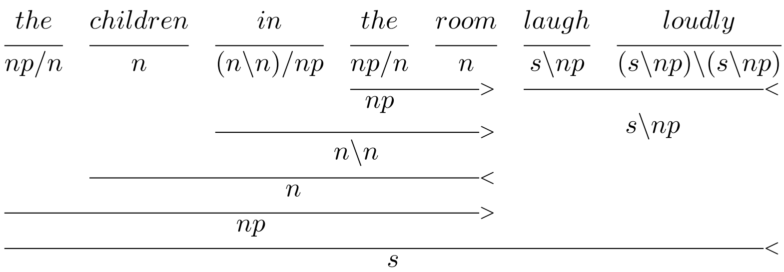


Figure A.4: Categorial Grammar analysis of *The children in the room laugh loudly.*

2. The analysis of *the picture of Mary* is given in Figure A.5. n/pp corresponds to N^0, n corresponds to $\overline{\text{N}}$ and np corresponds to NP.

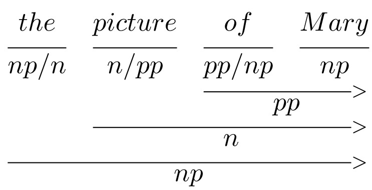


Figure A.5: Categorial Grammar analysis of *the picture of Mary*

A.8 Head-Driven Phrase Structure Grammar

1. The solution is on the facing page.

$$
\left[\begin{array}{ll}
\text{PHON} & \langle \textit{Max lacht} \rangle \\
\text{SYNSEM|LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \boxed{2}\,\langle\rangle \end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \text{IND} & \boxed{3} \\ \text{RELS} & \langle \boxed{4}, \boxed{5} \rangle \end{array}\right]
\end{array}\right] \\
\text{HEAD-DTR} & \left[\begin{array}{ll}
\text{PHON} & \langle \textit{lacht} \rangle \\
\text{SYNSEM|LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \boxed{1} \left[\begin{array}{ll} \text{INITIAL} & - \\ \text{VFORM} & \textit{fin} \\ \textit{verb} & \end{array}\right] \\
\text{SUBCAT} & \boxed{2} \oplus \langle \boxed{6} \rangle
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll}
\text{IND} & \boxed{3}\ \textit{event} \\
\text{RELS} & \left\langle \boxed{4} \left[\begin{array}{ll} \text{EVENT} & \boxed{3} \\ \text{AGENS} & \boxed{7} \\ \textit{lachen} & \end{array}\right] \right\rangle
\end{array}\right]
\end{array}\right] \\
\textit{word} &
\end{array}\right] \\
\text{NON-HEAD-DTRS} & \left\langle \left[\begin{array}{ll}
\text{PHON} & \langle \textit{Max} \rangle \\
\text{SYNSEM}\ \boxed{6} & \left[\begin{array}{ll}
\text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \text{CAS} & \textit{nom} \\ \textit{noun} & \end{array}\right] \\
\text{SUBCAT} & \langle\rangle
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll}
\text{IND} & \boxed{7} \left[\begin{array}{ll} \text{PER} & \mathit{3} \\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{mas} \end{array}\right] \\
\text{RELS} & \left\langle \boxed{5} \left[\begin{array}{ll} \text{NAME} & \textit{max} \\ \text{INST} & \boxed{7} \\ \textit{named} & \end{array}\right] \right\rangle
\end{array}\right]
\end{array}\right]
\end{array}\right] \\
\textit{word} &
\end{array}\right] \right\rangle \\
\textit{head-argument-phrase} &
\end{array}\right]
$$

2. An analysis of the difference in (30) has to capture the fact that the case of the adjective has to agree with that of the noun. In (30a), the genitive form of *interessant* 'interesting' is used, whereas (30b) contains a form that is incompatible with the genitive singular.

   (30) a. eines interessanten Mannes
   one.GEN interesting.GEN man.GEN

   b. * eines interessanter Mannes
   one.GEN interesting.NOM man.GEN

   (31) shows the CAT value of *interessanten*.

   (31) CAT value of *interessanten* 'interesting' with case information:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}}\ [\text{CASE}\ \boxed{1}] \\ \text{CASE} & \boxed{1}\ \textit{gen} \\ \textit{adj} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix}$$

   The structure sharing of the case value of the adjective with the $\overline{\text{N}}$ under MOD identifies the case values of the noun and the adjective. *interessanten* can therefore be combined with *Mannes*, but not with*Mann*. Similarly, *interessanter* can only be combined with the nominative *Mann*, but not with the genitive *Mannes*.

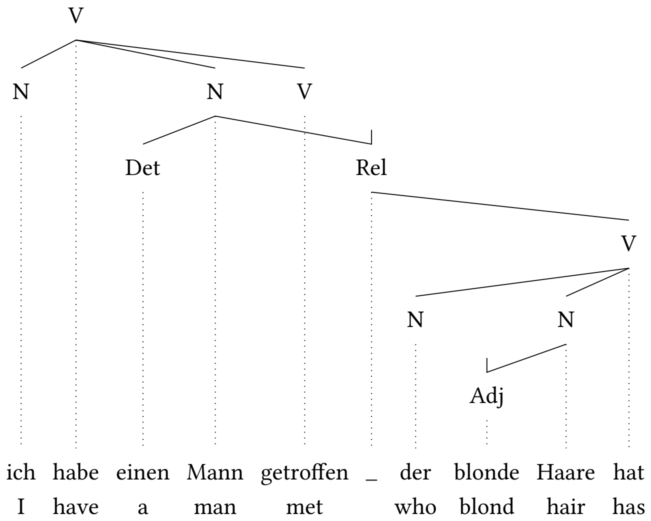
   For a refinement of the analysis of agreement inside the noun phrase, see Müller (2007b: Abschnitt 13.2).

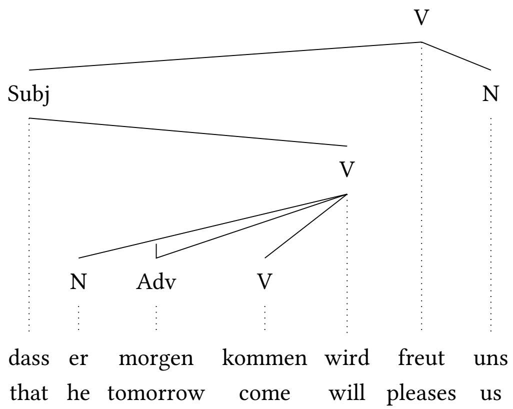
A.9 Construction Grammar

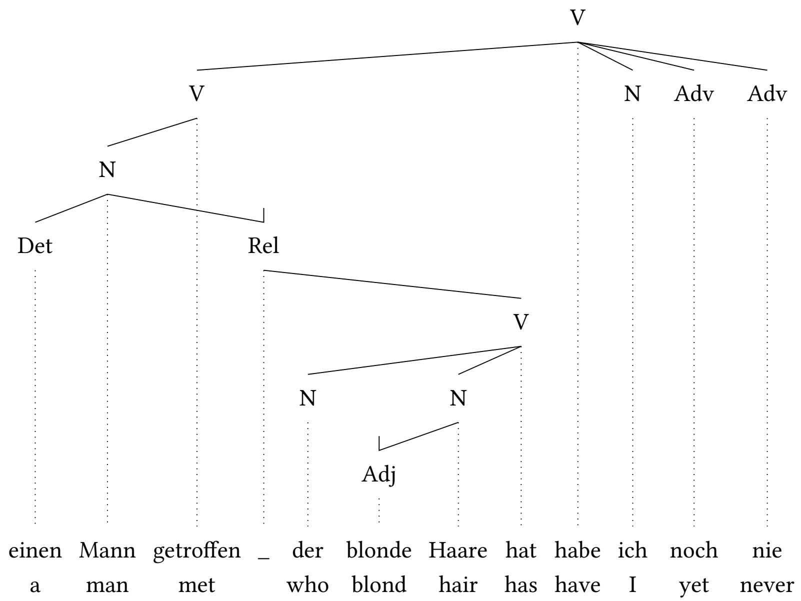
Idioms can be found by reading the newspaper carefully. The less exciting method is to look them up a dictionary of idioms such as the Free Dictionary of Idioms and Phrases[1].

[1] http://idioms.thefreedictionary.com/, 04.03.2015.

A.10 Dependency Grammar







A.11 Tree Adjoining Grammar

The elementary trees in Figure A.6 are needed for the analysis of (32).

(32) der dem König treue Diener
the.NOM the.DAT king loyal servant
'the servant loyal to the king'

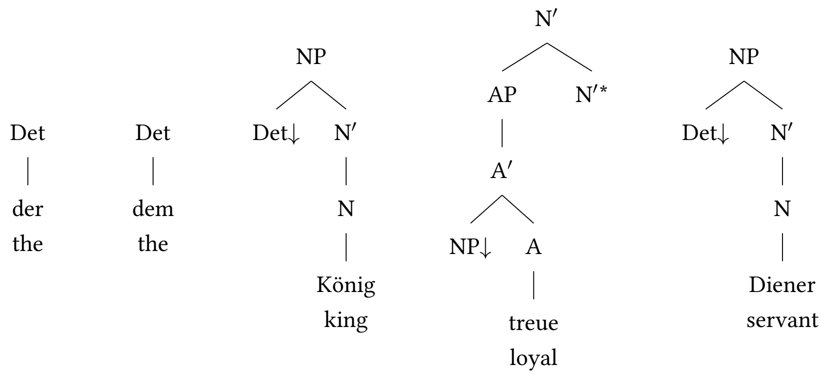


Figure A.6: Elementary trees for *der dem König treue Diener*

By substituting the tree for *dem* 'the' in the substitution node of *König* 'king', one then arrives at a full NP. This can then by inserted into the substitution node of *treue* 'loyal'. Similarly, the tree for *der* 'the' can be combined with the one for *Diener.* One then has both of the trees in Figure A.7.

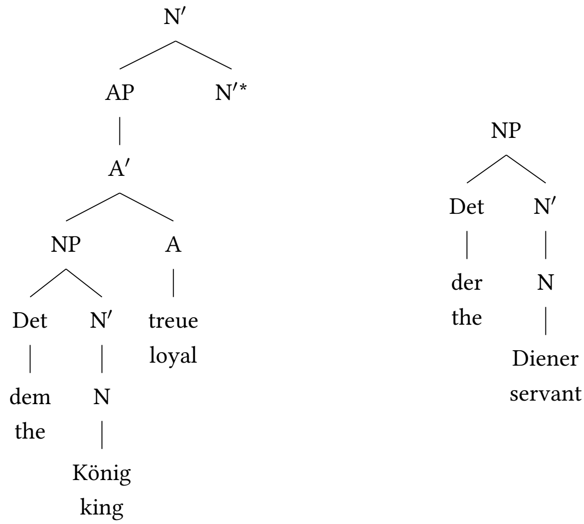


Figure A.7: Trees for *der dem König treue Diener* after substitution

The adjective tree can then be adjoined to the N′-node of *der Diener*, which yields the structure in Figure A.8 on the next page.

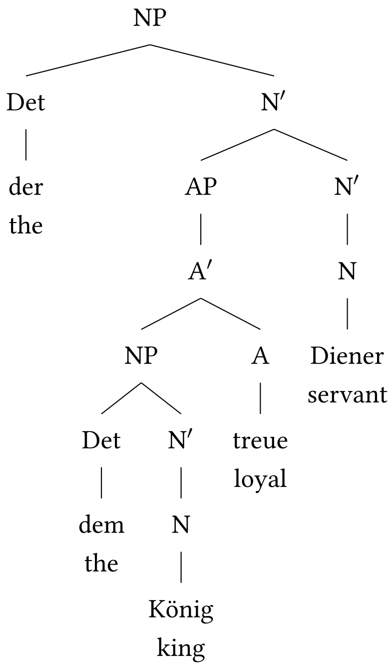


Figure A.8: Result of adjunction of the AP to the N′-node

Bibliography

Abbott, Barbara. 1976. Right node raising as a test for constituenthood. *Linguistic Inquiry* 7(4). 639–642.

Abeillé, Anne. 1988. Parsing French with Tree Adjoining Grammar: Some linguistic accounts. In *Proceedings of coling*, 7–12. Budapest. http://www.aclweb.org/anthology/C/C88/C88-1002.pdf.

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

Abeillé, Anne & Owen Rambow (eds.). 2000. *Tree Adjoining Grammars: Formalisms, linguistic analysis and processing* (CSLI Lecture Notes 156). Stanford, CA: CSLI Publications.

Abeillé, Anne & Yves Schabes. 1989. Parsing idioms in Lexicalized TAG. In Harold Somers & Mary McGee Wood (eds.), *Proceedings of the fourth conference of the european chapter of the association for computational linguistics*, 1–9. Manchester, England: Association for Computational Linguistics.

Abney, Steven P. 1987. *The english noun phrase in its sentential aspect.* Cambridge, MA: MIT dissertation. http://www.vinartus.net/spa/publications.html.

Abney, Steven P. 1996. Statistical methods and linguistics. In Judith L. Klavans & Philip Resnik (eds.), *The balancing act: Combining symbolic and statistical approaches to language* (Language, Speech, and Communication), 1–26. London, England/Cambridge, MA: MIT Press.

Abney, Steven P. & Jennifer Cole. 1986. A government-binding parser. In *Proceedings of north eastern linguistic society*, vol. 16, 1–17. University of Massachusetts, Amherst: GLSA.

Abraham, Werner. 1995. *Deutsche Syntax im Sprachenvergleich. Grundlegung einer typologischen Syntax des Deutschen.* Tübingen: Stauffenburg Verlag.

Abraham, Werner. 2003. The syntactic link between thema and rhema: the syntax-discourse interface. *Folia Linguistica* 37(1–2). 13–34.

Abraham, Werner. 2005. *Deutsche Syntax im Sprachenvergleich. Grundlegung einer typologischen Syntax des Deutschen* (Studien zur deutschen Grammatik 41). Tübingen: Stauffenburg Verlag 2nd edn.

Abzianidze, Lasha. 2011. *An HPSG-based formal grammar of a core fragment of Georgian implemented in TRALE.* Charles University in Prague MA thesis. https://sites.google.com/site/lashabzianidze/thesis.

Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates* (CSLI Lecture Notes 76). Stanford, CA: CSLI Publications.

Adams, Marianne. 1984. Multiple interrogation in Italian. *The Linguistic Review* 4(1). 1–27.

Ades, Anthony E. & Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4. 517–558.

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

Adger, David. 2010. A minimalist theory of feature structure. http://ling.auf.net/lingbuzz/000583/v2.pdf.

Adger, David. 2013. Constructions and grammatical explanation: Comments on goldberg. *Mind and Language* 28(4). 466–478.

Ágel, Vilmos. 2000. *Valenztheorie* (Narr Studienbücher). Tübingen: Gunter Narr Verlag.

Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2003. *Dependenz und Valenz / dependency and valency. Ein internationales Handbuch der zeitgenössischen Forschung / an international handbook of contemporary research*, vol. 25.1 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter.

Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2006. *Dependenz und Valenz / dependency and valency. Ein internationales Handbuch der zeitgenössischen Forschung / an international handbook of contemporary research / dependency and valency. ein internationales handbuch der zeitgenössischen forschung / an international handbook of contemporary research*, vol. 25.2 (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter.

Ágel, Vilmos & Klaus Fischer. 2010. 50 Jahre Valenztheorie und Dependenzgrammatik. *Zeitschrift für Germanistische Linguistik* 38(2). 249–290.

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

Alencar, Leonel Figueiredo de. 2004. Complementos verbais oracionais – uma análise léxicofuncional. *Lingua(gem)* 1(1). 173–218.

Alsina, Alex. 1996. Resultatives: A joint operation of semantic and syntactic structures. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference, rank xerox, grenoble*, Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/1/lfg1.html.

Alsina, Alex, KP Mohanan & Tara Mohanan. 2005. How to get rid of the COMP. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2005 conference*, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/10/lfg05amm.pdf.

Altmann, Hans & Ute Hofman. 2004. *Topologie fürs Examen: Verbstellung, Klammerstruktur, Stellungsfelder, Satzglied- und Wortstellung* (Linguistik fürs Examen 4). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH.

Ambridge, Ben & Adele E. Goldberg. 2008. The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19. 349–381. http://www.princeton.edu/~adele/Publications_files/08Ambridge%26Goldberg-islands.pdf.

Ambridge, Ben, Caroline F. Rowland & Julian M. Pine. 2008. Is structure dependence an innate constraint? new experimental evidence from children's complex-question production. *Cognitive Science: A Multidisciplinary Journal* 32(1). 222–255.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Aoun, Joseph & David W. Lightfoot. 1984. Government and contraction. *Linguistic Inquiry* 15(3). 465–473.

Aoun, Joseph & Dominique Sportiche. 1983. On the formal theory of government. *The Linguistic Review* 2(3). 211–236.

Arka, I Wayan, Avery Andrews, Mary Dalrymple, Meladel Mistica & Jane Simpson. 2009. A linguistic and computational morphosyntactic analysis for the applicative *-i* in Indonesian. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2009 conference*, 85–105. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/14/.

Arnold, Jennifer E., Michael K. Tanenhaus, Rebecca J. Altmann & Maria Fagnano. 2004. The old and thee, uh, new. *Psychological Science* 15(9). 578–582.

Askedal, John Ole. 1986. Zur vergleichenden Stellungsfelderanalyse von Verbalsätzen und nichtverbalen Satzgliedern. *Deutsch als Fremdsprache* 23. 269–273 and 342–348.

Asudeh, Ash. 2004. *Resumption as resource management*: Stanford University dissertation.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2008. Constructions with lexical integrity: Templates as the lexicon-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2008 conference*, Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/13/.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1(1). 1–54.

Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 68–88. Stanford, CA: CSLI Publications.

Asudeh, Ash & Ida Toivonen. 2014. *With* lexical integrity. *Theoretical Linguistics* 40(1–2). In Print.

Attardi, Giuseppe. 2006. Experiments with a multilanguage non-projective dependency parser. In *Proceedings of the 10th conference on computational natural language learning (conll-x)*, 166–170. Association for Computational Linguistics.

Attia, Mohammed A. 2008. *Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation*: School of Languages, Linguistics and Cultures, University of Manchester dissertation.

Avgustinova, Tania & Yi Zhang. 2009. Exploiting the russian national corpus in the development of a russian resource grammar. In Núria Bel, Erhard Hinrichs, Kiril Simov & Petya Osenova (eds.), *Adaptation of language resources and technology to new domains at the RANLP 2009 conference, borovets, bulgaria*, 1–11. Shoumen, Bulgaria: INCOMA Ltd.

Bach, Emmon. 1962. The order of elements in a Transformational Grammar of German. *Language* 8(3). 263–269.

Bach, Emmon. 1976. An extension of classical transformation grammar. In *Problems in linguistic metatheory, proceedings of the 1976 conference at the michigan state university*, 183–224.

Bahrani, Mohammad, Hossein Sameti & Mehdi Hafezi Manshadi. 2011. A computational grammar for Persian based on GPSG. *Language Resources and Evaluation* 45(4). 387–408.

Baker, Carl Lee. 1978. *Introduction to generative-transformational syntax*. Englewood Cliffs, NJ: Prentice-Hall.

Baker, Mark C. 2003. Linguistic differences and language design. *Trends in Cognitive Sciences* 7(8). 349–353.

Baker, Mark C. 2009. Language universals: Abstract but not mythological. *The Behavioral and Brain Sciences* 32(5). 448–449.

Baldridge, Jason. 2002. *Lexically specified derivational control in Combinatory Categorial Grammar*: University of Edinburgh dissertation. http://comp.ling.utexas.edu/jbaldrid/papers/Baldridge_dissertation.pdf.

Baldridge, Jason, Sudipta Chatterjee, Alexis Palmer & Ben Wing. 2007. Dotccg and VisCCG: Wiki and programming paradigms for improved grammar engineering with OpenCCG. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE), 5–25. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/.

Baldridge, Jason & Geert-Jan M. Kruijff. 2002. Coupling ccg and hybrid logic dependency semantics. In Pierre Isabelle (ed.), *40th annual meeting of the association for computational linguistics. proceedings of the conference*, 319–326. University of Pennsylvania, Philadelphia: Association for Computational Linguistics. http://www.aclweb.org/.

Ballweg, Joachim. 1997. Stellungsregularitäten in der Nominalphrase. In Eroms et al. (1997) 2062–2072. http://www.ids-mannheim.de/gra/grammar/.

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

Baltin, Mark. 2004. Remarks on the relation between language typology and Universal Grammar: Commentary on Newmeyer. *Studies in Language* 28(3). 549–553.

Baltin, Mark. 2006. Extraposition. In Martin Everaert, Henk van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds.), *The blackwell companion to syntax* (Blackwell Handbooks in Linguistics), 237–271. Oxford: Blackwell Publishing Ltd. http://www.nyu.edu/gsas/dept/lingu/people/faculty/baltin/papers/extrapos.pdf.

Bangalore, Srinivas, Aravind K. Joshi & Owen Rambow. 2003. Dependency and valency in other theories: Tree Adjoining Grammar. In Ágel et al. (2003) 669–678.

Bannard, Colin, Elena Lieven & Michael Tomasello. 2009. Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences* 106(41). 17284–17289.

Bar-Hillel, Yehoshua, Micha A. Perles & Eliahu Shamir. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und*

Kommunikationsforschung 14(2). 143–172.

Bartsch, Renate & Theo Vennemann. 1972. *Semantic structures. a study in the relation between semantics and syntax* (Athenäum-Skripten Linguistik 9). Frankfurt/Main: Athenäum.

Barwise, Jon & John Perry. 1983. *Situations and attitudes.* Cambridge, MA/London, England: MIT Press.

Barwise, Jon & John Perry. 1987. *Situationen und einstellungen – grundlagen der situationssemantik.* Berlin, New York: de Gruyter.

Baschung, K., G. G. Bes, A. Corluy & T. Guillotin. 1987. Auxiliaries and clitics in French UCG grammar. In Bente Maegaard (ed.), *Proceedings of the third conference of the european chapter of the association for computational linguistics*, 173–178. Copenhagen, Denmark: Association for Computational Linguistics.

Bates, Elizabeth A. 1984. Bioprograms and the innateness hypothesis. *The Behavioral and Brain Sciences* 7(2). 188–190.

Baumgärtner, Klaus. 1965. Spracherklärung mit den mitteln der abhängigkeitsstruktur. *Beiträge zur Sprachkunde und Informationsverarbeitung* 5. 31–53.

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

Bausewein, Karin. 1990. Haben kopflose Relativsätze tatsächlich keine Köpfe? In Gisbert Fanselow & Sascha W. Felix (eds.), *Strukturen und merkmale syntaktischer kategorien* (Studien zur deutschen Grammatik 39), 144–158. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Bayer, Josef & Jaklin Kornfilt. 1989. Restructuring effects in German. DYANA Report University of Edinburgh.

Beavers, John. 2003. A CCG implementation for the LKB. LinGO Working Paper 2002-08 CSLI Stanford Stanford, CA. https://webspace.utexas.edu/jbeavers/www/tccg-doc.ps.gz.

Beavers, John. 2004. Type-inheritance combinatory categorial grammar. In *Proceedings of COLING 2004*, 57–63. Geneva, Switzerland: COLING.

Beavers, John, Elias Ponvert & Stephen Mark Wechsler. 2008. Possession of a controlled substantive. In T. Friedman & S. Ito (eds.), *Proceedings of semantics and linguistic theory (salt) xviii*, 108–125. Ithaca, NY: Cornell University.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, center for computational linguistics, katholieke universiteit leuven*, 48–69. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2004/.

Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum* (Linguistische Arbeiten 139). Tübingen: Max Niemeyer Verlag. 2nd unchanged edition 1983.

Becker, Tilman, Aravind K. Joshi & Owen Rambow. 1991. Long-distance scrambling and Tree Adjoining Grammars. In *Fifth conference of the european chapter of the association* *for computational linguistics. proceedings of the conference*, 21–26. Berlin: Association for Computational Linguistics. http://www.aclweb.org/anthology/E91-1005.pdf.

Beermann, Dorothee & Lars Hellan. 2004. A treatment of directionals in two implemented HPSG grammars. In Stefan Müller (ed.), *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, center for computational linguistics, katholieke universiteit leuven*, 357–377. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2004/.

Beghelli, Filippo & Timothy Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In Anna Szabolcsi (ed.), *Ways of scope taking*, 71–107. Dordrecht/Boston/London: Kluwer Academic Publishers.

Behaghel, Otto. 1909. Beziehung zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.

Behaghel, Otto. 1930. Von deutscher Wortstellung. *Zeitschrift für Deutschkunde* 44. 81–89.

Behrens, Heike. 2009. Konstruktionen im Spracherwerb. *Zeitschrift für Germanistische Linguistik* 37(3). 427–444.

Bellugi, Ursula, Liz Lichtenberger, Wendy Jones, Zona Lai & Marie St. George. 2000. The neurocognitive profile of Williams Syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience* 12(Supplement). 7–29.

Bender, Emily & Daniel P. Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 199–214. Stanford, CA: CSLI Publications.

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

Bender, Emily M. 2008a. Evaluating a crosslinguistic grammar resource: A case study of Wambaya. In Johanna D. Moore, Simone Teufel, James Allan & Sadaoki Furui (eds.), *Proceedings of the 46th annual meeting of the association for computational linguistics: Human language technologies*, 977–985. Columbus, Ohio: Association for Computational Linguistics. http://aclweb.org/anthology-new/P/P08/P08-1111.pdf.

Bender, Emily M. 2008b. Grammar engineering for linguistic hypothesis testing. In Nicholas Gaylord, Alexis Palmer & Elias Ponvert (eds.), *Proceedings of the Texas Linguistics Society X conference: Computational linguistics for less-studied languages*, 16–36. Stanford CA: CSLI Publications ONLINE.

Bender, Emily M. 2008c. Radical non-configurationality without shuffle operators: An analysis of Wambaya. In Stefan Müller (ed.), *Proceedings of the 15th international conference on Head-Driven Phrase Structure Grammar*, 6–24. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2008/.

Bender, Emily M. 2010. Reweaving a grammar for Wambaya: A case study in grammar engineering for linguistic hypothesis testing. *Linguistic Issues in Language Technology – LiLT* 3(3). 1–34. http://elanguage.net/journals/index.php/lilt/article/view/662.

Bender, Emily M., Daniel P. Flickinger & Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the workshop on grammar engineering and evaluation at the 19th international conference on computational linguistics*, 8–14. Taipei, Taiwan.

Bender, Emily M. & Melanie Siegel. 2005. Implementing the syntax of Japanese numeral classifiers. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural language processing ijcnlp 2004* (Lecture Notes in Artificial Intelligence 3248), 626–635. Berlin/Heidelberg/New York, NY: Springer Verlag. http://www.melaniesiegel.de/publications/syntax-of-numeral-classifiers.pdf.

Bergen, Benjamin K. & Nancy Chang. 2005. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam/Philadelphia: John Benjamins Publishing Co.

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

Berman, Judith. 1999. Does German satisfy the Subject Condition? In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '99 conference, University of Manchester*, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/4/.

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

Berman, Judith. 2003b. Zum Einfluss der strukturellen Position auf die syntaktische Funktion der Komplementsätze. *Deutsche Sprache* 3. 263–286.

Berman, Judith. 2007. Functional identification of complement clauses in German and the specification of COMP. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling & Chris Manning (eds.), *Architectures, rules, and preferences. variations on themes by joan w. bresnan*, Stanford, CA: CSLI Publications.

Berwick, Robert C. 1982. Computational complexity and lexical-functional grammar. *American Journal of Computational Linguistics* 8(3–4). 97–109.

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

Berwick, Robert C. & Partha Niyogi. 1996. Learning from triggers. *Linguistic Inquiry* 27. 605–622.

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

Bickerton, Derek. 1984a. Creol is still king. *The Behavioral and Brain Sciences* 7(2). 212–218.

Bickerton, Derek. 1984b. The language bioprogram hypothesis. *The Behavioral and Brain Sciences* 7(2). 173–188.

Bickerton, Derek. 1997. How to acquire language without positive evidence: What acquisitionists can learn from creoles. In Michel DeGraff (ed.), *Language creation and language change: Creolization, diachrony, and development* (Learning, Development, and Conceptual Change), 49–74. Cambridge, MA: MIT Press.

Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs* (studia grammatica 2). Berlin: Akademie Verlag.

Bierwisch, Manfred. 1992. Grammatikforschung in der DDR: Auch ein Rückblick. *Linguistische Berichte* 139. 169–181.

Bildhauer, Felix. 2008. *Representing information structure in an HPSG grammar of Spanish*: Universität Bremen Dissertation.

Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topichood. 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.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491. http://www.hcrc.ed.ac.uk/~ewan/Papers/.

Bishop, Dorothy V.M. 2002. Putting language genes in perspective. *TRENDS in Genetics* 18(2).

Bjerre, Tavs. 2006. Object positions in a topological sentence model for Danish – a linearization-based HPSG approach. Presentation at Ph.D.-course at Sandbjerg, Denmark. http://www.hum.au.dk/engelsk/engsv/objectpositions/workshop/Bjerre.pdf.

Blackburn, Patrick & Johan Bos. 2005. *Representation and inference for natural language. a first course in computational semantics.* Stanford, CA: CSLI Publications.

Blackburn, Patrick, Claire Gardent & Wilfried Meyer-Viol. 1993. Talking about trees. In Steven Krauwer, Michael Moortgat & Louis des Tombe (eds.), *Sixth conference of the european chapter of the association for computational linguistics. proceedings of the conference*, 21–29. Uetrecht: Association for Computational Linguistics.

Błaszczak, Joanna & Hans-Martin Gärtner. 2005. Intonational phrasing, discontinuity, and the scope of negation. *Syntax* 8(1). 1–22.

Blevins, James P. 2003. Passives and impersonals. *Journal of Linguistics* 39(3). 473–520.

Block, Hans-Ulrich & Rudolf Hunze. 1986. Incremental construction of c- and f-structure in a LFG-parser. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 490–493. University of Bonn: Association for Computational Linguistics.

Blom, Corrien. 2005. *Complex predicates in Dutch. synchrony and diachrony* (LOT Dissertation Series 111). Utrecht University. http://www.lotpublications.nl/publish/issues/BlomC/.

Bloom, Paul. 1993. Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry* 24(4). 721–734.

Boas, Hans C. 2003. *A constructional approach to resultatives* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Boas, Hans C. 2014. Lexical approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40(1–2). 89–112.

Bobaljik, Jonathan. 1999. Adverbs: The hierarchy paradox. *Glot International* 4(9/10). 27–28.

Bod, Rens. 2009a. Constructions at work or at rest? *Cognitive Linguistics* 20(1). 129–134.

Bod, Rens. 2009b. From exemplar to grammar: Integrating analogy and probability in language learning. *Cognitive Science* 33(4). 752–793. http://staff.science.uva.nl/~rens/analogy.pdf.

Bögel, Tina, Miriam Butt & Sebastian Sulger. 2008. Urdu ezafe and the morphology-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2008 conference*, 129–149. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/13/.

Bolc, Leonard, Krzysztof Czuba, Anna Kupść, Małgorzata Marciniak, Agnieszka Mykowiecka & Adam Przepiórkowski. 1996. A survey of systems for implementing hpsg grammars. Tech. Rep. 814 Institute of Computer Science, Polish Academy of Sciences Warsaw, Poland. http://www.cs.cmu.edu/~kczuba/systems-wide.ps.gz.

Booij, Geert E. 2002. Separable complex verbs in Dutch: A case of periphrastic word formation. In Nicole Dehé, Ray S. Jackendoff, Andrew McIntyre & Silke Urban (eds.), *Verb-particle explorations* (Interface Explorations 1), 21–41. Berlin/New York, NY: Mouton de Gruyter.

Booij, Geert E. 2005. Construction-dependent morphology. *Lingue e linguaggio* 4. 31–46.

Booij, Geert E. 2009. Lexical integrity as a formal universal: A constructionist view. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today* (Studies in Natural Language and Linguistic Theory 76), 83–100. Berlin/Heidelberg/New York, NY: Springer Verlag.

Booij, Geert E. 2010. Construction morphology. *Language and Linguistics Compass* 4(7). 543–555. doi:10.1111/j.1749-818X.2010.00213.x.

Booij, Geert E. To appear. Construction morphology. In Andrew Spencer & Arnold Zwicky (eds.), *The handbook of morphology*, London: Blackwell Publishing Ltd 2nd edn. http://www.hum2.leidenuniv.nl/booijge/pdf/Spencer%20Zwicky%20handbook%20Construction%20morphology.pdf.

Borer, Hagit. 1994. The projection of arguments. In E. Benedicto & J. Runner (eds.), *Functional projections* (UMass Occasional Papers in Linguistics (UMOP) 17), 19–47. Massachusetts: University of Massachusetts Graduate Linguistic Student Association.

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory*, 31–67. Stanford, CA: CSLI Publications.

Borer, Hagit. 2005. *Structuring sense*, vol. I. In Name Only. Oxford: Oxford University Press.

Borsley, Robert D. 1989. Phrase-structure grammar and the barriers conception of clause structure. *Linguistics* 27. 843–863.

Borsley, Robert D. 1991. *Syntax-theory: A unified approach.* London: Edward Arnold.

Borsley, Robert D. 1999. *Syntactic theory: A unified approach*. London: Edward Arnold 2nd edn.

Borsley, Robert D. 2005. Against ConjP. *Lingua* 115(4). 461–482.

Borsley, Robert D. 2006. Syntactic and lexical approaches to unbounded dependencies. Essex Research Reports in Linguistics 49 University of Essex. http://www.essex.ac.uk/linguistics/errl/errl_49.pdf.

Borsley, Robert D. 2007. Hang on again! are we 'on the right track'? In Andrew Radford (ed.), *Martin Atkinson – the Minimalist muse* (Essex Research Reports in Linguistics 53), 43–69. Department of Language and Linguistics, University of Essex.

Borsley, Robert D. 2012. Don't move! *Iberia: An International Journal of Theoretical Linguistics* 4(1). 110–139.

Bos, Johan. 1996. Predicate logic unplugged. In Paul J. E. Dekker & M. Stokhof (eds.), *Proceedings of the tenth amsterdam colloquium*, 133–143. Amsterdam: ILLC/Department of Philosophy, University of Amsterdam.

Boukedi, Sirine & Kais Haddar. 2014. Hpsg grammar treating of different forms of arabic coordination. *Research in Computing Science* 86: Advances in Computational Linguistics and Intelligent Decision Making. 25–41. http://www.micai.org/rcs/2014_86/HPSG%20Grammar%20Treating%20of%20Different%20Forms%20of%20Arabic%20Coordination.pdf.

Boullier, Pierre & Benoît Sagot. 2005a. Analyse syntaxique profonde à grande échelle: SxLfg. *Traitement Automatique des Langues (T.A.L.)* 46(2). 65–89.

Boullier, Pierre & Benoît Sagot. 2005b. Efficient and robust LFG parsing: SxLfg. In *Proceedings of iwpt 2005*, 1–10. Vancouver, Canada: Association for Computational Linguistics. http://atoll.inria.fr/~sagot/pub/IWPT05.pdf.

Boullier, Pierre, Benoît Sagot & Lionel Clément. 2005. Un analyseur LFG efficace pour le français: SxLfg. In *Actes de taln 05*, 403–408. Dourdan, France.

Bouma, Gosse. 1996. Extraposition as a nonlocal dependency. In Geert-Jan Kruijff, Glynn V. Morrill & Dick Oehrle (eds.), *Proceedings of formal grammar 96*, 1–14. Prag. http://www.let.rug.nl/gosse/papers.html.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001a. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65. http://ftp-linguistics.stanford.edu/sag/bms-nllt.pdf.

Bouma, Gosse & Gertjan van Noord. 1994. Constraint-based Categorial Grammar. In James Pustejovsky (ed.), *32th annual meeting of the association for computational linguistics. proceedings of the conference*, 147–154. Las Cruses: Association for Computational Linguistics.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press. http://www.let.rug.nl/~vannoord/papers/.

Bouma, Gosse, Gertjan van Noord & Robert Malouf. 2001b. Alpino: Wide-coverage computational analysis of Dutch. In Walter Daelemans, Khalil Sima'an, Jorn Veenstra & Jakub Zavrel (eds.), *Computational linguistics in the netherlands 2000. selected papers* *from the eleventh clin meeting* (Language and Computers 37), Amsterdam/New York, NY: Rodopi.

Braine, Martin D. S. 1987. What is learned in acquiring word classes—a step toward an acquisition theory. In Brian MacWhinny (ed.), *Mechanisms of language acquisition*, 65–87. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Branco, António & Francisco Costa. 2008a. A computational grammar for deep linguistic processing of portuguese: LXGram, version A.4.1. Tech. Rep. TR-2008-17 Universidade de Lisboa, Faculdade de Ciências, Departamento de Informática. http://hdl.handle.net/10455/3008.

Branco, António & Francisco Costa. 2008b. LXGram in the shared task 'comparing semantic representations' of STEP 2008. In Johan Bos & Rodolfo Delmonte (eds.), *Semantics in text processing. STEP 2008 conference proceedings*, vol. 1 (Research in Computational Semantics), 299–314. College Publications. http://www.aclweb.org/anthology/W08-2224.

Brants, Sabine, Stefanie Dipper, Peter Eisenberg, Silvia Hansen-Schirra, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith & Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. *Research on Language and Computation* 2(4). 597–620.

Bresnan, Joan. 1974. The position of certain clause-particles in phrase structure. *Linguistic Inquiry* 5(4). 614–619.

Bresnan, Joan. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan & G. A. Miller (eds.), *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: MIT Press.

Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). 343–434.

Bresnan, Joan. 1982b. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 3–86. Cambridge, MA/London: MIT Press.

Bresnan, Joan. 2001. *Lexical-functional syntax*. Oxford, UK/Cambridge, USA: Blackwell.

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

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

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

Bresnan, Joan & Sam A. Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–254.

Bresnan, Joan & Annie Zaenen. 1990. Deep unaccusativity in LFG. In Katarzyna Dziwirek, Patrick Farrell & Errapel Mejías-Bikandi (eds.), *Grammatical relations: A cross-theoretical perspective*, 45–57. Stanford, CA: CSLI Publications.

Brew, Chris. 1995. Stochastic HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the seventh conference of the European Chapter of the Association for*

Computational Linguistics, 83–89. Dublin: Association for Computational Linguistics.
Briscoe, Ted J. 1997. Review of edward p. stabler, jr., the logical approach to syntax: Foundations, specifications, and implementations of theories of Government and Binding. *Journal of Linguistics* 33(1). 223–225.
Briscoe, Ted J. 2000. Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language* 76(2). 245–296.
Briscoe, Ted J. & Ann Copestake. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25(4). 487–526. http://acl.ldc.upenn.edu/J/J99/J99-4002.pdf.
Bröker, Norbert. 2003. Formal foundations of dependency grammar. In Ágel et al. (2003) 294–310.
Brosziewski, Ulf. 2003. *Syntactic derivations: A nontransformational view* (Linguistische Arbeiten 470). Tübingen: Max Niemeyer Verlag.
Brown, Roger & Camille Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In John R. Hayes (ed.), *Cognition and the development of language*, New York: John Wiley & Sons, Inc.
Bruening, Benjamin. 2009. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. In Laurel MacKenzie (ed.), *Proceedings of the 32th annual penn linguistics colloquium* (Penn Working Papers in Linguistics 15.1), 26–35. Philadelphia.
Bryant, John. 2003. *Constructional analysis*. University of Californa at Berkeley MA thesis. http://www.icsi.berkeley.edu/~jbryant/old-analyzer.html.
Budde, Monika. 2010. Konstruktionen integrativ: Topik-Konstruktionen als reinsyntaktisches Pendant zu den lexikalisch verankerten Komplement-Konstruktionen. Vortrag auf der Tagung Konstruktionsgrammatik: Neue Perspektiven zur Untersuchung des Deutschen und Englischen. Internationale Fachtagung an der Christian-Albrechts-Universität zu Kiel vom 18. bis 20. Februar 2010. http://snipurl.com/konstruktionsgrammatik.
Burzio, Luigi. 1981. *Intransitive verbs and italian auxiliaries*: MIT dissertation.
Burzio, Luigi. 1986. *Italian syntax: A government-binding approach* (Studies in Natural Language and Linguistic Theory 1). Dordrecht: Holland, Boston: U.S.A, London: England: D. Reidel Publishing Company.
Busemann, Stephan. 1992. *Generierung natürlicher sprache mit generalisierten phrasenstrukturgrammatiken*, vol. 313 (Informatik-Fachberichte). Berlin/Heidelberg/New York, NY: Springer Verlag.
Bußmann, Hadumod (ed.). 1983. *Lexikon der Sprachwissenschaft* (Kröners Taschenausgabe 452). Stuttgart: Alfred Kröner Verlag.
Bußmann, Hadumod (ed.). 1990. *Lexikon der Sprachwissenschaft* (Kröners Taschenausgabe 452). Stuttgart: Alfred Kröner Verlag 2nd edn.
Butt, Miriam. 2003. The light verb jungle. In *Papers from the harvard/dudley house light verb workshop* (Harvard Working Papers in Linguistics 9), 1–49. Harvard University, Department of Linguistics.
Butt, Miriam, Stefanie Dipper, Anette Frank & Tracy Holloway King. 1999a. Writing large-scale parallel grammars for English, French and German. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '99 conference, University of Manch-*

ester, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/4/.

Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi & Christian Rohrer. 2002. The parallel grammar project. In *Proceedings of coling-2002 workshop on grammar engineering and evaluation*, 1–7.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño & Frédérique Segond. 1999b. *A grammar writer's cookbook* (CSLI Lecture Notes 95). Stanford, CA: CSLI Publications.

Butt, Miriam, Tracy Holloway King & Sebastian Roth. 2007. Urdu correlatives: Theoretical and implementational issues. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2007 conference*, 107–127. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/12/lfg07.html.

Cahill, Aoife, Michael Burke, Martin Forst, Ruth O'Donovan, Christian Rohrer, Josef van Genabith & Andy Way. 2005. Treebank-based acquisition of multilingual unification grammar resources. *Research on Language and Computation* 3(2). 247–279.

Cahill, Aoife, Michael Burke, Ruth O'Donovan, Stefan Riezler, Josef van Genabith & Andy Way. 2008. Wide-coverage deep statistical parsing using automatic dependency structure annotation. *Computational Linguistics* 34(1). 81–124.

Calder, Jonathan, Ewan Klein & Henk Zeevat. 1988. Unification categorial grammar: A concise, extendable grammar for natural language processing. In Dénes Vargha (ed.), *Proceedings of COLING 88*, 83–86. University of Budapest: Association for Computational Linguistics.

Callmeier, Ulrich. 2000. PET—A platform for experimentation with efficient HPSG processingtechniques. *Journal of Natural Language Engineering* 1(6). 99–108. (Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation).

Candito, Marie-Hélène. 1996. A principle-based hierarchical representation of LTAGs. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 194–199. Copenhagen, Denmark: Association for Computational Linguistics.

Candito, Marie-Hélène. 1998. Building parallel LTAG for French and Italian. In *Proceedings of the 36th annual meeting of the association for computational linguistics and 17th international conference on computational linguistics*, 211–217. Montreal, Quebec, Canada: Association for Computational Linguistics.

Candito, Marie-Hélène. 1999. *Organisation modulaire et paramétrable de grammaires électroniques lexicalisées. application au français et à l'italien*: Université Paris 7 dissertation.

Candito, Marie-Hélène & Sylvain Kahane. 1998. Can the TAG derivation tree represent a semantic graph? An answer in the light of meaning-text theory. In *Tag+4*, 25–28.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions". *Constructions online* (SV1–7). 1–28.

Cappelle, Bert, Yury Shtyrov & Friedemann Pulvermüller. 2010. *Heating up* or *cooling up* the brain? MEG evidence that phrasal verbs are lexical units. *Brain and Language* 115. 189–201.

Carlson, Gregory N. & Michael K. Tanenhaus. 1988. Thematic roles and language comprehension. In Wendy Wilkins (ed.), *Thematic relations* (Syntax and Semantics 21),

263–289. San Diego: Academic Press.
Carpenter, Bob. 1992. *The logic of typed feature structures* (Tracts in Theoretical Computer Science). Cambridge: Cambridge University Press.
Carpenter, Bob. 1994. A natural deduction theorem prover for type-theoretic Categorial Grammars. Tech. rep. Carnegie Mellon Laboratory for Computational Linguistics. http://www.essex.ac.uk/linguistics/external/clmt/papers/cg/carp_cgparser_doc.ps.
Carpenter, Bob. 1998. *Type-logical semantics*. Cambridge, MA/London, England: MIT Press.
Carpenter, Bob & Gerald Penn. 1996. Efficient parsing of compiled typed attribute value logic grammars. In Harry Bunt & Masaru Tomita (eds.), *Recent advances in parsing technology* (Text, Speech and Language Technology 1), 145–168. Dordrecht/Boston/London: Kluwer Academic Publishers.
Çetinoğlu, Özlem & Kemal Oflazer. 2006. Morphology-syntax interface for Turkish LFG. In Nicoletta Calzolari, Claire Cardie & Pierre Isabelle (eds.), *Proceedings of the 21st international conference on computational linguistics and 44th annual meeting of the association for computational linguistics*, 153–160. Sydney, Australia: Association for Computational Linguistics. http://www.aclweb.org/anthology/P/P06/P06-1/.
Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 47–67. Stanford, CA: CSLI Publications.
Choi, Hye-Won. 1999. *Optimizing structure in scrambling. scrambling and information structure* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.
Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.
Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2(2). 137–167.
Chomsky, Noam. 1964. Degrees of grammaticalness. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The structure of language*, 384–389. Englewood Cliffs, NJ: Prentice-Hall.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1968. Language and the mind. *Psychology Today* 1(9). 48–68. Reprint as: Chomsky (1976a).
Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in english transformational grammar*, chap. 12, 184–221. Waltham, MA/Toronto/London: Ginn and Company.
Chomsky, Noam. 1971. *Problems of knowledge and freedom*. London: Fontana.
Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparski (eds.), *A festschrift for morris halle*, 232–286. New York: Holt, Rinehart & Winston.
Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.

Chomsky, Noam. 1976a. Language and the mind. In Diane D. Borstein (ed.), *Readings in the theory of grammar: From the 17th to the 20th century*, 241–251. Cambridge, MA: Winthrop. Reprint from: Chomsky (1968).

Chomsky, Noam. 1976b. *Reflections on language*. New York: Pantheon Books.

Chomsky, Noam. 1977. *Essays on form and interpretation*. New York: North Holland.

Chomsky, Noam. 1980. *Rules and representations*. Oxford: Basil Blackwell.

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

Chomsky, Noam. 1981b. Reply to comments of thompson. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 295(1077). 277–281.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of Government and Binding* (Linguistic Inquiry Monographs 5). Cambridge, MA/London, England: MIT Press.

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

Chomsky, Noam. 1986b. *Knowledge of language – its nature, origin, and use* (Convergence). New York, Westport: Connecticut, London: Praeger.

Chomsky, Noam. 1988. *Language and problems of knowledge. the Managua lectures* (Current Studies in Linguistics 16). Cambridge, MA: MIT Press.

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

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in generative grammar*, 417–454. Cambridge, MA: MIT Press. Reprint as: Chomsky (1995: 129–166).

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

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

Chomsky, Noam. 1998. Noam Chomsky's Minimalist Program and the philosophy of mind. An interview [with] Camilo J. Cela-Conde and Gisde Marty. *Syntax* 1(1). 19–36.

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

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

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

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

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

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? chomsky's minimalism and the view from syntax-semantics* (Studies in Generative Grammar 89), 1–29. Berlin/New York, NY: Mouton de Gruyter.

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

Chomsky, Noam. 2010. Restricting stipulations: Consequences and challenges. Vortrag in Stuttgart.

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

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: John Wiley & Sons, Inc.

Chouinard, Michelle M. & Eve V. Clark. 2003. Adult reformulations of child errors as negative evidence. *Journal of Child Language* 30. 637–669.

Chrupala, Grzegorz & Josef van Genabith. 2006. Using machine-learning to assign function labels to parser output for spanish. In Nicoletta Calzolari, Claire Cardie & Pierre Isabelle (eds.), *Proceedings of the 21st international conference on computational linguistics and 44th annual meeting of the association for computational linguistics*, 136–143. Sydney, Australia: Association for Computational Linguistics. http://www.aclweb.org/anthology/P/P06/P06-1/.

Chung, Sandra & James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14. 704–713.

Church, Kenneth. 2011. A pendulum swung too far. *Linguistic Issues in Language Technology* 6(5). http://elanguage.net/journals/index.php/lilt/issue/view/330. Special Issue on Interaction of Linguistics and Computational Linguistics.

Cinque, Guglielmo. 1994. On the evidence for partial n movement in the romance dp. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raffaella Zanuttini (eds.), *Paths towards universal grammar. studies in honor of richard s. kayne*, 85–110. Washington, D.C.: Georgetown University Press.

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

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

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

Clark, Herbert H. & Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84(1). 73–111. http://www-psych.stanford.edu/~herb/2000s/Clark.FoxTree.02.pdf.

Clark, Herbert H. & Thomas Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37(3). 201–242. http://www-psych.stanford.edu/~herb/1990s/Clark.Wasow.98.pdf.

Clark, Stephen & James Curran. 2007. Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics* 33(4). 493–552.

Clark, Stephen, Julia Hockenmaier & Mark J. Steedman. 2002. Building deep dependency structures with a wide-coverage CCG parser. In Pierre Isabelle (ed.), *40th annual meeting of the association for computational linguistics. proceedings of the conference*, 327–334. University of Pennsylvania, Philadelphia: Association for Computational Linguistics. http://www.aclweb.org/.

Clément, Lionel. 2009. XLFG5 documentation. Translated from French by Olivier Bonami. http://www.xlfg.org/.

Clément, Lionel & Alexandra Kinyon. 2001. XLFG—an LFG parsing scheme for French. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2001 conference*, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/6/lfg01.html.

Clément, Lionel & Alexandra Kinyon. 2003. Generating parallel multilingual LFG-TAG grammars from a MetaGrammar. In Erhard Hinrichs & Dan Roth (eds.), *Proceedings of the 41st annual meeting of the association for computational linguistics*, 184–191. Sapporo, Japan: Association for Computational Linguistics.

Clifton, Charles Jr. & Penelope Odom. 1966. Similarity relations among certain English sentence constructions. *Psychological Monographs: General and Applied* 80(5). 1–35.

Cook, Philippa Helen. 2001. *Coherence in German: An information structure approach*: Departments of Linguistics and German, University of Manchester dissertation.

Cook, Philippa Helen. 2006. The datives that aren't born equal: Beneficiaries and the dative passive. In Daniel Hole, André Meinunger & Werner Abraham (eds.), *Datives and similar cases: Between argument structure and event structure*, 141–184. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Cooper, Robin, Kuniaki Mukai & John Perry (eds.). 1990. *Situation theory and its applications, volume 1* (CSLI Lecture Notes 22). Stanford, CA: CSLI Publications.

Coopmans, Peter. 1989. Where stylistic and syntactic processes meet: Locative inversion in English. *Language* 65(4). 728–751.

Copestake, Ann. 2002. *Implementing typed feature structure grammars* (CSLI Lecture Notes 110). Stanford, CA: CSLI Publications.

Copestake, Ann. 2007. Applying robust semantics. In *Proceedings of the 10th conference of the Pacific Assocation for Computational Linguistics (PACLING)*, 1–12.

Copestake, Ann & Ted Briscoe. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12(1). 15–67.

Copestake, Ann & Ted J. Briscoe. 1992. Lexical operations in a unification based framework. In James Pustejovsky & Sabine Bergler (eds.), *Lexical semantics and knowledge representation* (Lecture Notes in Artificial Intelligence 627), 101–119. Berlin/Heidelberg/New York, NY: Springer Verlag. http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html.

Copestake, Ann & Daniel P. Flickinger. 2000. An open-source grammar development environment and broad-coverage English grammar using HPSG. In *Proceedings of the Second Linguistic Resources and Evaluation Conference*, 591–600. Athens, Greece.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal recursion semantics: an introduction. *Research on Language and Computation* 4(3). 281–332. http://lingo.stanford.edu/sag/papers/copestake.pdf.

Corbett, Greville G. 2008. Number of genders. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures (wals)*, chap. 30, 126–129. Oxford: Oxford University Press. http://wals.info/feature/description/30.

Correa, Nelson. 1987. An attribute-grammar implementation of Government-Binding Theory. In Candy Sidner (ed.), *25th annual meeting of the association for computational linguistics*, 45–51. Stanford, CA: Association for Computational Linguistics. http://acl.ldc.upenn.edu/P/P87/P87-1007.pdf.

Costa, Francisco & António Branco. 2010. LXGram: A deep linguistic processing grammar for Portuguese. In Thiago A.S. Pardo (ed.), *Computational processing of the portuguese language: 9th international conference, propor 2010, porto alegre, rs, brazil, april 27-30, 2010. proceedings* (Lecture Notes in Artificial Intelligence 6001), 86–89. Berlin/Heidelberg/New York, NY: Springer Verlag.

Covington, Michael A. 1990. Parsing discontinuous constituents in Dependency Grammar. *Computational Linguistics* 16(4). 234–236. http://dl.acm.org/citation.cfm?id=124992.124997.

Crabbé, Benoit. 2005. *Représentation informatique de grammaires d'arbres fortement lexicalisées: le cas de la grammaire d'arbres adjoints*: Université Nancy 2 dissertation.

Crain, Stephen, Drew Khlentzos & Rosalind Thornton. 2010. Universal Grammar versus language diversity. *Lingua* 120(12). 2668–2672.

Crain, Stephen & Mineharu Nakayama. 1987. Structure dependence in grammar formation. *Language* 63(3). 522–543.

Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, Cambridge, UK: Cambridge University Press.

Crain, Stephen, Rosalind Thornton & Drew Khlentzos. 2009. The case of the missing generalizations. *Cognitive Linguistics* 20(1). 145–155.

Cramer, Bart & Yi Zhang. 2009. Construction of a German HPSG grammar from a detailed treebank. In Tracy Holloway King & Marianne Santaholma (eds.), *Proceedings of the 2009 workshop on grammar engineering across frameworks (geaf 2009)*, 37–45. Suntec, Singapore: Association for Computational Linguistics. http://www.aclweb.org/anthology/W/W09/#2600.

Crocker, Matthew Walter & Ian Lewin. 1992. Parsing as deduction: Rules versus principles. In Bernd Neumann (ed.), *Ecai 92. l0th european conference on artificial intelligence*, 508–512. John Wiley and Sons, Ltd.

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language:*

Studies in honour of Günter Radden, 49–68. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Croft, William. 2009. Syntax is more diverse, and evolutionary linguistics is already here. *The Behavioral and Brain Sciences* 32(5). 457–458.

Crysmann, Berthold. 2001. Clitics and coordination in linear structure. In Birgit Gerlach & Janet Grijzenhout (eds.), *Clitics in phonology, morphology and syntax* (Linguistik Aktuell / Linguistics Today 36), Amsterdam/Philadelphia: John Benjamins Publishing Co. http://www.coli.uni-sb.de/~crysmann/papers/ClCoord.html.

Crysmann, Berthold. 2002. *Constraint-based co-analysis: Portuguese cliticisation and morphology-syntax interaction in HPSG* (Saarbrücken Dissertations in Computational Linguistics and Language Technology 15). Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz undUniversität des Saarlandes.

Crysmann, Berthold. 2003. On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, 112–116. Borovets, Bulgaria.

Crysmann, Berthold. 2004. Underspecification of intersective modifier attachment: Some arguments from German. In Stefan Müller (ed.), *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, center for computational linguistics, katholieke universiteit leuven*, 378–392. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2004/.

Crysmann, Berthold. 2005a. An inflectional approach to Hausa final vowel shortening. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2004*, 73–112. Dordrecht: Kluwer Academic Publishers.

Crysmann, Berthold. 2005b. Relative clause extraposition in German: An efficient and portable implementation. *Research on Language and Computation* 1(3). 61–82.

Crysmann, Berthold. 2005c. Syncretism in German: a unified approach to underspecification, indeterminacy, and likeness of case. In Stefan Müller (ed.), *Proceedings of the 12th international conference on Head-Driven Phrase Structure Grammar, department of informatics, university of lisbon*, 91–107. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2005/.

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

Crysmann, Berthold. 2009. Autosegmental representations in an HPSG of Hausa. In Tracy Holloway King & Marianne Santaholma (eds.), *Proceedings of the 2009 workshop on grammar engineering across frameworks (geaf 2009)*, 28–36. Suntec, Singapore: Association for Computational Linguistics. http://www.aclweb.org/anthology/W/W09/#2600.

Crysmann, Berthold. 2011. A unified account of Hausa genitive constructions. In Philippe de Groote, Markus Egg & Laura Kallmeyer (eds.), *Formal grammar. 14th international conference, fg 2009, bordeaux, france, july 25–26, 2009, revised selected papers* (Lecture Notes in Artificial Intelligence 5591), 102–117. Berlin/Heidelberg/New York, NY: Springer Verlag.

Crysmann, Berthold. 2012. Hag – a computational grammar of Hausa. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Selected proceedings of the 42nd annual conference on african linguistics (acal 42)*, 321–337. Somerville, MA: Cascadilla Press. http://www.lingref.com/cpp/acal/42/paper2780.pdf.

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*, vol. 1 (Foundations of Syntax). Oxford: Oxford University Press.

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

Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8. 345–351.

Curtiss, Susan. 1977. *Genie: A psycholinguistic study of a modern-day "wild child"*. Academic Press.

Dąbrowska, Ewa. 2001. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11(1–2). 83–102.

Dąbrowska, Ewa. 2004. *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Washington, D.C.: Georgetown University Press.

Dahl, Östen & Viveka Velupillai. 2013a. *The past tense*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/66.

Dahl, Östen & Viveka Velupillai. 2013b. *Perfective/imperfective aspect*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/65.

Dahllöf, Mats. 2002. Token dependency semantics and the paratactic analysis of intensional constructions. *Journal of Semantics* 19(4). 333–368.

Dahllöf, Mats. 2003. Two reports on computational syntax and semantics. Reports from Uppsala University (RUUL) 36 Department of Linguistics. http://stp.ling.uu.se/~matsd/pub/ruul36.pdf.

Dalrymple, Mary. 1993. *The syntax of anaphoric binding* (CSLI Lecture Notes 36). Stanford, CA: CSLI Publications.

Dalrymple, Mary (ed.). 1999. *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary. 2001. *Lexical functional grammar* (Syntax and Semantics 34). New York: Academic Press.

Dalrymple, Mary. 2006. Lexical functional grammar. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, Oxford: Elsevier Science Publisher B.V. (North-Holland) 2nd edn. http://eprints.ouls.ox.ac.uk/archive/00001101/.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2004 conference*, 199–208. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/9/lfg04.html.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in lexical-functional grammar* (CSLI Lecture Notes 47). Stanford, CA: CSLI Publications.

Dalrymple, Mary, John Lamping & Vijay Saraswat. 1993. LFG semantics via constraints. In Steven Krauwer, Michael Moortgat & Louis des Tombe (eds.), *Sixth conference of the european chapter of the association for computational linguistics. proceedings of the conference*, 97–105. Uetrecht: Association for Computational Linguistics. doi:http://dx.doi.org/10.3115/976744.976757.

Dalrymple, Mary, Maria Liakata & Lisa Mackie. 2006. Tokenization and morphological analysis for malagasy. *Computational Linguistics and Chinese Language Processing* 11(4). 315–332.

Dalrymple, Mary & Helge Lødrup. 2000. The grammatical functions of complement clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2000 conference*, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/5/lfg00dalrympl-lodrup.pdf.

Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81–95. Pittsburg University Press.

Davis, Anthony R. 1996. *Lexical semantics and linking in the hierarchical lexicon*: Stanford University dissertation.

Davis, Anthony R. & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.

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

De Kuthy, Kordula. 2001. Splitting PPs from NPs. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 31–76. Stanford, CA: CSLI Publications.

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

De Kuthy, Kordula, Vanessa Metcalf & Walt Detmar Meurers. 2004. Documentation of the implementation of the Milca English Resource Grammar in theTrale system. Ohio State University, ms.

De Kuthy, Kordula & Walt Detmar Meurers. 2001. On partial constituent fronting in German. *Journal of Comparative Germanic Linguistics* 3(3). 143–205. http://www.ling.ohio-state.edu/~kdk/papers/dekuthy-meurers-jcgs.html.

De Kuthy, Kordula & Walt Detmar Meurers. 2003a. Dealing with optional complements in HPSG-based grammar implementations. In Stefan Müller (ed.), *Proceedings of the 10th international conference on Head-Driven Phrase Structure Grammar, michigan state university, east lansing*, 88–96. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2003/.

De Kuthy, Kordula & Walt Detmar Meurers. 2003b. The secret life of focus exponents, and what it tells us about fronted verbal projections. In Stefan Müller (ed.), *Proceedings of the 10th international conference on Head-Driven Phrase Structure Grammar, michigan state university, east lansing*, 97–110. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2003/.

de Saussure, Ferdinand. 1916a. *Cours de linguistique générale* (Bibliothèque Scientifique Payot). Payot. Edited by Charles Bally and Albert Sechehaye.

de Saussure, Ferdinand. 1916b. *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin: Walter de Gruyter & Co. 2nd edition 1967.

Delmonte, Rodolfo. 1990. Semantic parsing with an LFG-based lexicon and conceptual representations. *Computers and the Humanities* 24(5–6). 461–488.

Demske, Ulrike. 2001. *Merkmale und Relationen. Diachrone Studien zur Nominalphrase des Deutschen* (Studia Linguistica Germanica 56). Berlin/New York, NY: Walter de Gruyter Verlag.

den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In Werner Abraham (ed.), *On the formal syntax of the westgermania: Papers from the 3rd groningen grammar talks, groningen, january 1981* (Linguistik Aktuell/ Linguistics Today 3), 47–131. Amsterdam/Philadelphia: John Benjamins Publishing Co. http://www.dbnl.org/tekst/best005inte01_01/.

den Besten, Hans. 1985. Some remarks on the Ergative Hypothesis. In Werner Abraham (ed.), *Erklärende Syntax des Deutschen* (Studien zur deutschen Grammatik 25), 53–74. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Deppermann, Arnulf. 2006. Construction Grammar – eine Grammatik für die Interaktion? In Arnulf Deppermann, Reinhard Fiehler & Thomas Spranz-Fogasy (eds.), *Grammatik und Interaktion*, 43–65. Radolfzell: Verlag für Gesprächsforschung.

Derbyshire, Desmond C. 1979. *Hixkaryana* (Lingua Descriptive Series 1). Amsterdam: North Holland.

Devlin, Keith. 1992. *Logic and information*. Cambridge: Cambridge University Press.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA/London, England: MIT Press.

Dione, Cheikh Mouhamadou Bamba. 2013. Handling Wolof Clitics in LFG. In Christine Meklenborg Salvesen & Hans Petter Helland (eds.), *Challenging clitics* (Linguistik Aktuell/Linguistics Today 206), 87–118. Amsterdam: John Benjamins Publishing Co.

Dione, Cheikh Mouhamadou Bamba. 2014. An LFG approach to Wolof cleft constructions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 157–176. Stanford, CA: CSLI Publications.

Dipper, Stefanie. 2003. *Implementing and documenting large-scale grammars – German LFG*: IMS, University of Stuttgart dissertation. http://www.ims.uni-stuttgart.de/projekte/pargram/paper/diss2003.pdf. Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (AIMS), Volume 9, Number 1.

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

Doran, Christine, Beth Ann Hockey, Anoop Sarkar, Bangalore Srinivas & Fei Xia. 2000. Evolution of the XTAG system. In Abeillé & Rambow (2000) 371–403.

Dörre, Jochen & Michael Dorna. 1993. CUF: A formalism for linguistic knowledge representation. DYANA 2 deliverable R.1.2A IMS Stuttgart, Germany. http://www.ims.uni-stuttgart.de/projekte/cuf/dyana2/R1.2.A/R1.2.A-Dorre1.ps.gz.

Dörre, Jochen & Roland Seiffert. 1991. A formalism for natural language — STUF. In Otthein Herzog & Claus-Rainer Rollinger (eds.), *Text understanding in lilog* (Lecture Notes in Artificial Intelligence 546), 29–38. Berlin/Heidelberg/New York, NY: Springer Verlag.

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

Dowty, David R. 1979. *Word meaning and montague grammar* (Synthese Language Library 7). Dordrecht/Boston/London: D. Reidel Publishing Company.

Dowty, David R. 1988. Type raising, functional composition, and nonconstituent coordination. In Richard Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial grammars and natural language structures*, 153–198. Dordrecht/Boston/London: D. Reidel Publishing Company.

Dowty, David R. 1989. On the semantic content of the notion 'thematic role'. In Gennaro Chierchia, Barbara H. Partee & Raymond Turner (eds.), *Properties, types and meaning*, vol. 2 (Studies in Linguistics and Philosophy), 69–130. Dordrecht/Boston/London: Kluwer Academic Publishers.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Dowty, David R. 2003. The dual analysis of adjuncts and complements in categorial grammar. In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.), *Modifying adjuncts* (Interface Explorations 4), Berlin: Mouton de Gruyter.

Dras, Mark, François Lareau, Benjamin Börschinger, Robert Dale, Yasaman Motazedi, Owen Rambow, Myfany Turpin & Morgan Ulinski. 2014. Complex predicates in Arrernte. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 177–197. Stanford, CA: CSLI Publications.

Drellishak, Scott. 2009. *Widespread but not universal: Improving the typological coverage of the grammar matrix*: University of Washington Doctoral dissertation.

Dryer, Matthew S. 1997. Are grammatical relations universal? In Joan Bybee, John Haiman & Sandra Thompson (eds.), *Essays on language function and language type: Dedicated to t. givon*, 115–143. John Benjamins Publishing Co.

Duden. 1966. *Grammatik der deutschen gegenwartssprache*, vol. 4. Mannheim, Wien, Zürich: Dudenverlag 2nd edn.

Duden. 1984. *Grammatik der deutschen gegenwartssprache*, vol. 4. Mannheim, Wien, Zürich: Dudenverlag 4th edn.

Duden. 1995. *Die grammatik*, vol. 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag 5th edn.

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

Dürscheid, Christa. 1989. *Zur Vorfeldbesetzung in deutschen Verbzweit-Strukturen* (FOKUS 1). Trier: Wissenschaftlicher Verlag.

Dyvik, Helge, Paul Meurer & Victoria Rosén. 2005. LFG, Minimal Recursion Semantics and translation. Paper preseneted at the LFG conference 2005.

Egg, Markus. 1999. Derivation and resolution of ambiguities in *wieder*-sentences. In Paul J. E. Dekker (ed.), *Proceedings of the 12th amsterdam colloquium*, 109–114.

Eisele, Andreas & Jochen Dorre. 1986. A Lexical Functional Grammar system in Prolog. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 551–553. University of Bonn: Association for Computational Linguistics.

Eisenberg, Peter. 1992. Platos Problem und die Lernbarkeit der Syntax. In Peter Suchsland (ed.), *Biologische und soziale Grundlagen der Sprache* (Linguistische Arbeiten 280), 371–378. Tübingen: Max Niemeyer Verlag.

Eisenberg, Peter. 1994a. German. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages* (Routledge Language Family Descriptions), 349–387. London / New York: Routledge.

Eisenberg, Peter. 1994b. *Grundriß der deutschen grammatik*. Stuttgart, Weimar: Verlag J. B. Metzler 3rd edn.

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

Ellefson, Michelle R. & Morten Christiansen. 2000. Subjacency constraints without universal grammar: Evidence from artificial language learning and connectionist modeling. In *Proceedings of the 22nd annual conference of the cognitive science society*, 645–650. Mahwah, NJ: Lawrence Erlbaum Associates. http://cnl.psych.cornell.edu/abstracts/EandC-cogsci2000.html.

Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48(1). 71–99. http://www3.isrl.uiuc.edu/~junwang4/langev/localcopy/pdf/elman93cognition.pdf.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: Bradford Books/MIT Press.

Embick, David. 2004. On the structure of resultative participles in english. *Linguistic Inquiry* 35(3). 355–392.

Emirkanian, Louisette, Lyne Da Sylva & Lorne H. Bouchard. 1996. The implementation of a computational grammar of French using the Grammar Development Environment. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 1024–1027. Copenhagen, Denmark: Association for Computational Linguistics.

Enard, Wolfgang, Molly Przeworski, Simon E. Fisher, Cecilia S. L. Lai, Victor Wiebe, Takashi Kitano, Anthony P. Monaco & Svante Pääbo. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418. 869–872.

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

Engel, Ulrich. 1970. Regeln zur Wortstellung. Forschungsberichte des Instituts für deutsche Sprache 5 Institut für deutsche Sprache Mannheim.

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

Engel, Ulrich. 1994. *Syntax der deutschen Gegenwartssprache*, vol. 22 (Grundlagen der Germanistik). Berlin: Erich Schmidt Verlag. 3., völlig überarbeitete Auflage.

Engel, Ulrich. 1996. Tesnière mißverstanden. In Gertrud Gréciano & Helmut Schumacher (eds.), *Lucien Tesnière – Syntaxe Structurale et Opèrations Mentales. Akten des deutsch-französischen Kolloquiums anläßlich der 100. Wiederkehr seines Geburtstages. Strasbourg 1993* (Linguistische Arbeiten 348), 53–61. Tübingen: Max Niemeyer Verlag.

Engel, Ulrich. 2014. Die dependenzielle Verbgrammatik (DVG). In Jörg Hagemann & Sven Staffeldt (eds.), *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28), 43–62. Tübingen: Stauffenburg.

Erbach, Gregor. 1995. ProFIT: Prolog with features, inheritance and templates. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the seventh conference of the European Chapter of the Association for Computational Linguistics*, 180–187. Dublin: Association for Computational Linguistics.

Eroms, Hans-Werner. 1985. Eine reine Dependenzgrammatik für das Deutsche. *Deutsche Sprache* 13. 306–326.

Eroms, Hans-Werner. 1987. Passiv und Passivfunktionen im Rahmen einer Dependenzgrammatik. In Centre de Recherche en Linguistique Germanique (Nice) (ed.), *Das passiv im deutschen* (Linguistische Arbeiten 183), 73–95. Tübingen: Max Niemeyer Verlag.

Eroms, Hans-Werner. 2000. *Syntax der deutschen sprache* (de Gruyter Studienbuch). Berlin/New York, NY: Walter de Gruyter Verlag.

Eroms, Hans-Werner & Hans Jürgen Heringer. 2003. Dependenz und lineare Ordnung. In Ágel et al. (2003) 247–263.

Eroms, Hans-Werner, Gerhard Stickel & Gisela Zifonun (eds.). 1997. *Grammatik der deutschen sprache*, vol. 7 (Schriften des Instituts für deutsche Sprache). Berlin/New York, NY: Walter de Gruyter. http://www.ids-mannheim.de/gra/grammar/.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.

Erteschik-Shir, Nomi. 1981. More on extractability from quasi-nps. *Linguistic Inquiry* 12(4). 665–670.

Erteschik-Shir, Nomi & Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6(1–3). 41–86.

Estigarribia, Bruno. 2009. Facilitation by variation: Right-to-left learning of English yes/no questions. *Cognitive Science* 34(1). 68–93.

Evans, Nicholas & Stephen C. Levinson. 2009a. The myth of language universals: Language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences* 32(5). 429–448.

Evans, Nicholas & Stephen C. Levinson. 2009b. With diversity in mind: Freeing the language sciences from Universal Grammar. *The Behavioral and Brain Sciences* 32(5). 472–492.

Evans, Roger. 1985. ProGram – a development tool for GPSG grammars. *Linguistics* 23(2). 213–244.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46(4). 621–646. http://www.eva.mpg.de/psycho/pdf/Publications_2005_PDF/Commentary_on_D.Everett_05.pdf.

Everett, Daniel L. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85(2). 405–442.

Evers, Arnold. 1975. *The transformational cycle in Dutch and German*: University of Utrecht dissertation.

Faaß, Gertrud. 2010. *A morphosyntactic description of Northern Sotho as a basis for an automated translation from Northern Sotho into English*. Pretoria, South Africa: University of Pretoria dissertation. http://hdl.handle.net/2263/28569.

Falk, Yehuda N. 1984. The English auxiliary system: A lexical-functional analysis. *Language* 60(3). 483–509.

Fang, Ji & Tracy Holloway King. 2007. An LFG Chinese grammar for machine use. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE), 144–160. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/.

Fanselow, Gisbert. 1981. *Zur Syntax und Semantik der Nominalkomposition* (Linguistische Arbeiten 107). Tübingen: Max Niemeyer Verlag.

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

Fanselow, Gisbert. 1988. Aufspaltung von NPn und das Problem der ‚freien' Wortstellung. *Linguistische Berichte* 114. 91–113.

Fanselow, Gisbert. 1990. Scrambling as NP-movement. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers* (Linguistik Aktuell/Linguistics Today 5), 113–140. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Fanselow, Gisbert. 1992a. ‚Ergative' Verben und die Struktur des deutschen Mittelfelds. In Ludger Hoffmann (ed.), *Deutsche syntax – ansichten und aussichten* (Institut für deutsche Sprache, Jahrbuch 1991), 276–303. Berlin/New York, NY: de Gruyter.

Fanselow, Gisbert. 1992b. Zur biologischen Autonomie der Grammatik. In Peter Suchsland (ed.), *Biologische und soziale Grundlagen der Sprache* (Linguistische Arbeiten 280), 335–356. Tübingen: Max Niemeyer Verlag.

Fanselow, Gisbert. 1993. Die Rückkehr der Basisgenerierer. *Groninger Arbeiten zur Germanistischen Linguistik* 36. 1–74.

Fanselow, Gisbert. 2000a. Does constituent length predict German word order in the Middle Field? In Josef Bayer & Christine Römer (eds.), *Von der Philologie zur Grammatiktheorie: Peter Suchsland zum 65. Geburtstag*, 63–77. Tübingen: Max Niemeyer Verlag.

Fanselow, Gisbert. 2000b. Optimal exceptions. In Barbara Stiebels & Dieter Wunderlich (eds.), *The lexicon in focus* (studia grammatica 45), 173–209. Berlin: Akademie Verlag.

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

Fanselow, Gisbert. 2002. Against remnant VP-movement. In Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner (eds.), *Dimensions of move-*

ment: From features to remnants (Linguistik Aktuell/Linguistics Today 48), 91–127. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Fanselow, Gisbert. 2003a. Free constituent order: A Minimalist interface account. *Folia Linguistica* 37(1–2). 191–231.

Fanselow, Gisbert. 2003b. Münchhausen-style head movement and the analysis of verb second. In Anoop Mahajan (ed.), *Proceedings of the workshop on head movement* (UCLA Working Papers in Linguistics 10), Los Angeles: UCLA, Linguistics Department.

Fanselow, Gisbert. 2003c. Zur Generierung der Abfolge der Satzglieder im Deutschen. *Neue Beiträge zur Germanistik* 112. 3–47.

Fanselow, Gisbert. 2004a. Cyclic phonology-syntax-interaction: PPT Movement in German (and other languages). In Shinichiro Ishihara, Michaela Schmitz & Anne Schwarz (eds.), *Interdisciplinary studies on information structure* (Working papers of the SFB 632 1), Potsdam: Universitätsverlag.

Fanselow, Gisbert. 2004b. Fakten, Fakten, Fakten! *Linguistische Berichte* 200. 481–493.

Fanselow, Gisbert. 2006. On pure syntax (uncontaminated by information structure). In Patrick Brandt & Eric Fuss (eds.), *Form, structure and grammar: A festschrift presented to Günther Grewendorf on occasion of his 60th birthday* (Studia grammatica 63), 137–157. Berlin: Akademie Verlag.

Fanselow, Gisbert. 2009. Die (generative) Syntax in den Zeiten der Empiriediskussion. *Zeitschrift für Sprachwissenschaft* 28(1). 133–139.

Fanselow, Gisbert & Sascha W. Felix. 1987. *Sprachtheorie 2. Die Rektions- und Bindungstheorie* (UTB für Wissenschaft: Uni-Taschenbücher 1442). Tübingen: A. Francke Verlag GmbH.

Fanselow, Gisbert, Matthias Schlesewsky, Damir Cavar & Reinhold Kliegl. 1999. Optimal parsing, syntactic parsing preferences, and optimality theory. Rutgers Optimality Archive (ROA) 367 Universität Potsdam. http://roa.rutgers.edu/view.php3?roa=367.

Feldhaus, Anke. 1997. Eine HPSG-Analyse ausgewählter Phänomene des deutschen *w*-Fragesatzes. Working Papers of the Institute for Logic and Linguistics 27 Institute for Logic and Linguistics IBM Scientific Center Heidelberg.

Feldman, Jerome. 1972. Some decidability results on grammatical inference and complexity. *Information and Control* 20(3). 244–262.

Fillmore, Charles J. 1968. The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals of linguistic theory*, 1–88. New York: Holt, Rinehart, and Winston.

Fillmore, Charles J. 1971. Plädoyer für Kasus. In Werner Abraham (ed.), *Kasustheorie* (Schwerpunkte Linguistik und Kommunikationswissenschaft 2), 1–118. Frankfurt/Main: Athenäum.

Fillmore, Charles J. 1988. The mechanisms of "construction grammar". In Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds.), *Proceedings of the 14th annual meeting of the berkeley linguistics society*, 35–55. Berkeley, CA: Berkeley Linguistics Society.

Fillmore, Charles J. 1999. Inversion and constructional inheritance. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 113–128. Stanford, CA: CSLI Publications.

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.

Fillmore, Charles J., Russell R. Lee-Goldmann & Russell Rhomieux. 2012. The FrameNet constructicon. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar* (CSLI Lecture Notes 193), 309–372. Stanford, CA: CSLI Publications.

Fischer, Ingrid & Martina Keil. 1996. Parsing decomposable idioms. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 388–393. Copenhagen, Denmark: Association for Computational Linguistics.

Fischer, Kerstin & Anatol Stefanowitsch (eds.). 2006. *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40). Tübingen: Stauffenburg Verlag.

Fisher, Simon E. & Gary F. Marcus. 2005. The eloquent ape: Genes, brains and the evolution of language. *Nature Reviews Genetics* 7(1). 9–20.

Fisher, Simon E., Faraneh Vargha-Khadem, Kate E. Watkins, Anthony P. Monaco & Marcus E. Pembrey. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics* 18(2). 168–170.

Fitch, W. Tecumseh. 2010. Three meanings of "recursion": Key distinctions for biolinguistics. In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The evolution of human language: Biolinguistic perspectives* (Approaches to the Evolution of Language 2), 73–90. Cambridge, UK: Cambridge University Press.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97(2). 179–210.

Flickinger, Daniel P. 1983. Lexical heads and phrasal gaps. In *Proceedings of the west coast conference on formal linguistics*, vol. 2, Stanford University Linguistics Dept.

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

Flickinger, Daniel P. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(1). 15–28.

Flickinger, Daniel P. 2008. Transparent heads. In Stefan Müller (ed.), *Proceedings of the 15th international conference on Head-Driven Phrase Structure Grammar*, 87–94. Stanford, CA: CSLI Publications. .

Flickinger, Daniel P. & Emily M. Bender. 2003. Compositional semantics in a multilingual grammar resource. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the esslli 2003 workshop "ideas and strategies for multilingual grammar development"*, 33–42. Vienna, Austria.

Flickinger, Daniel P., Ann Copestake & Ivan A. Sag. 2000. HPSG analysis of English. In Wahlster (2000) 254–263.

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

Fodor, Janet Dean. 1998a. Parsing to learn. *Journal of Psycholinguistic Research* 27(3). 339–374.

Fodor, Janet Dean. 1998b. Unambiguous triggers. *Linguistic Inquiry* 29(1). 1–36.

Fodor, Janet Dean. 2001. Parameters and the periphery: Reflections on *syntactic nuts*. *Journal of Linguistics* 37. 367–392.

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

Fokkens, Antske. 2011. Metagrammar engineering: Towards systematic exploration of implemented grammars. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, 1066–1076. Portland, Oregon, USA: Association for Computational Linguistics. http://www.aclweb.org/anthology/P11-1107.

Fokkens, Antske, Laurie Poulson & Emily M. Bender. 2009. Inflectional morphology in Turkish VP coordination. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 110–130. Stanford, CA: CSLI Publications.

Fong, Sandiway. 1991. *Computational properties of principle-based grammatical theories*: MIT Artificial Intelligence Lab dissertation. http://www.neci.nec.com/homepages/sandiway/pappi/index.html.

Fong, Sandiway. 2014. Unification and efficient computation in the minimalist program. In L. Francis & L. Laurent (eds.), *Language and recursion*, 129–138. Berlin/Heidelberg/New York, NY: Springer Verlag.

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

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

Forst, Martin. 2006. COMP in (parallel) grammar writing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2006 conference*, Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/11/lfg06forst.pdf.

Forst, Martin & Christian Rohrer. 2009. Problems of German VP coordination. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2009 conference*, 297–316. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/14/.

Fortmann, Christian. 1996. *Konstituentenbewegung in der DP-Struktur: zur funktionalen Analyse der Nominalphrase im Deutschen* (Linguistische Arbeiten 347). Tübingen: Max Niemeyer Verlag.

Fourquet, Jean. 1957. Review of: Heinz Anstock: Deutsche Syntax – Lehr- und Übungsbuch. *Wirkendes Wort* 8. 120–122.

Fourquet, Jean. 1970. *Prolegomena zu einer deutschen Grammatik* (Sprache der Gegenwart – Schriften des Instituts für deutsche Sprache in Mannheim 7). Düsseldorf: Pädagogischer Verlag Schwann.

Fouvry, Frederik. 2003. Lexicon acquisition with a large-coverage unification-based grammar. In *Proceedings of eacl 03, 10th conference of the european chapter of the association for computational linguistics, research notes and demos, april 12–17, 2003, budapest, hungary*, 87–90.

Fraj, Fériel Ben, Chiraz Zribi & Mohamed Ben Ahmed. 2008. ArabTAG: A tree adjoining grammar for Arabic syntactic structures. In *Proceedings of the international arab conference on information technology*, Sfax, Tunisia.

Frank, Anette. 1994. Verb second by lexical rule or by underspecification. Arbeitspapiere des SFB 340 No. 43 IBM Deutschland GmbH Heidelberg. ftp://ftp.ims.uni-stuttgart.de/pub/papers/anette/v2-usp.ps.gz.

Frank, Anette. 1996. Eine LFG-Grammatik des Französischen. In *Deutsche und französische Syntax im Formalismus der LFG* (Linguistische Arbeiten 344), 97–244. Tübingen: Max Niemeyer Verlag.

Frank, Anette. 2006. (discourse-) functional analysis of asymmetric coordination. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 259–285. Stanford, CA: CSLI Publications.

Frank, Anette & Uwe Reyle. 1995. Principle based semantics for HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the seventh conference of the European Chapter of the Association for Computational Linguistics*, 9–16. Dublin: Association for Computational Linguistics.

Frank, Anette & Annie Zaenen. 2002. Tense in LFG: Syntax and morphology. In Hand Kamp & Uwe Reyle (eds.), *How we say when it happens. contributions to the theory of temporal reference in natural language*, Tübingen: Max Niemeyer Verlag. Reprint as: Frank & Zaenen (2004).

Frank, Anette & Annie Zaenen. 2004. Tense in LFG: Syntax and morphology. In Louisa Sadler & Andrew Spencer (eds.), *Projecting morphology*, Stanford, CA: CSLI Publications.

Frank, Robert. 2002. *Phrase structure composition and syntactic dependencies* (Current Studies in Linguistics 38). Cambridge, MA/London: MIT Press.

Franks, Steven. 1995. *Parameters in Slavic morphosyntax*. New York, Oxford: Oxford University Press.

Frazier, Lyn. 1985. Syntactic complexity. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, 129–189. Cambridge, UK: Cambridge University Press.

Frazier, Lyn & Charles Clifton. 1996. *Construal*. Cambridge, MA: MIT Press.

Freidin, Robert. 1975. The analysis of passives. *Language* 51(2). 384–405.

Freidin, Robert. 1997. Review article: The minimalist program. *Language* 73(3). 571–582.

Freidin, Robert. 2009. A note on methodology in linguistics. *The Behavioral and Brain Sciences* 32(5). 454–455.

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.

Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2006. Modeling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science* 30(2). 277–310.

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.

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

Frey, Werner. 2000. Über die syntaktische Position der Satztopiks im Deutschen. In Ewald Lang, Marzena Rochon, Kerstin Schwabe & Oliver Teuber (eds.), *Issues on topics* (ZAS Papers in Linguistics 20), 137–172. ZAS, Humboldt-Universität zu Berlin. http://www.zas.gwz-berlin.de/mitarb/homepage/frey/files/1130927092.pdf.

Frey, Werner. 2001. About the whereabouts of indefinites. *Theoretical Linguistics* 27(2/3). Special Issue: NP Interpretation and Information Structure, Edited by Klaus von Heusinger and Kerstin Schwabe.

Frey, Werner. 2004a. The grammar-pragmatics interface and the German prefield. Forschungsprogramm Sprache und Pragmatik 52 Germanistisches Institut der Universität Lund.

Frey, Werner. 2004b. A medial topic position for German. *Linguistische Berichte* 198. 153–190.

Frey, Werner. 2005. Pragmatic properties of certain German and English left peripheral constructions. *Linguistics* 43(1). 89–129.

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

Frey, Werner & Uwe Reyle. 1983a. Lexical Functional Grammar und Diskursrepräsentationstheorie als Grundlagen eines sprachverarbeitenden Systems. *Linguistische Berichte* 88. 79–100.

Frey, Werner & Uwe Reyle. 1983b. A Prolog implementation of Lexical Functional Grammar as a base for a natural language processing system. In Antonio Zampoili (ed.), *Proceedings of the first conference on european chapter of the association for computational linguistics*, 52–57. Pisa, Italy: Association for Computational Linguistics. http://aclweb.org/anthology/E/E83/.

Fried, Mirjam. 2015. Construction grammar. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science), Berlin/New York, NY: Mouton de Gruyter 2nd edn.

Friederici, Angela D. 2009. Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences* 13(4). 175–181.

Friedman, Joyce. 1969. Applications of a computer system for Transformational Grammar. In Research Group for Quantitative Linguistics (ed.), *Proceedings of COLING 69*, 1–27.

Friedman, Joyce, Thomas H. Bredt, Robert W. Doran, Bary W. Pollack & Theodore S. Martner. 1971. *A computer model of transformational grammar* (Mathematical Linguistics and Automatic Language Processing 9). New York: Elsevier.

Fries, Norbert. 1988. Über das Null-Topik im Deutschen. Forschungsprogramm Sprache und Pragmatik 3 Germanistisches Institut der Universität Lund Lund.

Fukui, Naoki & Margaret Speas. 1986. Specifiers and projection. In N. Fukui, T. R. Rapoport & E. Sagey (eds.), *Papers in theoretical linguistics* (MIT Working Papers 8), 128–172. Cambridge, MA: MIT.

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

Gardner, R. Allen. 1957. Probability-learning with two and three choices. *The American Journal of Psychology* 70(2). 174–185.

Gärtner, Hans-Martin & Jens Michaelis. 2007. Some remarks on locality conditions and minimalist grammars. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? chomsky's minimalism and the view from syntax-semantics* (Studies in Generative Grammar 89), 161–195. Berlin/New York, NY: Mouton de Gruyter.

Gärtner, Hans-Martin & Markus Steinbach. 1997. Anmerkungen zur Vorfeldphobie pronominaler Elemente. In Franz-Josef d'Avis & Uli Lutz (eds.), *Zur satzstruktur im deutschen* (Arbeitspapiere des SFB 340 No. 90), 1–30. Tübingen: Eberhard-Karls-Universität Tübingen.

Gazdar, Gerald. 1981a. On syntactic categories. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 295(1077). 267–283.

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

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gazdar, Gerald, Geoffrey K. Pullum, Bob Carpenter, Ewan Klein, Thomas E. Hukari & Robert D. Levine. 1988. Category structures. *Computational Linguistics* 14(1). 1–19.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

Geißler, Stefan & Tibor Kiss. 1994. Erläuterungen zur Umsetzung einer HPSG im Basisformalismus STUF III. Tech. rep. IBM Informationssysteme GmbH – Institut fuer Logik und Linguistik (VerbundvorhabenVerbmobil) Heidelberg.

Gerdes, Kim. 2002a. DTAG? In *Proceedings of the sixth international workshop on tree adjoining grammar and related frameworks (tag+6)*, 242–251. Universitá di Venezia.

Gerdes, Kim. 2002b. *Topologie et grammaires formelles de l'allemand*: Ecole doctorale Science du langage, UFR de linguistique, Université Paris 7 dissertation.

Gerdes, Kim & Sylvain Kahane. 2001. Word order in german: A formal dependency grammar using a topological hierarchy. In *Proceedings of the 39th annual meeting on association for computational linguistics*, 220–227. Stroudsburg, PA, USA: Association for Computational Linguistics. doi:10.3115/1073012.1073041. http://dx.doi.org/10.3115/

1073012.1073041.

Gerken, LouAnn. 1991. The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30. 431–451.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

Gibson, Edward & James Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14(3). 225–248.

Gibson, Edward & Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25(3). 407–454.

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

Gold, Mark E. 1967. Language identification in the limit. *Information and Control* 10(5). 447–474.

Goldberg, Adele E. 1995. *Constructions. a construction grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago/London: The University of Chicago Press.

Goldberg, Adele E. 2003a. Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.

Goldberg, Adele E. 2003b. Words by default: the Persian complex predicate construction. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (CSLI Lecture Notes 163), 117–146. Stanford, CA: CSLI Publications.

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

Goldberg, Adele E. 2009. Constructions work. [response]. *Cognitive Linguistics* 20(1). 201–224.

Goldberg, Adele E. 2013a. Argument structure constructions vs. lexical rules or derivational verb templates. *Mind and Language* 28(4). 435–465.

Goldberg, Adele E. 2013b. Explanation and constructions: Response to Adger. *Mind and Language* 28(4).

Goldberg, Adele E. 2014. Fitting a slim dime between the verb template and argument structure construction approaches. *Theoretical Linguistics* 40(1–2). To appear.

Goldberg, Adele E. & Farrell Ackerman. 2001. The pragmatics of obligatory adjuncts. *Language* 77(4). 798–814.

Goldberg, Adele E., Devin Casenhiser & Nitya Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15(3). 289–316.

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

Gopnik, Myrna & Martha B. Cargo. 1991. Familial aggregation of a developmental language disorder. *Cognition* 39(1). 1–50.

Gordon, Peter. 1986. Level ordering in lexical development. *Cognition* 21(2). 73–93.

Gosch, Angela, Gabriele Städing & Rainer Pankau. 1994. Linguistic abilities in children with Williams-Beuren Syndrome. *American Journal of Medical Genetics* 52(3). 291–296.

Götz, Thilo, Walt Detmar Meurers & Dale Gerdemann. 1997. The ConTroll manual. (ConTroll v.1.0 beta, XTroll v.5.0 beta). User's manual Seminar für Sprachwissenschaft Universität Tübingen. http://www.sfs.uni-tuebingen.de/controll/code.html.

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

Grewendorf, Günther. 1983. Reflexivierungen in deutschen A.c.I.-Konstruktionen – kein transformationsgrammatisches Dilemma mehr. *Groninger Arbeiten zur Germanistischen Linguistik* 23. 120–196.

Grewendorf, Günther. 1985. Anaphern bei Objekt-Koreferenz im Deutschen. Ein Problem für die Rektions-Bindungs-Theorie. In Werner Abraham (ed.), *Erklärende Syntax des Deutschen* (Studien zur deutschen Grammatik 25), 137–171. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Grewendorf, Günther. 1987. Kohärenz und Restrukturierung. Zu verbalen Komplexen im Deutschen. In Brigitte Asbach-Schnitker & Johannes Roggenhofer (eds.), *Neuere forschungen zur wortbildung und histographie. festgabe für herbert e. brekle zum 50. geburtstag* (Tübinger Beiträge zur Linguistik 284), 123–144. Tübingen: Gunter Narr Verlag.

Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse* (Studien zur deutschen Grammatik 33). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Grewendorf, Günther. 1989. *Ergativity in German* (Studies in Generative Grammar 35). Dordrecht: Holland, Providence: U.S.A.: Foris Publications.

Grewendorf, Günther. 1993. German. A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax – ein internationales handbuch zeitgenössischer forschung*, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin/New York, NY: Walter de Gruyter Verlag.

Grewendorf, Günther. 2002. *Minimalistische Syntax* (UTB für Wissenschaft: Uni-Taschenbücher 2313). Tübingen, Basel: A. Francke Verlag GmbH.

Grewendorf, Güther. 2009. The left clausal periphery. Clitic left dislocation in Italian and left dislocation in German. In Benjamin Shear, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated elements in discourse. syntactic, semantic, and pragmatic perspectives* (Routledge Studies in Germanic Linguistics), 49–94. New York: Routledge.

Grimshaw, Jane. 1986. Subjacency and the s/s′ parameter. *Linguistic Inquiry* 17(2). 364–369.

Grimshaw, Jane. 1997. Projections, heads, and optimality. *Linguistic Inquiry* 28. 373–422.

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

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

Groß, Thomas Michael. 2003. Dependency grammar's limits – and ways of extending them. In Ágel et al. (2003) 331–351.

Grosu, Alexander. 1973. On the status of the so-called right roof constraint. *Language* 49(2). 294–311.

Grover, Claire, John Carroll & Ted J. Briscoe. 1993. The Alvey Natural Language Tools grammar (4th release). Technical Report 284 Computer Laboratory, Cambridge University, UK.

Grubačić, Emilija. 1965. *Untersuchungen zur Frage der Wortstellung in der deutschen Prosadichtung der letzten Jahrzehnte*. Zagreb: Philosophische Fakultät dissertation.

Gruber, Jeffrey. 1965. *Studies in lexical relations*: MIT dissertation.

Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the second international workshop on Japanese syntax*, 1–21. Stanford, CA: CSLI Publications.

Günther, Carsten, Claudia Maienborn & Andrea Schopp. 1999. The processing of information structure. In Peter Bosch & Rob van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives* (Studies in Natural Language Processing), 18–42. Cambridge, UK: Cambridge University Press. Rev. papers orig. presented at a conference held 1994, Schloss Wolfsbrunnen, Germany.

Guo, Yuqing, Haifeng Wang & Josef van Genabith. 2007. Recovering non-local dependencies for chinese. In *Proceedings of the joint conference on empirical methods in natural language processing and natural language learning, (emnlp-conll 2007)*, 257–266. Prague, Czech Republic: Association for Computational Linguistics.

Haddar, Kais, Sirine Boukedi & Ines Zalila. 2010. Construction of an HPSG grammar for the Arabic language and its specification in TDL. *International Journal on Information and Communication Technologies, Vol. 3, No. 3, June 2010* 3(3). 52–64.

Haegeman, Liliane. 1990. Understood subjects in English diaries. *Multilingua* 9. 157–199.

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

Haegeman, Liliane. 1995. *The syntax of negation*. Cambridge, UK: Cambridge University Press.

Haftka, Brigitta. 1995. Syntactic positions for topic and contrastive focus in the German middlefield. In Inga Kohlhof, Susanne Winkler & Hans-Bernhard Drubig (eds.), *Proceedings of the göttingen focus workshop, 17 DGfS, march 1–3* (Arbeitspapiere des SFB 340 No. 69), 137–157. Eberhard-Karls-Universität Tübingen.

Haftka, Brigitta. 1996. Deutsch ist eine V/2-Sprache mit Verbendstellung und freier Wortfolge. In Ewald Lang & Gisela Zifonun (eds.), *Deutsch – typologisch* (Institut für deutsche Sprache, Jahrbuch 1995), 121–141. Berlin/New York, NY: Walter de Gruyter.

Hahn, Michael. 2011. Null conjuncts and bound pronouns in Arabic. In Stefan Müller (ed.), *Proceedings of the 18th international conference on Head-Driven Phrase Structure Grammar, university of washington*, 60–80. Stanford, CA: CSLI Publications.

Haider, Hubert. 1982. Dependenzen und Konfigurationen: Zur deutschen V-Projektion. *Groninger Arbeiten zur Germanistischen Linguistik* 21. 1–60.

Haider, Hubert. 1984. Was zu haben ist und was zu sein hat – Bemerkungen zum Infinitiv. *Papiere zur Linguistik* 30(1). 23–36.

Haider, Hubert. 1985a. The case of German. In Jindřich Toman (ed.), *Studies in german grammar* (Studies in Generative Grammar 21), 23–64. Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

Haider, Hubert. 1985b. Über *sein* oder nicht *sein*: Zur Grammatik des Pronomens *sich*. In Werner Abraham (ed.), *Erklärende Syntax des Deutschen* (Studien zur deutschen Grammatik 25), 223–254. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Haider, Hubert. 1986a. Fehlende Argumente: vom Passiv zu kohärenten Infinitiven. *Linguistische Berichte* 101. 3–33.

Haider, Hubert. 1986b. Nicht-sententiale Infinitive. *Groninger Arbeiten zur Germanistischen Linguistik* 28. 73–114.

Haider, Hubert. 1990a. Pro-bleme? In Gisbert Fanselow & Sascha W. Felix (eds.), *Strukturen und merkmale syntaktischer kategorien* (Studien zur deutschen Grammatik 39), 121–143. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Haider, Hubert. 1990b. Topicalization and other puzzles of German syntax. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers* (Linguistik Aktuell/ Linguistics Today 5), 93–112. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Haider, Hubert. 1991. Fakultativ kohärente Infinitivkonstruktionen im Deutschen. Arbeitspapiere des SFB 340 No. 17 IBM Deutschland GmbH Heidelberg.

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

Haider, Hubert. 1994. (Un-)heimliche Subjekte – Anmerkungen zur Pro-drop Causa, im Anschluß an die Lektüre von Osvaldo Jaeggli & Kenneth J. Safir, eds., The Null Subject Parameter. *Linguistische Berichte* 153. 372–385.

Haider, Hubert. 1995. Studies on phrase structure and economy. Arbeitspapiere des SFB 340 No. 70 Universität Stuttgart Stuttgart.

Haider, Hubert. 1997a. Projective economy. On the minimal functional structure of the German clause. In Werner Abraham & Elly van Gelderen (eds.), *German: Syntactic problems—problematic syntax* (Linguistische Arbeiten 374), 83–103. Tübingen: Max Niemeyer Verlag.

Haider, Hubert. 1997b. Typological implications of a directionality constraint on projections. In Artemis Alexiadou & T. Alan Hall (eds.), *Studies on universal grammar and typological variation* (Linguistik Aktuell/Linguistics Today 13), 17–33. Amsterdam/ Philadelphia: John Benjamins Publishing Co.

Haider, Hubert. 1999. The license to license: Structural case plus economy yields Burzio's Generalization. In Eric Reuland (ed.), *Arguments and case: Explaining Burzio's Generalization* (Linguistik Aktuell/Linguistics Today 34), 31–55. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Haider, Hubert. 2000. OV is more basic than VO. In Peter Svenonius (ed.), *The derivation of vo and ov*, 45–67. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Haider, Hubert. 2001. Parametrisierung in der Generativen Grammatik. In Martin Haspelmath, Eckehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Sprachtypologie und sprachliche universalien – language typology and language universals. ein internationales handbuch – an international handbook*, 283–294. Berlin: Mouton de Gruyter.

Haider, Hubert. 2014. Scientific ideology and unscientific conduct – metaphors one lives by in the minimalist program. Ms. Universität Salzburg — Dept. of Linguistics and Centre for Cognitive Neuroscience.

Hajičová, Eva & Petr Sgall. 2003. Dependency syntax in functional generative description. In Ágel et al. (2003) 570–592.

Hakuta, Kenji, Ellen Bialystok & Edward Wiley. 2003. Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science* 14(1). 31–38.

Hale, Kenneth. 1976. The adjoined relative clause in Australia. In R.M.W. Dixon (ed.), *Grammatical catgeories of australian languages* (Linguistic Series 22), 78–105. New Jersey: Humanities Press.

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

Hale, Kenneth & Samuel Jay Keyser. 1997. On the complex nature of simple predicators. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates* (CSLI Lecture Notes 64), 29–65. Stanford, CA: CSLI Publications.

Han, Chung-hye, Juntae Yoon, Nari Kim & Martha Palmer. 2000. A feature-based Lexicalized Tree Adjoining Grammar for Korean. Technical Report IRCS-00-04 University of Pennsylvania Institute for Research in Cognitive Science. http://repository.upenn.edu/ircs_reports/35/.

Harbour, Daniel. 2011. Mythomania? methods and morals from 'the myth of language universals'. *Lingua* 121(12). 1820–1830.

Harley, Heidi & Rolf Noyer. 2000. Formal versus encyclopedic properties of vocabulary: Evidence from nominalizations. In Bert Peeters (ed.), *The lexicon–encyclopedia interface*, 349–374. Amsterdam: Elsevier.

Harman, Gilbert. 1963. Generative grammars without transformation rules: A defence of phrase structure. *Language* 39. 597–616.

Harris, Zellig S. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33(3). 283–340.

Haspelmath, Martin. 2008. Parametric versus functional explanations of syntactic universals. In T. Biberauer (ed.), *The limits of syntactic variation*, 75–107. Amsterdam/ Philadelphia: John Benjamins Publishing Co.

Haspelmath, Martin. 2009. The best-supported language universals refer to scalar patterns deriving from processing costs. *The Behavioral and Brain Sciences* 32(5). 457–458.

Haspelmath, Martin. 2010a. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.

Haspelmath, Martin. 2010b. The interplay between comparative concepts and descriptive categories (reply to Newmeyer). *Language* 86(3). 696–699.

Haugereid, Petter. 2007. Decomposed phrasal constructions. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/.

Haugereid, Petter. 2009. *Phrasal subconstructions: A constructionalist grammar design, exemplified with Norwegian and English*: Norwegian University of Science and Technology dissertation.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579. doi:10.1126/science.298.5598.1569. http://www.chomsky.info/articles/20021122.pdf.

Hausser, Roland. 1992. Complexity in left-associative grammar. *Theoretical Computer Science* 106(2). 283–308.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–285.

Hawkins, John A. 2004. *Efficiency and complexity in grammars.* Oxford: Oxford University Press.

Hays, David G. & T. W. Ziehe. 1960. *Studies in machine translation: 10–russian sentence-structure determination.* Rand Corporation.

Heinz, Wolfgang & Johannes Matiasek. 1994. Argument structure and case assignment in German. In Nerbonne et al. (1994) 199–236.

Helbig, Gerhard & Joachim Buscha. 1969. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht.* Leipzig: VEB Verlag Enzyklopädie.

Helbig, Gerhard & Joachim Buscha. 1998. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht.* Leipzig Berlin München: Langenscheidt Verlag Enzyklopädie 18th edn.

Helbig, Gerhard & Wolfgang Schenkel. 1969. *Wörterbuch zur Valenz und Distribution deutscher Verben.* Leipzig: VEB Bibliographisches Institut Leipzig.

Helbig, Gerhard & Wolfgang Schenkel. 1973. *Wörterbuch zur Valenz und Distribution deutscher Verben.* Leipzig: VEB Bibliographisches Institut Leipzig 2nd edn.

Hellan, Lars. 1986. The headedness of NPs in Norwegian. In Peter Muysken & Henk van Riemsdijk (eds.), *Features and projections*, 89–122. Dordrecht/Cinnaminson, U.S.A.: Foris Publications.

Hellan, Lars. 2007. On 'deep evaluation' for individual computational grammars and for cross-framework comparison. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE), 161–181. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/.

Hellan, Lars & Dorothee Beermann. 2006. The 'specifier' in an HPSG grammar implementation of Norwegian. In S. Werner (ed.), *Proceedings of the 15th nodalida conference, joensuu 2005* (Ling@JoY: University of Joensuu electronic publications in linguistics and language technology 1), 57–64. Joensuu: University of Joensuu.

Hellan, Lars & Petter Haugereid. 2003. Norsource – an excercise in the Matrix Grammar building design. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the esslli 2003 workshop "ideas and strategies for multilingual grammar development"*, Vienna, Austria. http://www.hf.ntnu.no/hf/isk/Ansatte/petter.haugereid/homepage/NorsourceMultilinggramm1.pdf.

Hellwig, Peter. 1978. PLAIN – Ein Programmsystem zur Sprachbeschreibung und maschinellen Sprachbearbeitung. *Sprache und Datenverarbeitung* 1(2).

Hellwig, Peter. 1986. Dependency unification grammar. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 195–198. University of Bonn: Association for Computational Linguistics. http://acl.eldoc.ub.rug.nl/mirror/C/C86/C86-1046.pdf.

Hellwig, Peter. 2003. Dependency unification grammar. In Ágel et al. (2003) 593–635.

Hellwig, Peter. 2006. Parsing with dependency grammars. In Ágel et al. (2006) 1081–1109.

Her, One-Soon, Dan Higinbotham & Joseph Pentheroudakis. 1991. An LFG-based machine translation system. *Computer Processing of Chinese and Oriental Languages* 5(3–4). 285–297. http://nccur.lib.nccu.edu.tw/bitstream/140.119/2105/1/J05.pdf.

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

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4). 547–593.

Hinrichs, Erhard W., Walt Detmar Meurers, Frank Richter, Manfred Sailer & Heike Winhart (eds.). 1997. *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, vol. No. 95 (Arbeitspapiere des SFB 340). Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989a. Flipped out: AUX in German. In *Aspects of German VP structure* (SfS-Report-01-93), Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989b. Subcategorization and VP structure in German. In *Aspects of German VP structure* (SfS-Report-01-93), Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In Nerbonne et al. (1994) 11–38.

Hinterhölzel, Roland. 2004. Language change versus grammar change: What diachronic data reveal about the interaction between core grammar and periphery. In Carola Trips & Eric Fuß (eds.), *Diachronic clues to synchronic grammar*, 131–160. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache* (Heutiges Deutsch. Linguistische Grundlagen. Forschungen des Instituts für deutsche Sprache 10). München: Max Hueber Verlag.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203. 88–96.

Hoeksema, Jack. 1991. Theoretische Aspekten van Partikelvooropplaatsing. *TABU Bulletin voor Taalwetenschap* 21(1). 18–26.

Hoffman, Beryl Ann. 1995. *The computational analysis of the syntax and interpretation of "free" word order in Turkish*: University of Pennsylvania dissertation.

Höhle, Tilman N. 1978. *Lexikalische Syntax: Die Aktiv-Passiv-Relation und andere Infinitkonstruktionen im Deutschen* (Linguistische Arbeiten 67). Tübingen: Max Niemeyer Verlag.

Höhle, Tilman N. 1982. Explikation für „normale Betonung“ und „normale Wortstellung“. In Werner Abraham (ed.), *Satzglieder im deutschen – vorschläge zur syntaktischen, semantischen und pragmatischen fundierung* (Studien zur deutschen Grammatik 15), 75–153. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Höhle, Tilman N. 1983. Topologische Felder. Köln, ms.

Höhle, Tilman N. 1986a. Der Begriff ‚Mittelfeld‘, Anmerkungen über die Theorie der topologischen Felder. In Walter Weiss, Herbert Ernst Wiegand & Marga Reis (eds.), *Akten des VII. Kongresses der Internationalen Vereinigung für germanische Sprach-und Literaturwissenschaft. Göttingen 1985. Band 3. Textlinguistik contra Stilistik?– Wortschatz und Wörterbuch – Grammatische oder pragmatische Organisation von Rede?* (Kontroversen, alte und neue 4), 329–340. Tübingen: Max Niemeyer Verlag.

Höhle, Tilman N. 1986b. Der Begriff „Mittelfeld“, Anmerkungen über die Theorie der topologischen Felder. In Walter Weiss, Herbert Ernst Wiegand & Marga Reis (eds.), *Akten des VII. Kongresses der Internationalen Vereinigung für germanische Sprach-und Literaturwissenschaft. Göttingen 1985. Band 3. Textlinguistik contra Stilistik?– Wortschatz und Wörterbuch – Grammatische oder pragmatische Organisation von Rede?* (Kontroversen, alte und neue 4), 329–340. Tübingen: Max Niemeyer Verlag.

Höhle, Tilman N. 1988. Verum-Fokus. Netzwerk Sprache und Pragmatik 5 Universität Lund-German. Inst. Lund.

Höhle, Tilman N. 1991. On reconstruction and coordination. In Hubert Haider & Klaus Netter (eds.), *Representation and derivation in the theory of grammar* (Studies in Natural Language and Linguistic Theory 22), 139–197. Dordrecht/Boston/London: Kluwer Academic Publishers.

Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Christa Dürscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im fokus. festschrift für heinz vater zum 65. geburtstag*, 107–120. Tübingen: Max Niemeyer Verlag.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Stanford, CA: CSLI Publications.

Holler, Anke. 2005. *Weiterführende Relativsätze. Empirische und theoretische Aspekte* (studia grammatica 60). Berlin: Akademie Verlag.

Hornstein, Norbert. 2013. Three grades of grammatical involvement: Syntax from a Minimalist perspective. *Mind and Language* 28(4). 392–420.

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

Hudson, Carla L. & Elissa L. Newport. 1999. Creolization: Could adults really have done it all? In Annabel Greenhill, Heather Littlefield & Cheryl Tano (eds.), *Proceedings of the boston university conference on language development*, vol. 23, 265–276. Somerville, MA: Cascadilla Press.

Hudson, Richard. 1980. Constituency and dependency. *Linguistics* 18. 179–198. Covington90a:17 genügend Restriktionen -> fest Stellung wie in Englisch.

Hudson, Richard. 1984. *Word grammar*. Oxford: Basil Blackwell. Goldberg95a:72 Vererbung, Hudson2004a:11 Determiners are Heads.

Hudson, Richard. 1989. Towards a computer-testable Word Grammar of English. *UCL Working Papers in Linguistics* 1. 321–339.

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), Stanford, CA: CSLI Publications. http://www.phon.ucl.ac.uk/home/dick/mismatch.htm.

Hudson, Richard. 2004. Are determiners heads? *Functions of Language* 11(1). 7–42. http://www.phon.ucl.ac.uk/home/dick/dets.htm.

Hudson, Richard. 2010. Reaction to: "The myth of language universals and cognitive science" on the choice between phrase structure and dependency structure. *Lingua* 120(12). 2676–2679.

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

Hudson, Richard A. 1997. German partial VP fronting. Ms. University College London. http://www.phon.ucl.ac.uk/home/dick/papers/germanpvp.htm.

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

Hudson, Richard A. 2007. *Language networks: The new word grammar*. Oxford: Oxford University Press.

Hudson Kam, Carla L. & Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1. 151–195.

Humboldt, Wilhelm von. 1988. *Gesammelte Werke*. Berlin, New York: Walter de Gruyter.

Hurford, James R. 2002. Expression/induction models of language evolution: Dimensions and issues. In Ted J. Briscoe (ed.), *Linguistic evolution through language acquisition*, 301–344. Cambridge, UK: Cambridge University Press.

Ingram, David & William Thompson. 1996. Early syntactic acquisition in German: Evidence for the modal hypothesis. *Language* 72(1). 97–120.

Islam, Md. Asfaqul, K. M. Azharul Hasan & Md. Mizanur Rahman. 2012. Basic HPSG structure for Bangla grammar. In *15th international conference on computer and information technology (iccit)*, 185–189. Curran Associates. http://dx.doi.org/10.1109/ICCITechn.2012.6509749.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA/London, England: MIT Press.

Jackendoff, Ray S. 1975. Morphological and semantic regularities in the lexikon. *Language* 51(3). 639–671.

Jackendoff, Ray S. 1977. $\overline{\text{X}}$ *syntax: A study of phrase structure*. Cambridge, MA/London, England: MIT Press.

Jackendoff, Ray S. 1997. *The architecture of the language faculty* (Linguistic Inquiry Monographs 28). Cambridge, MA/London: MIT Press.

Jackendoff, Ray S. 2000. Fodorian modularity and representational modularity. In Yosef Grodzinsky, Lewis P. Shapiro & David Swinney (eds.), *Language and the brain: Representation and processing* (Foundations of Neuropsychology), 3–30. San Diego: Academic Press.

Jackendoff, Ray S. 2002. *Foundations of language*. Oxford: Oxford University Press.

Jackendoff, Ray S. 2007. A parallel architecture perspective on language processing. *Brain Research* 1146. 2–22.

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

Jackendoff, Ray S. 2011. What is the human language faculty? two views. *Language* 87(3). 586–624.

Jackendoff, Ray S. & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97(2). 211–225.

Jackendoff, Ray S. & Steven Pinker. 2009. The reality of a universal language faculty. *The Behavioral and Brain Sciences* 32(5). 465–466.

Jacobs, Joachim. 1986. The syntax of focus and adverbials in German. In Werner Abraham & S. de Meij (eds.), *Topic, focus, and configurationality. papers from the 6th groningen grammar talks, groningen, 1984* (Linguistik Aktuell/Linguistics Today 4), 103–127. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Jacobs, Joachim. 1991. Bewegung als Valenztransfer. SFB 282: Theorie des Lexikons 1 Heinrich Heine Uni/BUGH Düsseldorf/Wuppertal.

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

Jacobson, Pauline. 1987a. Phrase structure, grammatical relations, and discontinuous constituents. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Discontinuous constituency* (Syntax and Semantics 20), 27–69. New York: Academic Press.

Jacobson, Pauline. 1987b. Review of generalized phrase structure grammar. *Linguistics and Philosophy* 10(3). 389–426.

Jaeggli, Osvaldo A. 1986. Passive. *Linguistic Inquiry* 17(4). 587–622.

Jäger, Gerhard & Reinhard Blutner. 2003. Competition and interpretation: The German adverb *wieder* („again“). In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.), *Modifying adjuncts* (Interface Explorations 4), 393–416. Berlin: Mouton de Gruyter.

Jäppinen, H., A. Lehtola & K. Valkonen. 1986. Functional structures for parsing dependency constraints. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 461–463. University of Bonn: Association for Computational Linguistics. doi:10.3115/991365.991501. http://dx.doi.org/10.3115/991365.991501.

Johnson, David E. & Shalom Lappin. 1997. A critique of the Minimalist Programm. *Linguistics and Philosophy* 20(3). 273–333.

Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Johnson, David E. & Paul M. Postal. 1980. *Arc Pair Grammar*. Princeton, NJ: Princeton University Press.

Johnson, Jacqueline S. & Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21(1). 60–99.

Johnson, Kent. 2004. Gold's theorem and cognitive science. *Philosophy of Science* 71(4). 571–592.

Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24(5). 871–882.

Johnson, Mark. 1988. *Attribute-value logic and the theory of grammar* (CSLI Lecture Notes 14). Stanford, CA: CSLI Publications.

Johnson, Mark. 1989. Parsing as deduction: The use of knowledge of language. *Journal of Psycholinguistic Research* 18(1). 105–128.

Johnson, Mark, Stuart Geman, Stephen Canon, Zhiyi Chi & Stefan Riezler. 1999. Estimators for stochastic 'unification-based' grammars. In Robert Dale & Ken Church (eds.), *Proceedings of the thirty-seventh annual meeting of the acl*, 535–541. Association for Computational Linguistics.

Joshi, Aravind K. 1987a. Introduction to tree adjoining grammar. In Alexis Manaster-Ramer (ed.), *The mathematics of language*, 87–114. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Joshi, Aravind K. 1987b. Word-order variation in natural language generation. In *Aaai 87, sixth national conference on artificial intelligence*, 550–555. Seattle.

Joshi, Aravind K., Tilman Becker & Owen Rambow. 2000. Complexity of scrambling: A new twist to the competence-performance distinction. In Abeillé & Rambow (2000) 167–181.

Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree adjunct grammar. *Journal of Computer and System Science* 10(2). 136–163.

Joshi, Aravind K. & Yves Schabes. 1997. Tree-adjoning grammars. In G. Rozenberg & A. Salomaa (eds.), *Handbook of formal languages*, 69–123. Berlin/Heidelberg/New York, NY: Springer Verlag.

Joshi, Aravind K., K. Vijay Shanker & David Weir. 1990. The convergence of mildly context-sensitive grammar formalisms. Tech. Rep. MS-CIS-90-01 Department of Computer and Information Science, University of Pennsylvania. http://repository.upenn.edu/cis_reports/539/.

Jungen, Oliver & Horst Lohnstein. 2006. *Einführung in die Grammatiktheorie* (UTB 2676). München: Wilhelm Fink Verlag.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20(2). 137–194.

Kahane, Sylvain. 1997. Bubble trees and syntactic representations. In Tilman Becker & Hans-Ulrich Krieger (eds.), *Proceedings of mathematics of language (mol5) meeting*, 70–76. Saarbrücken: DFKI.

Kahane, Sylvain. 2003. The meaning-text theory. In Ágel et al. (2003) 546–570.

Kahane, Sylvain. 2009. On the status of phrases in Head-Driven Phrase Structure Grammar: Illustration by a fully lexical treatment of extraction. In Alain Polguère & Igor A. Mel'čuk (eds.), *Dependency in linguistic description* (Studies in Language Companion Series 111), 111–150. John Benjamins Publishing Co.

Kahane, Sylvain, Alexis Nasr & Owen Rambow. 1998. Pseudo-projectivity: A polynomially parsable non-projective Dependency Grammar. In *Proceedings of the 36th annual meeting of the association for computational linguistics and 17th international conference on computational linguistics*, 646–652. Montreal, Quebec, Canada: Association for Computational Linguistics. doi:10.3115/980845.980953. http://www.aclweb.org/anthology/P98-1106.

Kallmeyer, Laura. 2005. Tree-local multicomponent tree adjoining grammars with shared nodes. *Computational Linguistics* 31(2). 187–225.

Kallmeyer, Laura & Aravind K. Joshi. 2003. Factoring predicate argument and scope semantics: Underspecified semantics with LTAG. *Research on Language and Computation* 1(1–2). 3–58. doi:10.1023/A:1024564228892.

Kallmeyer, Laura, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert & Kilian Evang. 2008. Tulipa: Towards a multi-formalism parsing environment for grammar engineering. In Stephen Clark & Tracy Holloway King (eds.), *Coling 2008: Proceedings of the workshop on grammar engineering across frameworks*, 1–8. Manchester, England: Association for Computational Linguistics. http://www.sfs.uni-tuebingen.de/~lk/papers/tulipa-geaf08.pdf.

Kallmeyer, Laura & Rainer Osswald. 2012. A frame-based semantics of the dative alternation in Lexicalized Tree Adjoining Grammars. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 9, 167–184. Paris: CNRS.

Kallmeyer, Laura & Maribel Romero. 2008. Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation* 6(1). 3–52.

Kallmeyer, Laura & Sinwon Yoon. 2004. Tree-local MCTAG with shared nodes: An analysis of word order variation in German and Korean. *Traitement automatique des langues TAL* 45(3). 49–69.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory* (Studies in Linguistics and Philosophy 42). Dordrecht/Boston/London: Kluwer Academic Publishers.

Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Dalrymple et al. (1995) 7–27. http://www.sx.ac.uk/linguistics/external/LFG/www-lfg.stanford.edu/pubs/papers/kaplan/kaplan-1989-0000.pdf.

Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 173–281. Cambridge, MA/London: MIT Press. http://www2.parc.com/istl/groups/nltt/papers/kb82-95.pdf. Reprint in: Dalrymple et al. (1995: 29–130).

Kaplan, Ronald M. & John T. Maxwell III. 1996. LFG grammar writer's workbench. Tech. rep. Xerox PARC. http://www2.parc.com/isl/groups/nltt/medley/.

Kaplan, Ronald M., Stefan Riezler, Tracy Holloway King, John T. Maxwell III, Alexander Vasserman & Richard Crouch. 2004. Speed and accuracy in shallow and deep stochastic parsing. In *Proceedings of the human language technology conference and the 4th annual meeting of the north american chapter of the association for computational linguistics (hlt-naacl'04)*, Boston, MA: Association for Computational Linguistics. http://www2.parc.com/isl/members/riezler/PAPERS/NAACL04.pdf.

Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure and functional uncertainty. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 17–42. Chicago/London: The University of Chicago Press. http://www2.parc.com/istl/members/zaenen/publications/LdDCSandFU.pdf.

Karimi, Simin. 2005. *A minimalist approach to scrambling. evidence from Persian* (Studies in Generative Grammar 76). Berlin, New York: Mouton de Gruyter.

Karimi-Doostan, Gholamhossein. 2005. Light verbs and structural case. *Lingua* 115(12). 1737–1756.

Karmiloff-Smith, Annette. 1998. Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences* 2(10). 389–398.

Karmiloff-Smith, Annette, Julia Grant, Ioanna Berthoud, Mark Davies, Patricia Howlin & Orlee Udwin. 1997. Language in Williams Syndrome: How intact is 'intact?'. *Child Development* 68(2). 246–262.

Karttunen, Lauri. 1986. Radical lexicalism. Report No. CSLI-86-68 Center for the Study of Language and Information.

Karttunen, Lauri. 1989. Radical lexicalism. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 43–65. Chicago/London: The University of Chicago Press. http://www2.parc.com/istl/members/karttune/publications/archive/RadicalLexicalism.pdf.

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In Nerbonne et al. (1994) 39–70.

Kasper, Robert T., Bernd Kiefer, Klaus Netter & Vijay K. Shanker. 1995. Compilation of HPSG to TAG. In Hans Uszkoreit (ed.), *33rd annual meeting of the association for computational linguistics. proceedings of the conference*, 92–99. Cambridge, MA: Association for Computational Linguistics.

Kathol, Andreas. 1995. *Linearization-based German syntax*: Ohio State University dissertation.

Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.

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

Kathol, Andreas & Carl J. Pollard. 1995. Extraposition via complex domain formation. In Hans Uszkoreit (ed.), *33rd annual meeting of the association for computational linguistics. proceedings of the conference*, 174–180. Cambridge, MA: Association for Computational Linguistics.

Kaufmann, Ingrid. 1995. *Konzeptuelle Grundlagen semantischer Dekompositionsstrukturen. Die Kombinatorik lokaler Verben und prädikativer Elemente* (Linguistische Arbeiten 335). Tübingen: Max Niemeyer Verlag.

Kaufmann, Ingrid & Dieter Wunderlich. 1998. Cross-linguistic patterns of resultatives. SFB 282: Theorie des Lexikons 109 Heinrich Heine Uni/BUGH Düsseldorf.

Kaufmann, Tobias. 2009. *A rule-based language model for speech recognition*: Computer Engineering and Networks Laboratory, ETH Zürich dissertation.

Kaufmann, Tobias & Beat Pfister. 2007. Applying licenser rules to a grammar with continuous constituents. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 150–162. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/.

Kaufmann, Tobias & Beat Pfister. 2008. Applying a grammar-based language model to a broadcast-news transcription task. In Johanna D. Moore, Simone Teufel, James Allan & Sadaoki Furui (eds.), *Proceedings of the 46th annual meeting of the association for computational linguistics: Human language technologies*, 106–113. Columbus, Ohio: Association for Computational Linguistics. http://www.aclweb.org/anthology/P/P08/P08-1013.pdf.

Kay, Martin. 1967. Experiments with a powerful parser. In *Proceedings of conference internationale sur le traitement automatique des langues (COLING 1967)*, Grenoble. http://aclweb.org/anthology/C/C67/C67-1009.pdf.

Kay, Martin. 2000. David g. hays. In William J. Hutchins (ed.), *Early years in machine translation* (Amsterdam Studies in the Theory and History of Linguistics Science Series 3), 165–170. John Benjamins Publishing Co.

Kay, Martin. 2011. Zipf's law and *L'Arbitraire du Signe*. *Linguistic Issues in Language Technology* 6(8). http://elanguage.net/journals/index.php/lilt/issue/view/330. Special Issue on Interaction of Linguistics and Computational Linguistics.

Kay, Paul. 2002. An informal sketch of a formal architecture for Construction Grammar. *Grammars* 5(1). 1–19. http://www.icsi.berkeley.edu/~kay/cg.arch.pdf.

Kay, Paul. 2005. Argument structure constructions and the argument-adjunct distinction. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical constructions: Back to the roots* (Constructional Approaches to Language 4), 71–98. Amsterdam/Philadelphia: John Benjamins Publishing Co. http://www.icsi.berkeley.edu/~kay/ASCs.pdf.

Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the what's x doing y? construction. *Language* 75(1). 1–33.

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

Kayne, Richard S. 2011. Why are there no directionality parameters? In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th west coast conference on formal linguistics*, 1–23. Somerville, MA: Cascadilla Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Keller, Frank. 1994. German functional HPSG – an experimental CUF encoding. Tech. rep. Institut für Maschinelle Sprachverarbeitung Stuttgart.

Keller, Frank. 1995. Towards an account of extraposition in HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the seventh conference of the European*

Chapter of the Association for Computational Linguistics, Dublin: Association for Computational Linguistics. http://homepages.inf.ed.ac.uk/keller/papers/eacl95.html.

Kettunen, Kimmo. 1986. On modelling dependency-oriented parsing. In Fred Karlsson (ed.), *Papers from the fifth scandinavian conference of computational linguistics*, 113–120. Helsinki.

Kiefer, Bernd, Hans-Ulrich Krieger & Mark-Jan Nederhof. 2000. Efficient and robust parsing of word hypotheses graphs. In Wahlster (2000) 280–295.

Kifle, Nazareth Amlesom. 2012. *Tigrinya applicatives in Lexical-Functional Grammar*: University of Bergen dissertation. http://hdl.handle.net/1956/5730.

Kim, Jong-Bok & Peter Sells. 2008. *English syntax: An introduction* (CSLI Lecture Notes 185). Stanford, CA: CSLI Publications.

Kim, Jong-Bok, Peter Sells & Jaehyung Yang. 2007. Parsing two types of multiple nominative constructions: A constructional approach. *Language and Information* 11(1). 25–37.

Kim, Jong-Bok & Jaehyung Yang. 2003. Korean phrase structure grammar and its implementations into the LKB system. In Dong Hong Ji & Kim Teng Lua (eds.), *Proceedings of the 17th pacific asia conference on language, information and computation*, 88–97. National University of Singapore: COLIPS Publications.

Kim, Jong-Bok & Jaehyung Yang. 2004. Projections from morphology to syntax in the Korean Resource Grammar: Implementing typed feature structures. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent text processing: 5th international conference, cicling 2004, seoul, korea, february 15-21, 2004, proceedings* (Lecture Notes in Computer Science 2945), 13–24. Berlin/Heidelberg/New York, NY: Springer Verlag.

Kim, Jong-Bok & Jaehyung Yang. 2006. Coordination structures in a typed feature structure grammar: Formalization and implementation. In Tapio Salakoski, Filip Ginter, Sampo Pyysalo & Tapio Pahikkala (eds.), *Advances in natural language processing advances in natural language processing: 5th international conference, fintal 2006 turku, finland, august 23-25, 2006 proceedings* (Lecture Notes in Artificial Intelligence 4139), 194–205. Berlin/Heidelberg/New York, NY: Springer Verlag.

Kim, Jong-Bok & Jaehyung Yang. 2009. Processing three types of Korean cleft constructions in a typed feature structure grammar. *Korean Journal of Cognitive Science* 20(1). 1–28.

Kim, Jong-Bok, Jaehyung Yang, Sanghoun Song & Francis Bond. 2011. Deep processing of Korean and the development of the Korean Resource Grammar. *Linguistic Research* 28(3). 635–672.

Kimball, John P. 1973. *The formal theory of grammar* (Foundations of Modern Linguistics). Englewood Cliffs, NJ: Prentice-Hall.

King, Paul. 1994. An expanded logical formalism for Head-Driven Phrase Structure Grammar. Arbeitspapiere des SFB 340 No. 59 Eberhard-Karls-Universität Tübingen. http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/59/59abs.html.

King, Paul. 1999. Towards truth in Head-Driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitsberichte des SFB 340 No. 132), 301–352. Tübingen: Universität Tübingen. http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/132/132abs.html.

King, Tracy Holloway & John T. Maxwell III. 2007. Overlay mechanisms for multi-level deep processing applications. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE), 182–202. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/.

Kinyon, Alexandra, Owen Rambow, Tatjana Scheffler, SinWon Yoon & Aravind K. Joshi. 2006. The metagrammar goes multilingual: A cross-linguistic look at the v2-phenomenon. In Laura Kallmeyer & Tilman Becker (eds.), *Tag+8. the eighth international workshop on tree adjoining grammar and related formalisms: Proceedings of the workshop*, 17–24. Sydney, Australia: Association for Computational Linguistics.

Kiparsky, Paul. 1987. Morphology and grammatical relations. Unpublished paper, Stanford University, Stanford.

Kiparsky, Paul. 1988. Agreement and linking theory. Unpublished paper, Stanford University, Stanford.

Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 143–173. The Hague/Paris: Mouton.

Kiss, Katalin E. 2003. Argument scrambling, focus movement and topic movement in Hungarian. In Simin Karimi (ed.), *Word order and scrambling*, 22–43. London: Blackwell.

Kiss, Tibor. 1991. The grammars of LILOG. In Otthein Herzog & Claus-Rainer Rollinger (eds.), *Text understanding in lilog* (Lecture Notes in Artificial Intelligence 546), 183–199. Berlin/Heidelberg/New York, NY: Springer Verlag.

Kiss, Tibor. 1992. Variable Subkategorisierung. Eine Theorie unpersönlicher Einbettungen im Deutschen. *Linguistische Berichte* 140. 256–293.

Kiss, Tibor. 1993. Infinite Komplementation – Neue Studien zum deutschen Verbum infinitum. Arbeiten des SFB 282 No. 42 Bergische Universität Gesamthochschule Wuppertal.

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

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

Kiss, Tibor. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2). 281–334.

Kiss, Tibor & Birgit Wesche. 1991. Verb order and head movement. In Otthein Herzog & Claus-Rainer Rollinger (eds.), *Text understanding in lilog* (Lecture Notes in Artificial Intelligence 546), 216–242. Berlin/Heidelberg/New York, NY: Springer Verlag.

Klann-Delius, Gisela. 2008. *Spracherwerb*. Stuttgart: J.B. Metzler–Verlag 2nd edn.

Klein, Wolfgang. 1971. *Parsing: Studien zur maschinellen Satzanalyse mit Abhängigkeitsgrammatiken und Transformationsgrammatiken*, vol. 2. Frankfurt a. M.: Athenäum Verlag.

Klein, Wolfgang. 1985. Ellipse, Fokusgliederung und thematischer Stand. In Reinhard Meyer-Hermann & Hannes Rieser (eds.), *Ellipsen und fragmentarische ausdrücke*, 1–24. Tübingen: Max Niemeyer Verlag.

Klein, Wolfgang. 1986. *Second language acquisition* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

Klein, Wolfgang. 2009. Finiteness, Universal Grammar and the language faculty. In Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura & Seyda Ozcaliskan (eds.), *Cross-linguistic approaches to the study of language: Research in the tradition of Dan Isaac Slobin* (Psychology Press Festschrift Series), 333–344. New York: Psychology Press.

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

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Helen Goodluck & Michael Rochemont (eds.), *Island constraints: Theory, acquisition, and processing*, 223–258. Dordrecht/Boston/London: Kluwer Academic Publishers.

Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8(4). 573–633.

Knecht, Laura. 1985. *Subject and object in Turkish*: M.I.T. dissertation.

Kobele, Gregory M. 2006. *Generating copies: An investigation into structural identity in language and grammar*: UCLA dissertation.

Kobele, Gregory M. 2008. Across-the-board extraction in minimalist grammars. In *Proceedings of the ninth international workshop on tree adjoining grammar and related formalisms (tag+9)*, 113–128.

Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Koenig, Jean-Pierre & Karin Michelson. 2010. Argument structure of Oneida kinship terms. *International Journal of American Linguistics* 76(2). 169–205.

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

Kohl, Dieter. 1992. Generation from under- and overspecified structures. In Antonio Zampoili (ed.), *14th international conference on computational linguistics (COLING '92), august23–28*, 686–692. Nantes, France: Association for Computational Linguistics.

Kohl, Dieter & Stefan Momma. 1992. LFG-based generation in ACORD. In Gabriel Bes (ed.), *The construction of a natural language and graphic interface: Results and perspectives from the ACORD project*, Berlin/Heidelberg/New York, NY: Springer Verlag.

Kohl, Karen T. 1999. *An analysis of finite parameter learning in linguistic spaces*. Massachusetts Institute of Technology MA thesis. http://karentkohl.org/SM.pdf.

Kohl, Karen T. 2000. Language learning in large parameter spaces. In *Proceedings of the seventeenth national conference on artificial intelligence and twelfth conference on innovative applications of artificial intelligence*, 1080. AAAI Press / The MIT Press.

Kolb, Hans-Peter. 1997. Gb blues: Two essays on procedures and structures in generative syntax. Arbeitspapiere des SFB 340 No. 110 Eberhard-Karls-Universität Tübingen.

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), Dordrecht/Boston/London: Kluwer Academic Publishers.

Konieczny, Lars. 1996. *Human sentence processing: a semantics-oriented parsing approach*: Universität Freiburg Dissertation. http://www.researchgate.net/publication/36150321_Human_sentence_processing__a_semantics-oriented_parsing_approach_/. IIG-Berichte 3/96.

König, Esther. 1999. LexGram – a practical Categorial Grammar formalism. *Journal of Language and Computation* 1(1). 33–52.

Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85(2–3). 211–258.

Kordoni, Valia (ed.). 1999. *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitspapiere des SFB 340, No. 132, Volume 1). Eberhard-Karls-Universität Tübingen.

Kordoni, Valia. 2001. Linking experiencer-subject psych verb constructions in Modern Greek. In Daniel P. Flickinger & Andreas Kathol (eds.), *Proceedings of the hpsg-2000 conference, university of california, berkeley*, 198–213. CSLI Publications. http://cslipublications.stanford.edu/HPSG/1/.

Kordoni, Valia & Julia Neu. 2005. Deep analysis of Modern Greek. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural language processing ijcnlp 2004* (Lecture Notes in Artificial Intelligence 3248), 674–683. Berlin/Heidelberg/New York, NY: Springer Verlag.

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

Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1(2). 111–136.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

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

Koster, Jan. 1987. *Domains and dynasties. the radical autonomy of syntax*. Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

Kratzer, Angelika. 1984. On deriving syntactic differences between German and English. TU Berlin, ms.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht/Boston/London: Kluwer Academic Publishers.

Krieger, Hans-Ulrich & John Nerbonne. 1993. Feature-based inheritance networks for computational lexicons. In Ted Briscoe, Ann Copestake & Valeria de Paiva (eds.), *Inheritance, defaults, and the lexicon*, 90–136. Cambridge, UK: Cambridge University Press. http://www.dfki.de/lt/publications_show.php?id=342. A version of this paper is available as DFKI Research Report RR-91-31. Also published in: Proceedings of the ACQUILEX Workshop on Default Inheritance in the Lexicon, Technical Report No. 238, University of Cambridge, Computer Laboratory, October 1991.

Kroch, Anthony S. 1987. Unbounded dependencies and subjacency in a Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *Mathematics of language*, 143–172. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Kroch, Anthony S. & Aravind K. Joshi. 1985. The linguistic relevance of Tree Adjoining Grammar. Tech. Rep. MS-CIS-85-16 University of Pennsylvania. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1706&context=cis_reports.

Kroch, Anthony S. & Aravind K. Joshi. 1987. Analyzing extraposition in a tree adjoining grammar. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Discontinuous constituency* (Syntax and Semantics 20), 107–149. New York: Academic Press.

Kropp Dakubu, Mary Esther, Lars Hellan & Dorothee Beermann. 2007. Verb sequencing constraints in ga: Serial verb constructions and the extended verb complex. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 99–119. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/.

Kruijff-Korbayová, Ivana & Mark J. Steedman. 2003. Discourse on information structure. *Journal of Logic, Language and Information: Special Issue on Discourse and Information Structure* 12(3). 249–259.

Kuhn, Jonas. 1995. Information packaging in German. some motivation from HPSG-based translation. Universität Stuttgart, ms. ftp://ftp.ims.uni-stuttgart.de/pub/papers/kuhn/Info-Pack.pdf.

Kuhn, Jonas. 1996. An underspecified HPSG representation for information structure. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 670–675. Copenhagen, Denmark: Association for Computational Linguistics.

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

Kuhn, Jonas & Christian Rohrer. 1997. Approaching ambiguity in real-life sentences – the application of an Optimality Theory-inspired constraint ranking in a large-scale LFG grammar. In *Proceedings of DGfS/CL 97*, Heidelberg.

Kuhns, Robert J. 1986. A PROLOG implementation of Government-Binding Theory. In Alan W. Biermann (ed.), *Proceedings of the twenty-fourth annual meeting of the association for computational linguistics*, 546–550. Columbia University, New York: Association for Computational Linguistics.

Kunze, Jürgen. 1968. The treatment of non-projective structures in the syntactic analysis and synthesis of English and German. *Computational Linguistics* 7. 67–77.

Kunze, Jürgen. 1975. *Abhängigkeitsgrammatik* (studia grammatica 12). Berlin: Akademie Verlag.

Kunze, Jürgen. 1991. *Kasusrelationen und semantische Emphase* (studia grammatica XXXII). Berlin: Akademie Verlag.

Kunze, Jürgen. 1993. *Sememstrukturen und Feldstrukturen* (studia grammatica 36). Berlin: Akademie Verlag. Unter Mitarbeit von Beate Firzlaff.

Labelle, Marie. 2007. Biolinguistics, the Minimalist Program, and psycholinguistic reality. *Snippets* 14. http://www.ledonline.it/snippets/.

Laczkó, Tibor, György Rákosi & Ágoston Tóth. 2010. Hungram vs. enggram in pargram: On the comparison of hungarian and english in an international computational linguistics project. In Irén Hegedűs & Sándor Martsa (eds.), *Selected papers in linguistics from the 9th husse conference*, vol. 1, 81–95. Pécs: Institute of English Studies, Faculty of Humanities, University of Pécs.

Laenzlinger, Christoph. 2004. A feature-based theory of adverb syntax. In Jennifer R. Austin, Stefan Engelberg & Gisa Rauh (eds.), *Adverbials: The Interplay Between Meaning, Context, and Syntactic Structure* (Linguistik Aktuell/Linguistics Today 70), 205–252. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Lai, Cecilia S. L., Simon E. Fisher, Jane A. Hurst, Faraneh Vargha-Khadem & Anthony P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413(6855). 519–523. doi:10.1038/35097076.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL/London: The University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. 1. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford, CA: CSLI Publications.

Langacker, Ronald W. 2008. *Cognitive grammar. a basic introduction.* Oxford: Oxford University Press.

Langacker, Ronald W. 2009. Cognitive (construction) grammar. *Cognitive Linguistics* 20(1). 167–176.

Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000a. The revolution confused: A response to our critics. *Natural Language and Linguistic Theory* 18(4). 873–890.

Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000b. The structure of unscientific revolutions. *Natural Language and Linguistic Theory* 18(3). 665–671.

Lappin, Shalom, Robert D. Levine & David E. Johnson. 2001. The revolution maximally confused. *Natural Language and Linguistic Theory* 19(4). 901–919.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391. http://semlab5.sbs.sunysb.edu/~rlarson/larson88do.pdf.

Lascarides, Alex & Ann Copestake. 1999. Default representation in constraint-based frameworks. *Computational Linguistics* 25(1). 55–105. http://acl.ldc.upenn.edu/J/J99/J99-1002.pdf.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on its application and output* (Current Studies in Linguistics 22). Cambridge, MA: MIT Press.

Lasnik, Howard & Juan Uriagereka. 2002. On the poverty of the challenge. *The Linguistic Review* 19(1–2). 147–150.

Le, Hong Phuong, Thi Minh Huyen Nguyen & Azim Roussanaly. 2008. Metagrammar for Vietnamese LTAG. In *Proceedings of the ninth international workshop on tree adjoining grammars and related formalisms (tag+9)*, 129–132. Tübingen.

Legate, Julie & Charles D. Yang. 2002. Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review* 19(1–2). 151–162.

Lehtola, Aarno. 1986. Dpl – a computational method for describing grammars and modelling parsers. In Fred Karlsson (ed.), *Papers from the fifth scandinavian conference of computational linguistics*, 151–159. Helsinki.

Leiss, Elisabeth. 2003. Empirische Argumente für Dependenz. In Ágel et al. (2003) 311–324.

Leiss, Elisabeth. 2009. *Sprachphilosophie* (de Gruyter Studienbuch). Berlin: Walter de Gruyter.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen* (Studien zur deutschen Grammatik 5). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Lenerz, Jürgen. 1994. Pronomenprobleme. In Brigitta Haftka (ed.), *Was determiniert Wortstellungsvariation? Studien zu einem Interaktionsfeld von Grammatik, Pragmatik und Sprachtypologie*, 161–174. Opladen: Westdeutscher Verlag.

Lenneberg, Eric H. 1964. The capacity for language acquisition. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The structure of language*, 579–603. Englewood Cliffs, NJ: Prentice-Hall.

Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: John Wiley & Sons, Inc.

Levelt, Willem J. M. 1989. *Speaking: From intonation to articulation* (ACL-MIT Press Series in Natural Language Processing). Cambridge, MA: MIT Press.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago, Illinois: University of Chicago Press.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge University Press.

Levine, Robert D. 2003. Adjunct valents, cumulative scopings and impossible descriptions. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th international conference on head-driven phrase structuregrammar*, 209–232. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/3/.

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

Levine, Robert D. & Walt Detmar Meurers. 2006. Head-Driven Phrase Structure Grammar: Linguistic approach, formal foundations, and computational realization. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, Oxford: Elsevier Science Publisher B.V. (North-Holland) 2nd edn. http://www.sfs.uni-tuebingen.de/~dm/papers/ell2-hpsg.html.

Lewis, Geoffrey L. 1967. *Turkish grammar*. Oxford: Clarendon Press.

Lewis, John D. & Jeffrey L. Elman. 2001. Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. In Barbora Skarabela, Sarah Fish & Anna H.-J. Do (eds.), *Proceedings of the 26th annual boston university conference on language development*, 359–370. http://crl.ucsd.edu/~elman/Papers/BU2001.pdf.

Li, Wei. 1996. Esperanto inflection and its interface in HPSG. Working papers of the linguistics circle University of Victoria.

Lichte, Timm. 2007. An MCTAG with tuples for coherent constructions in German. In Laura Kallmeyer, Paola Monachesi, Gerald Penn & Giorgio Satta (eds.), *Proceedings of the 12th conference on formal grammar 2007*, Dublin, Ireland.

Lieb, Hans-Heinrich. 1983. *Integrational linguistics. vol. i.: General outline* (Current Issues in Linguistic Theory 17). Amsterdam/Philadelphia: John Benjamins Publishing Co.

Lightfoot, David W. 1997. Catastrophic change and learning theory. *Lingua* 100(1). 171–192.

Link, Godehard. 1984. Hydras. on the logic of relative constructions with multiple heads. In Fred Landmann & Frank Veltman (eds.), *Varieties of formal semantics*, 245–257. Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

Lipenkova, Janna. 2009. *Serienverbkonstruktionen im Chinesischen und ihre Analyse im Rahmen von HPSG*. Institut für Sinologie, Freie Universität Berlin MA thesis. http://hpsg.fu-berlin.de/~lipenkov/magister.html.

Liu, Gang. 1997. *Eine unifikations-basierte Grammatik für das moderne Chinesisch – dargestellt in der HPSG*: FG Sprachwissenschaft, Universität Konstanz dissertation. http://www.ub.uni-konstanz.de/kops/volltexte/1999/191/.

Lloré, F. Xavier. 1995. *Un método de 'parsing' para gramáticas categoriales multimodales*: I.C.E. de la Universidad Politécnica de Catalunya dissertation.

Lobin, Henning. 1993. *Koordinationssyntax als strukturales phänomen* (Studien zur Grammatik 46). Tübingen: Gunter Narr Verlag.

Lobin, Henning. 2003. Dependenzgrammatik und Kategorialgrammatik. In Ágel et al. (2003) 325–330.

Löbner, Sebastian. 1986. In Sachen Nullartikel. *Linguistische Berichte* 101. 64–65. http://user.phil-fak.uni-duesseldorf.de/~loebner/publ/Nullartikel.html.

Lohndal, Terje. 2012. Toward the end of argument structure. In María Cristina Cuervo & Yves Roberge (eds.), *The end of argument structure?*, vol. 38 (Syntax and Semantics), 155–184. Bingley, UK: Emerald Group Publishing.

Lohnstein, Horst. 1993. *Projektion und linking. ein prinzipienbasierter parser fürs deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag.

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

Longobardi, Giuseppe & Ian Roberts. 2010. Universals, diversity and change in the science of language: Reaction to "the myth of language universals and cognitive science". *Lingua* 120(12). 2699–2703.

Lorenz, Konrad. 1970. *Studies in human and animal behavior*, vol. I. Cambridge, MA: Harvard University Press.

Lötscher, Andreas. 1985. Syntaktische Bedingungen der Topikalisierung. *Deutsche Sprache* 13(3). 207–229.

Lüdeling, Anke. 2001. *On particle verbs and similar constructions in German* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.

Luuk, Erkki & Hendrik Luuk. 2011. The redundancy of recursion and infinity for natural language. *Cognitive Processing* 12(1). 1–11.

Maas, Heinz Dieter. 1977. The Saarbrücken Automatic Translation System (SUSY). In *Proceedings of the third european congress on information systems and networks, overcoming the language barrier*, vol. 1, 585–592. München: Verlag Dokumentation.

Maché, Jakob. 2010. Towards a compositional analysis of verbless directives in German. Paper presented at the HPSG 2010 Conference.

MacWhinney, Brian. 1995. *The CHILDES project: Tools for analyzing talk.* Hillsdale, NJ: Erlbaum 2nd edn.

Maess, Burkhard, Stefan Koelsch, Thomas C. Gunter & Angela D. Friederici. 2001. Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience* 4(5). 540–545.

Marantz, Alec. 1984. *On the nature of grammatical relations* (Linguistic Inquiry Monographs 10). Cambridge, MA: MIT Press.

Marantz, Alec. 1997. No escape from syntax. don't try morphological analysis in the privacy of your own lexicon. *U. Penn Working Papers in Linguistics* 4(2). 201–225. http://www.ling.upenn.edu/papers/v4.2-contents.html.

Marantz, Alec. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22(2–4). 429–445.

Marcus, Gary F. 1993. Negative evidence in language acquisition. *Cognition* 46(1). 53–85.

Marcus, Gary F. & Simon E. Fisher. 2003. Foxp2 in focus: What can genes tell us about speech and language? *TRENDS in Cognitive Sciences* 7(6). 257–262.

Marcus, Mitchell P. 1980. *A theory of syntactic recognition for natural language.* London, England/Cambridge, MA: MIT Press.

Marimon, Montserrat. 2013. The Spanish DELPH-IN grammar. *Language Resources and Evaluation* 47(2). 371–397. http://dx.doi.org/10.1007/s10579-012-9199-7.

Marshall, Ian & Éva Sáfár. 2004. Sign Language generation in an ALE HPSG. In Stefan Müller (ed.), *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, center for computational linguistics, katholieke universiteit leuven*, 189–201. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2004/.

Marslen-Wilson, William. 1975. Sentence perception as an interactive parallel process. *Science* 189(4198). 226–228.

Masuichi, Hiroshi & Tomoko Ohkuma. 2003. Constructing a practical Japanese parser based on Lexical-Functional Grammar. *Journal of Natural Language Processing* 10. 79–109. In Japanese.

Masum, Mahmudul Hasan, Muhammad Sadiqul Islam, M. Sohel Rahman & Reaz Ahmed. 2012. HPSG analysis of type-based Arabic nominal declension. In *The 13th international Arab conference*, 272–279.

Mayo, Bruce. 1997. Die Konstanzer LFG-Umgebung. Arbeitspapier 82 des Fachbereichs Sprachwissenschaft der Universität Konstanz Universität Konstanz.

Mayo, Bruce. 1999. *A computational model of derivational morphology*: Universität Hamburg dissertation. http://www.sub.uni-hamburg.de/opus/volltexte/1999/386/.

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

Meisel, Jürgen. 1995. Parameters in acquisition. In Paul Fletcher & Brian MacWhinny (eds.), *The handbook of child language*, 10–35. Oxford: Blackwell Publishing Ltd.

Mel'čuk, Igor A. 1964. *Avtomatičeskij sintaksičeskij analiz*, vol. 1. Novosibirsk: Izdatel´stvo SO AN SSSR.

Mel'čuk, Igor A. 1981. Meaning-text models: A recent trend in Soviet linguistics. *Annual Review of Anthropology* 10. 27–62.

Mel'čuk, Igor A. 1988. *Dependency syntax: Theory and practice* (SUNY Series in Linguistics). Albany, NY: SUNY Press.

Melnik, Nurit. 2007. From 'hand-written' to computationally implemented HPSG theories. *Research on Language and Computation* 5(2). 199–236.

Mensching, Guido & Eva-Maria Remberger. 2011. Syntactic variation and change in romance: A minimalist approach. In Peter Siemund (ed.), *Linguistic universals and language variation* (Trends in Linguistics. Studies and Monographs 231), Berlin/New York, NY: Mouton de Gruyter.

Menzel, Wolfgang. 1998. Constraint satisfaction for robust parsing of spoken language. *Journal of Experimental & Theoretical Artificial Intelligence* 10(1). 77–89.

Menzel, Wolfgang & Ingo Schröder. 1998a. Constraint-based diagnosis for intelligent language tutoring systems. In *Proceedings of the itf & knows conference at the ıfip '98 congress*, Wien/Budapest.

Menzel, Wolfgang & Ingo Schröder. 1998b. Decision procedures for Dependency Parsing using graded constraints. In Sylvain Kahane & Alain Polguère (eds.), *Processing of dependency-based grammars: Proceedings of the workshop. coling-acl '98*, Association for Computational Linguistics.

Meurer, Paul. 2009. A computational grammar for Georgian. In *7th international tbilisi symposium on logic, language, and computation, tbillc 2007, tbilisi, georgia, october 1–5, 2007, revised selected papers* (Lecture Notes in Artificial Intelligence 5422), 1–15. Berlin/ Heidelberg/New York, NY: Springer Verlag.

Meurers, Walt Detmar. 1994. On implementing an HPSG theory. In Erhard W. Hinrichs, Walt Detmar Meurers & Tsuneko Nakazawa (eds.), *Partial-VP and Split-NP Topicalization in German – An HPSG Analysis and its Implementation* (Arbeitspapiere des SFB 340 No. 58), Tübingen: Eberhard-Karls-Universität. http://www.sfs.uni-tuebingen.de/~dm/on-implementing.html.

Meurers, Walt Detmar. 1999a. German partial-VP fronting revisited. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 129–144. Stanford, CA: CSLI Publications. http://www.sfs.uni-tuebingen.de/~dm/papers/hpsg-volume98/pvp-revisited.html.

Meurers, Walt Detmar. 1999b. *Lexical generalizations in the syntax of German non-finite constructions*. Tübingen: Eberhard-Karls-Universität dissertation.

Meurers, Walt Detmar. 1999c. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. http://www.sfs.uni-

tuebingen.de/~dm/papers/gagl99.html.
Meurers, Walt Detmar. 2000. Lexical generalizations in the syntax of German non-finite constructions. Arbeitspapiere des SFB 340 No. 145 Eberhard-Karls-Universität Tübingen. http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html.

Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217. http://www.sfs.uni-tuebingen.de/~dm/papers/lexical-generalizations.html.

Meurers, Walt Detmar, Kordula De Kuthy & Vanessa Metcalf. 2003. Modularity of grammatical constraints in HPSG-based grammar implementations. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the esslli 2003 workshop "ideas and strategies for multilingual grammar development"*, 83–90. Vienna, Austria. http://www.sfs.uni-tuebingen.de/~dm/papers/meurers-dekuthy-metcalf-03.html.

Meurers, Walt Detmar & Stefan Müller. 2009. Corpora and syntax. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics. an international handbook*, vol. 29.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), chap. 42, 920–933. Berlin: Mouton de Gruyter.

Meurers, Walt Detmar, Gerald Penn & Frank Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL*, 18–25. http://www.sfs.uni-tuebingen.de/~dm/papers/acl02.html. Proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA.

Micelli, Vanessa. 2012. Field topology and information structure: A case study for german constituent order. In Steels (2012) 178–211.

Michaelis, Jens. 2001. *On formal properties of Minimalist Grammars*: Universität Potsdam dissertation.

Michaelis, Laura A. 2006. Construction Grammar. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, Oxford: Elsevier Science Publisher B.V. (North-Holland) 2nd edn. http://spot.colorado.edu/~michaeli/publications.html.

Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond alternations: A constructional model of the German applicative pattern* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Miller, George A. & Kathryn Ojemann McKean. 1964. A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology* 16(4). 297–308.

Mittendorf, Ingo & Louisa Sadler. 2005. Numerals, nouns and number in welsh nps. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2005 conference*, 294–312. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/LFG/10/lfg05.html.

Miyao, Yusuke, Takashi Ninomiya & Jun'ichi Tsujii. 2005. Corpus-oriented grammar development for acquiring a Head-Driven Phrase Structure Grammar from the Penn Treebank. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural* *language processing ijcnlp 2004* (Lecture Notes in Artificial Intelligence 3248), 684–693. Berlin/Heidelberg/New York, NY: Springer Verlag. http://www-tsujii.is.s.u-tokyo.ac.jp/~yusuke/paper/ijcnlp04.pdf.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Moens, Marc, Jo Calder, Ewan Klein, Mike Reape & Henk Zeevat. 1989. Expressing generalizations in unification-based grammar formalisms. In Harold Somers & Mary McGee Wood (eds.), *Proceedings of the fourth conference of the european chapter of the association for computational linguistics*, 174–181. Manchester, England: Association for Computational Linguistics.

Montague, Richard. 1974. *Formal philosophy*. New Haven: Yale University Press.

Moortgat, Michael. 1989. *Categorical investigations. logical and linguistic aspects of the Lambek Calculus* (Groningen Amsterdam Studies in Semantics 9). Dordrecht/Cinnaminson, U.S.A.: Foris Publications.

Moot, Richard. 2002. *Proof nets for linguistic analysis*: University of Utrecht dissertation.

Morgan, James L. 1989. Learnability considerations and the nature of trigger experiences in language acquisition. *Behavioral and Brain Sciences* 12(2). 352–353.

Morin, Yves Ch. 1973. A computer tested Transformational Grammar of French. *Linguistics* 116(11). 49–114.

Morrill, Glyn V. 1994. *Type logical grammars: Categorial logic of signs*. Dordrecht/Boston/London: Kluwer Academic Publishers.

Morrill, Glyn V. 1995. Discontinuity in Categorial Grammar. *Linguistics and Philosophy* 18(2). 175–219.

Morrill, Glyn V. 2012. Catlog: A categorial parser/theorem-prover. In *Logical aspects of computational linguistics: System demonstrations*, 13–16. Nantes, France: University of Nantes.

Müller, Gereon. 1996a. A constraint on remnant movement. *Natural Language and Linguistic Theory* 14(2). 355–407.

Müller, Gereon. 1996b. On extraposition and successive cyclycity. In Uli Lutz & Jürgen Pafel (eds.), *On extraction and extraposition in German* (Linguistik Aktuell/Linguistics Today 11), 213–243. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Müller, Gereon. 1998. *Incomplete category fronting. a derivational approach to remnant movement in German* (Studies in Natural Language and Linguistic Theory 42). Dordrecht/Boston/London: Kluwer Academic Publishers.

Müller, Gereon. 2000a. *Elemente der optimalitätstheoretischen Syntax* (Stauffenburg Linguistik 20). Tübingen: Stauffenburg Verlag.

Müller, Gereon. 2009a. There are no constructions. Handout Ringvorlesung: *Algorithmen und Muster: Strukturen in der Sprache*. Freie Universität Berlin, 20. Mai. http://hpsg.fu-berlin.de/Events/Ring2009/gmueller.pdf.

Müller, Gereon. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Stefan Engelberg, Anke Holler & Kristel Proost (eds.), *Sprachliches wissen zwischen lexikon und grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), 211–249. Berlin/New York, NY: de Gruyter. http://www.uni-leipzig.de/~muellerg/mu242.pdf.

Müller, Gereon. 2014a. Syntactic buffers. Linguistische Arbeitsberichte 91 Institut für Linguistic Universität Leipzig. http://www.uni-leipzig.de/~muellerg/mu765.pdf.

Müller, Natascha & Beate Riemer. 1998. *Generative Syntax der romanischen Sprachen: Französisch, Italienisch, Portugiesisch, Spanisch* (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 1995. Scrambling in German – extraction into the *Mittelfeld*. In Benjamin K. T'sou & Tom Bong Yeung Lai (eds.), *Proceedings of the tenth pacific asia conference on language, information and computation*, 79–83. City University of Hong Kong. http://hpsg.fu-berlin.de/~stefan/Pub/scrambling.html.

Müller, Stefan. 1996c. The Babel-System—an HPSG fragment for German, a parser, and a dialogue component. In *Proceedings of the fourth international conference on the practical application of prolog*, 263–277. London. http://hpsg.fu-berlin.de/~stefan/Pub/babel.html.

Müller, Stefan. 1996d. Yet another paper about partial verb phrase fronting in German. In Jun-ichi Tsuji (ed.), *Proceedings of coling-96. 16th international conference on computational linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 800–805. Copenhagen, Denmark: Association for Computational Linguistics. http://hpsg.fu-berlin.de/~stefan/Pub/pvp.html.

Müller, Stefan. 1999a. *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag. http://hpsg.fu-berlin.de/~stefan/Pub/hpsg.html.

Müller, Stefan. 1999b. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105. http://hpsg.fu-berlin.de/~stefan/Pub/freeRel.html.

Müller, Stefan. 1999c. Restricting discontinuity. In *Proceedings of the 5th natural language processing pacific rim symposium 1999 (nlprs'99)*, 85–90. Peking. http://hpsg.fu-berlin.de/~stefan/Pub/restricting.html.

Müller, Stefan. 2000b. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German.* Saarbrücken: Universität des Saarlandes Habilitationsschrift. http://hpsg.fu-berlin.de/~stefan/Pub/complex.html.

Müller, Stefan. 2001. Case in German – towards an HPSG analysis. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 217–255. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/case.html.

Müller, Stefan. 2002a. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/complex.html.

Müller, Stefan. 2002b. Multiple frontings in German. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of formal grammar 2002*, 113–124. Trento. http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf.html.

Müller, Stefan. 2002c. Syntax or morphology: German particle verbs revisited. In Nicole Dehé, Ray S. Jackendoff, Andrew McIntyre & Silke Urban (eds.), *Verb-particle explorations* (Interface Explorations 1), 119–139. Berlin/New York, NY: Mouton de Gruyter.

http://hpsg.fu-berlin.de/~stefan/Pub/syn-morph-part.html.
Müller, Stefan. 2003a. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62. http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf-ds.html.
Müller, Stefan. 2003b. Object-to-subject-raising and lexical rule. an analysis of the German passive. In Stefan Müller (ed.), *Proceedings of the 10th international conference on Head-Driven Phrase Structure Grammar, michigan state university, east lansing*, vol. 4, 278–297. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/passiv-da.html.
Müller, Stefan. 2003c. Solving the bracketing paradox: an analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. http://hpsg.fu-berlin.de/~stefan/Pub/paradox.html.
Müller, Stefan. 2004a. An analysis of depictive secondary predicates in German without discontinuous constituents. In Stefan Müller (ed.), *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, center for computational linguistics, katholieke universiteit leuven*, 202–222. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/depiktiv-2004.html.
Müller, Stefan. 2004b. Complex NPs, subjacency, and extraposition. *Snippets* 8. 10–11. http://hpsg.fu-berlin.de/~stefan/Pub/subjacency.html.
Müller, Stefan. 2004c. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257. http://hpsg.fu-berlin.de/~stefan/Pub/discont.html.
Müller, Stefan. 2004d. An HPSG analysis of German depictive secondary predicates. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings of the joint meeting of the 6th conference on formal grammar and the 7th conference on mathematics of language* (Electronic Notes in Theoretical Computer Science 53), 233–245. Helsinki: Elsevier Science Publisher B.V. (North-Holland). doi:http://dx.doi.org/10.1016/S1571-0661(05)82585-X. http://hpsg.fu-berlin.de/~stefan/Pub/depiktive.html.
Müller, Stefan. 2005a. Resultative constructions: Syntax, world knowledge, and collocational restrictions. review of hans c. boas: A constructional approach to resultatives. *Studies in Language* 29(3). 651–681. http://hpsg.fu-berlin.de/~stefan/Pub/boas2003.html.
Müller, Stefan. 2005b. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39. http://hpsg.fu-berlin.de/~stefan/Pub/satz-lb.html.
Müller, Stefan. 2005c. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330. http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf-lb.html.
Müller, Stefan. 2006. Phrasal or lexical constructions? *Language* 82(4). 850–883. http://hpsg.fu-berlin.de/~stefan/Pub/phrasal.html.
Müller, Stefan. 2007a. The Grammix CD Rom. a software collection for developing typed feature structure grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE), Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/.

Müller, Stefan. 2007b. *Head-Driven Phrase Structure Grammar: Eine Einführung* (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag 1st edn. http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html.

Müller, Stefan. 2007c. Phrasal or lexical constructions: Some comments on underspecification of constituent order, compositionality, and control. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 373–393. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/phrasal-2.html.

Müller, Stefan. 2007d. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In Gisela Zifonun & Werner Kallmeyer (eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), 70–90. Berlin: Walter de Gruyter. http://hpsg.fu-berlin.de/~stefan/Pub/intro-korpus.html.

Müller, Stefan. 2008. Depictive secondary predicates in German and English. In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary predicates in Eastern European languages and beyond* (Studia Slavica Oldenburgensia 16), 255–273. Oldenburg: BIS-Verlag. http://hpsg.fu-berlin.de/~stefan/Pub/depiktiv-2006.html.

Müller, Stefan. 2009b. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics. papers from the 1st international conference on Maltese linguistics (Bremen/Germany, 18–20 october, 2007)* (Studies in Language Companion Series 113), 83–112. Amsterdam/Philadelphia: John Benjamins Publishing Co. http://hpsg.fu-berlin.de/~stefan/Pub/maltese-sketch.html.

Müller, Stefan. 2009c. On predication. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 213–233. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/predication.html.

Müller, Stefan. 2010a. *Grammatiktheorie* (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag. http://hpsg.fu-berlin.de/~stefan/Pub/grammatiktheorie.html.

Müller, Stefan. 2010b. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. http://hpsg.fu-berlin.de/~stefan/Pub/persian-cp.html.

Müller, Stefan. 2012. On the copula, specificational constructions and type shifting. Ms. Freie Universität Berlin. http://hpsg.fu-berlin.de/~stefan/Pub/copula.html.

Müller, Stefan. 2013a. The CoreGram project: A brief overview and motivation. In Denys Duchier & Yannick Parmentier (eds.), *Proceedings of the workshop on high-level methodologies for grammar engineering (HMGE 2013), Düsseldorf*, 93–104.

Müller, Stefan. 2013b. *Grammatiktheorie* (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag 2nd edn. http://hpsg.fu-berlin.de/~stefan/Pub/grammatiktheorie.html.

Müller, Stefan. 2013c. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. http://hpsg.fu-berlin.de/~stefan/Pub/unifying-everything.html.

Müller, Stefan. 2014b. Artenvielfalt und Head-Driven Phrase Structure Grammar. In *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28), 187–233. Tübingen: Stauffenburg. http://hpsg.fu-berlin.de/~stefan/Pub/artenvielfalt.html.

Müller, Stefan. 2014c. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/surface.html.

Müller, Stefan. 2014d. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung. In Andreas Nolda, Athina Sioupi & Antonio Machicao y Priemer (eds.), *Zwischen kern und peripherie* (studia grammatica 76), 25–39. Berlin/New York, NY: de Gruyter. http://hpsg.fu-berlin.de/~stefan/Pub/kernigkeit.html.

Müller, Stefan. 2014e. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse und Tendenzen), Berlin, Boston: de Gruyter. To appear.

Müller, Stefan. 2015a. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). http://hpsg.fu-berlin.de/~stefan/Pub/coregram.html. To appear.

Müller, Stefan. 2015b. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. http://hpsg.fu-berlin.de/~stefan/Pub/gs.html. In Preparation.

Müller, Stefan. 2015c. HPSG – a synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), chap. 27, 937–973. Berlin/New York, NY: Walter de Gruyter 2nd edn. http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-hsk.html.

Müller, Stefan. 2016. *Germanic syntax* (Lecture Notes in Language Sciences). Berlin: Language Science Press. In Preparation.

Müller, Stefan & Masood Ghayoomi. 2010. PerGram: A TRALE implementation of an HPSG fragment of Persian. In *Proceedings of 2010 ieee international multiconference on computer science and information technology – computational linguistics applications (cla'10). wisła, poland, 18–20 october 2010*, vol. 5, 461–467. Polnish Information Processing Society. http://hpsg.fu-berlin.de/~stefan/Pub/pergram.html.

Müller, Stefan & Walter Kasper. 2000. HPSG analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 238–253. Berlin/Heidelberg/New York, NY: Springer Verlag.

Müller, Stefan & Janna Lipenkova. 2009. Serial verb constructions in Chinese: An HPSG account. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 234–254. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/chinese-svc.html.

Müller, Stefan & Janna Lipenkova. 2013. Chingram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei, Taiwan: Department of English, National Chengchi University.

Müller, Stefan & Janna Lipenkova. 2015. *Mandarin Chinese in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.

Müller, Stefan & Bjarne Ørsnes. 2011. Positional expletives in Danish, German, and Yiddish. In Stefan Müller (ed.), *Proceedings of the 18th international conference on Head-Driven Phrase Structure Grammar, University of Washington, U.S.A.*, 167–187. Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/expletives.html.

Müller, Stefan & Bjarne Ørsnes. 2013a. Passive in Danish, English, and German. In Stefan Müller (ed.), *Proceedings of the 20th international conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 140–160. Stanford, CA: CSLI Publications.

Müller, Stefan & Bjarne Ørsnes. 2013b. Towards an HPSG analysis of object shift in Danish. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal grammar: 17th and 18th international conferences, FG 2012, Opole, Poland, August 2012, revised selected papers, FG 2013, Düsseldorf, Germany, August 2013. proceedings* (Lecture Notes in Computer Science 8036), 69–89. Berlin/Heidelberg/New York, NY: Springer Verlag. http://hpsg.fu-berlin.de/~stefan/Pub/object-shift.html.

Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. http://hpsg.fu-berlin.de/~stefan/Pub/danish.html. In Preparation.

Müller, Stefan, Pollet Samvelian & Olivier Bonami. In Preparation. *Persian in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. http://hpsg.fu-berlin.de/~stefan/Pub/persian.html.

Müller, Stefan & Stephen Mark Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. http://hpsg.fu-berlin.de/~stefan/Pub/arg-st.html.

Müller, Stefan & Stephen Mark 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. http://hpsg.fu-berlin.de/~stefan/Pub/arg-st-reply.html.

Muraki, Kazunori, Shunji Ichiyama & Yasutomo Fukumochi. 1985. Augmented Dependency Grammar: A Simple Interface between the Grammar Rule and the Knowledge. In Maghi King (ed.), *Proceedings of the 2nd european meeting of the association for computational linguistics*, 198–204. Geneva: Association for Computational Linguistics. http://aclweb.org/anthology/E/E85/.

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

Muysken, Peter. 1982a. Parameterizing the notion of 'head'. *Journal of Linguistic Research* 2. 57–75.

Muysken, Peter. 1982b. Parameterizing the notion of "head". *Journal of Linguistic Research* 2. 57–75.

Mykowiecka, Agnieszka, Małgorzata Marciniak, Adam Przepiórkowski & Anna Kupść. 2003. An implementation of a generative grammar of Polish. In Peter Kosta, Joanna

Błaszczak, Jens Frasek, Ljudmila Geist & Marzena Żygis (eds.), *Investigations into formal Slavic linguistics: Contributions of the fourth European conference on formal description of Slavic languages – FDSL IV held at Potsdam University, November 28–30, 2001*, 271–285. Frankfurt am Main: Peter Lang.

Naumann, Sven. 1987. Ein einfacher Parser für generalisierte Phrasenstrukturgrammatiken. *Zeitschrift für Sprachwissenschaft* 6(2). 206–226.

Naumann, Sven. 1988. *Generalisierte Phrasenstrukturgrammatik: Parsingstrategien, Regelorganisation und Unifikation* (Linguistische Arbeiten 212). Tübingen: Max Niemeyer Verlag.

Neeleman, Ad. 1994. *Complex predicates.* Utrecht: Onderzoeksinstituut voor Taal en Spraak (OTS) dissertation. http://www.phon.ucl.ac.uk/home/ad/pubs.html.

Nelimarkka, Esa, Harri Jäppinen & Aarno Lehtola. 1984. Two-way finite automata and Dependency Grammar: A parsing method fo inflectional free word order languages. In Yorick Wilks (ed.), *Proceedings of the 10th international conference on computational linguistics and 22nd annual meeting of the association for computational linguistics*, 389–392. Stanford University, California: Association for Computational Linguistics.

Nerbonne, John. 1986a. 'Phantoms' and German fronting: Poltergeist constituents? *Linguistics* 24(5). 857–870.

Nerbonne, John. 1986b. A phrase-structure grammar for German passives. *Linguistics* 24(5). 907–938.

Nerbonne, John. 1993. A feature-based syntax/semantics interface. *Annals of Mathematics and Artificial Intelligence* 8(1–2). 107–132. Special issue on Mathematics of Language edited by Alexis Manaster-Ramer and Wlodek Zadrozsny, selected from the 2nd Conference on Mathematics of Language. Also published as DFKI Research Report RR-92-42.

Nerbonne, John, Klaus Netter & Carl J. Pollard (eds.). 1994. *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46). Stanford, CA: CSLI Publications.

Netter, Klaus. 1991. Clause union phenomena and complex predicates in German. In Klaus Netter & Mike Reape (eds.), *Clause structure and word order variation in germanic* (DYANA Report, Deliverable R1.1.B), University of Edinburgh.

Netter, Klaus. 1992. On non-head non-movement. An HPSG treatment of finite verb position in German. In Günther Görz (ed.), *Konvens 92. 1. konferenz „verarbeitung natürlicher sprache". nürnberg 7.–9. oktober 1992* (Informatik aktuell), 218–227. Berlin/Heidelberg/New York, NY: Springer Verlag. http://www.dfki.de/lt/publications_show.php?id=420.

Netter, Klaus. 1993. Architecture and coverage of the disco grammar. In Stephan Busemann & Karin Harbusch (eds.), *Dfki workshop on natural language systems: Reusability and modularity, october23* (DFKI Document D-93-03), 1–10. Saarbrücken, Germany: DFKI.

Netter, Klaus. 1994. Towards a theory of functional heads: German nominal phrases. In Nerbonne et al. (1994) 297–340.

Netter, Klaus. 1996. *Functional categories in an HPSG for German.* Saarbrücken: Universität des Saarlandes Dissertation.

Netter, Klaus. 1998. *Functional categories in an HPSG for German* (Saarbrücken Dissertations in Computational Linguistics and Language Technology 3). Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes.

Neville, Anne & Patrizia Paggio. 2004. Developing a Danish grammar in the GRASP project: A construction-based approach to topology and extraction in Danish. In Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings of the joint meeting of the 6th conference on formal grammar and the 7th conference on mathematics of language* (Electronic Notes in Theoretical Computer Science 53), 246–259. Helsinki: Elsevier Science Publisher B.V. (North-Holland).

Nevins, Andrew Ira, David Pesetsky & Cilene Rodrigues. 2009. Pirahã exceptionality: a reassessment. *Language* 85(2). 355–404.

Newmeyer, Frederick J. 2004a. Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4. 181–234.

Newmeyer, Frederick J. 2004b. Typological evidence and Universal Grammar. *Studies in Language* 28(3). 527–548.

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

Newmeyer, Frederick J. 2010. On comparative concepts and descriptive categories: A reply to haspelmath. *Language* 86(3). 688–695.

Newport, Elissa L. 1990. Maturational constraints on language learning. *Cognitive Science* 14. 11–28. http://www.bcs.rochester.edu/people/newport/pdf/Newport_CogSci90.pdf.

Ng, Say Kiat. 1997. *A double-specifier account of Chinese NPs using Head-Driven Phrase Structure Grammar*. University of Edinburgh, Department of Linguistics MSc speech and language processing. http://home2.pacific.net.sg/~nskiat/ae.ps.

Nivre, Joakim. 2003. An efficient algorithm for projective dependency parsing. In Gertjan van Noord (ed.), *Proceedings of the 8th international workshop on parsing technologies (iwpt 03)*, Nancy.

Niyogi, Sourabh & Robert C. Berwick. 2005. A Minimalist implementation of Hale-Keyser incorporation theory. In Anna Maria Di Sciullo (ed.), *Ug and external systems: Language, brain and computation* (Linguistik Aktuell/Linguistics Today 75), 269–288. John Benjamins Publishing Co.

Nolda, Andreas. 2007. *Die Thema-Integration: Syntax und Semantik der gespaltenen Topikalisierung im Deutschen* (Studien zur deutschen Grammatik 72). Tübingen: Stauffenburg.

Noonan, Michael. 1994. A tale of two passives in Irish. In Barbara Fox & Paul J. Hopper (eds.), *Voice: Form and function* (Typological Studies in Language 27), 279–311. Amsterdam/Philadelphia: John Benjamins Publishing Co.

van Noord, Gertjan & Gosse Bouma. 1994. The scope of adjuncts and the processing of lexical rules. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 250–256. Kyoto, Japan: Association for Computational Linguistics. http://grid.let.rug.nl/~vannoord/papers/coling94.ps.gz.

Nordgård, Torbjørn. 1994. E-Parser: An implementation of a deterministic GB-related parsing system. *Computers and the Humanities* 28(4–5). 259–272.
Nordlinger, Rachel. 1998. *Constructive case: Evidence from australia* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.
Nowak, Martin A., Natalia L. Komarova & Partha Niyogi. 2001. Evolution of Universal Grammar. *Science* 291(5501). 114–118.
Nozohoor-Farshi, R. 1986. On formalizations of marcus' parser. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 533–535. University of Bonn: Association for Computational Linguistics.
Nozohoor-Farshi, R. 1987. Context-freeness of the language accepted by Marcus' parser. In Candy Sidner (ed.), *25th annual meeting of the association for computational linguistics*, 117–122. Stanford, CA: Association for Computational Linguistics.
Nunberg, Geoffrey. 1995. Transfers of meaning. *Journal of semantics* 12(2). 109–132.
Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.
Nunes, Jairo. 2004. *Linearization of chains and Sideward Movement* (Linguistic Inquiry Monographs 43). MIT Press.
Ochs, Elinor. 1982. Talking to children in Western Samoa. *Language and Society* 11(1). 77–104.
Ochs, Elinor & Bambi B. Schieffelin. 1984. Language acquisition and socialization: three developmental stories. In R. Schweder & R. Levin (eds.), *Culture theory: Essays in mind, self and emotion*, Cambridge, UK: Cambridge University Press.
O'Donovan, Ruth, Michael Burke, Aoife Cahill, Josef van Genabith & Andy Way. 2005. Large-scale induction and evaluation of lexical resources from the penn-ii and penn-iii treebanks. *Computational Linguistics* 31(3). 328–365.
Oepen, Stephan & Daniel P. Flickinger. 1998. Towards systematic grammar profiling. Test suite technology ten years after. *Journal of Computer Speech and Language* 12(4). 411–436. http://www.delph-in.net/itsdb/publications/profiling.ps.gz. (Special Issue on Evaluation).
Özkaragöz, İnci. 1986. Monoclausal double passives in Turkish. In Dan I. Slobin & Karl Zimmer (eds.), *Studies in turkish linguistics* (Typological Studies in Language 8), Amsterdam/Philadelphia: John Benjamins Publishing Co.
Oliva, Karel. 1992. Word order constraints in binary branching syntactic structures. CLAUS-Report 20 Universität des Saarlandes Saarbrücken.
Oliva, Karel. 2003. Dependency, valency and head-driven phrase-structure grammar. In Ágel et al. (2003) 660–668.
O'Neill, Michael & Randall Wood. 2012. The grammar of happiness. *Essential Media & Entertainment / Smithsonian Networks* https://www.youtube.com/watch?v=er2VAk4uXUs#t=1519.
Oppenrieder, Wilhelm. 1991. *Von Subjekten, Sätzen und Subjektsätzen* (Linguisitische Arbeiten 241). Tübingen: Max Niemeyer Verlag.
Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*: University of California, Berkeley dissertation.

Osborne, Miles & Ted J. Briscoe. 1997. Learning Stochastic Categorial Grammars. In T. Mark Ellison (ed.), *Conll97: Computational natural language learning. proceedings of the 1997 meeting of the acl special interest group in natural language learning*, 80–87. Madrid: Association for Computational Linguistics. http://www.aclweb.org/anthology-new/signll.html#1997-0.

Osborne, Timothy. 2016. Ellipsis in dependency grammar. In Jeroen van Craenenbrok (ed.), *Oxford handbook of ellipsis*, Oxford University Press. To appear.

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

Osborne, Timothy & Sylvain Kahane. 2015. Translators' introduction. In Tesnière (2015). Translated by Timothy Osborne and Sylvain Kahane.

Osborne, Timothy, Michael Putnam & Thomas Groß. 2012. Catenae: Introducing a novel unit of syntactic analysis. *Syntax* 15(4). 354–396.

Osborne, Timothy, Michael Putnam & Thomas M. Groß. 2011. Bare phrase structure, label-less trees, and specifier-less syntax. is minimalism becoming a dependency grammar? *The Linguistic Review* 28(3). 315–364.

Osenova, Petya. 2010a. Bulgarian resource grammar – efficient and realistic (burger). Tech. rep. LingoLab, CSLI Stanford. http://www.bultreebank.org/BURGER/BURGER3.pdf.

Osenova, Petya. 2010b. *Bulgarian Resource Grammar: Modeling Bulgarian in HPSG*. Saarbrücken: VDM Verlag Dr. Müller.

Osenova, Petya. 2011. Localizing a core hpsg-based grammar for bulgarian. In Hanna Hedeland, Thomas Schmidt & Kai Wörner (eds.), *Multilingual resources and multilingual applications: Proceedings of the conference of the german society for computational linguistics and language technology (GSCL) 2011* (Arbeiten zur Mehrsprachigkeit/Working Papers in Multilingualism, Folge B/Series B 96), 175–182. Universitat Hamburg. http://www.corpora.uni-hamburg.de/gscl2011/downloads/AZM96.pdf.

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

Ørsnes, Bjarne. 1995. *The derivation and compounding of complex event nominals in Modern Danish – an HPSG approach with an implementation in Prolog*: University of Copenhagen dissertation.

Ørsnes, Bjarne. 2002. Case marking and subject extraction in danish. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2002 conference*, 333–353. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/7/lfg02.html.

Ørsnes, Bjarne. 2009a. Das Verbalfeldmodell – ein Stellungsfeldermodell für den kontrastiven DaF-Unterricht. *Deutsch als Fremdsprache* 46(3). 143–149.

Ørsnes, Bjarne. 2009b. Preposed negation in Danish. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 255–275. Stanford, CA: CSLI Publications.

Ørsnes, Bjarne & Jürgen Wedekind. 2003. Paralelle datamatiske grammatikker for Norsk og Dansk [parallel computational grammars for Norwegian and Danish]. In Henrik Holmboe (ed.), *Årbog for nordisk sprogteknologisk forskningsprogram 2000–2004*, Kopenhagen: Museum Tusculanums Forlag.

Ørsnes, Bjarne & Jürgen Wedekind. 2004. Paralelle datamatiske grammatikker for Norsk og Dansk: Analyse og disambiguering af modalverber [parallel computational grammars for Norwegian and Danish: Analysis and disambiguation of modal verbs]. In Henrik Holmboe (ed.), *Årbog for nordisk sprogteknologisk forskningsprogram 2000–2004*, Kopenhagen: Museum Tusculanums Forlag.

Pafel, Jürgen. 1993. Ein Überblick über die Extraktion aus Nominalphrasen im Deutschen. In Franz-Josef d'Avis, Sigrid Beck, Uli Lutz, Jürgen Pafel & Susanne Trissler (eds.), *Extraktion im Deutschen I* (Arbeitspapiere des SFB 340 No. 34), 191–245. Eberhard-Karls-Universität Tübingen.

Paggio, Patrizia. 2005. Representing information structure in a formal grammar of Danish. In *Proceedings of the 2nd international workshop on logic and engineering of natural language semantics (lenls2005). kitakyushu, japan. june 13–14*, http://cst.dk/patrizia/paggio/.

Parmentier, Yannick, Laura Kallmeyer, Wolfgang Maier, Timm Lichte & Johannes Dellert. 2008. TuLiPA: A syntax-semantics parsing environment for mildly context-sensitive formalisms. In *Proceedings of the ninth international workshop on tree adjoining grammars and related formalisms (tag+9)*, 121–128. Tübingen. http://www.sfs.uni-tuebingen.de/~lk/papers/tag+9-parmentier-et-al.pdf.

Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen A. G. Groenendijk, Dick de Jongh & Martin J. B. Stokhof (eds.), *Studies in discourse representation theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris Publications.

Patejuk, Agnieszka & Adam Przepiórkowski. 2012. Towards an LFG parser for Polish: An exercise in parasitic grammar development. In *Proceedings of the eighth international Conference on Language Resources and Evaluation, LREC 2012*, 3849–3852. Istanbul, Turkey: ELRA.

Paul, Hermann. 1919. *Deutsche Grammatik. Teil IV: Syntax*, vol. 3. Halle an der Saale: Max Niemeyer Verlag. 2nd unchanged edition 1968, Tübingen: Max Niemeyer Verlag.

Paul, Soma. 2004. *An hpsg account of Bangla compound verbs with lkb implementation*. Hyderabad, India: CALTS, University of Hyderabad, India dissertation.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In Donia Scott (ed.), *Proceedings of the 42nd meeting of the association for computational linguistics (acl'04), main volume*, 239–246. Barcelona, Spain.

Penn, Gerald & Bob Carpenter. 1999. ALE for speech: a translation prototype. In Géza Gordos (ed.), *Proceedings of the 6th conference on speech communication and technology (eurospeech)*, Budapest, Hungary.

Perlmutter, David M. 1978. Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the 4th annual meeting of the berkeley linguistics society*, 157–189. Berkeley Linguistic Society.

Perlmutter, David M. (ed.). 1983. *Studies in relational grammar*, vol. 1. The University of Chicago Press.

Perlmutter, David M. (ed.). 1984. *Studies in relational grammar*, vol. 2. The University of Chicago Press.

Perlmutter, David M. & John Robert Ross. 1970. Relative clauses with split antecedents. *Linguistic Inquiry* 1(3). 350.

Pesetsky, David. 1996. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Peters, Stanley & R. W. Ritchie. 1973. On the generative power of Transformational Grammar. *Information Sciences* 6(C). 49–83.

Petrick, Stanley Roy. 1965. *A recognition procedure for transformational grammars*: Massachusetts Institute of Technology. Dept. of Modern Languages dissertation. http://hdl.handle.net/1721.1/13013.

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

Phillips, John D. 1992. A computational representation for generalised phrase structure grammars. *Linguistics and Philosophy* 15(3). 255–287.

Phillips, John D. & Henry S. Thompson. 1985. Gpsgp – a parser for Generalized Phrase Structure Grammar. *Linguistics* 23(2). 245–261.

Piattelli-Palmarini, Massimo (ed.). 1980. *Language and learning: The debate between jean piaget and noam chomsky*. Cambridge: Harvard University Press.

Pickering, Martin & Guy Barry. 1993. Dependency Categorial Grammar and coordination. *Linguistics* 31(5). 855–902.

Pienemann, Manfred. 2005. An introduction to Processability Theory. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processablity theory*, 1–60. John Benjamins Publishing Co.

Piñango, Maria Mercedes, Jennifer Mack & Ray S. Jackendoff. 2006. Semantic combinatorial processes in argument structure: Evidence from light-verbs. In *Proceedings of the 32nd annual meeting of the berkeley linguistics society: Theoretical approaches to argument structure*, vol. 32, Berkeley, CA: BLS.

Pineda, Luis Alberto & Iván V. Meza. 2005a. A computational model of the spanish clitic system. In Alexander Gelbkuh (ed.), *Computational linguistics and intelligent language processing*, 73–82. Berlin/Heidelberg/New York, NY: Springer Verlag.

Pineda, Luis Alberto & Iván V. Meza. 2005b. The Spanish pronominal clitic system. *Procesamiento del Lenguaje Natural* 34. 67–103. http://www.sepln.org/revistaSEPLN/revista/34/06.pdf.

Pinker, Steven. 1984. *Learnability and cognition: the acquisition of argument structure*. London/Cambridge, MA: MIT Press.

Pinker, Steven. 1994. *The language instinct. how the mind creates language*. New York: William Morrow.

Pinker, Steven & Ray S. Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95(2). 201–236.

Pittner, Karin. 1995. Regeln für die Bildung von freien Relativsätzen. Eine Antwort an Oddleif Leirbukt. *Deutsch als Fremdsprache* 32(4). 195–200.

Plank, Frans & Elena Filimonova. 2000. The universals archive: A brief introduction for prospective users. *Sprachtypologie und Universalienforschung* 53(1). 109–123.

Poletto, Cecilia. 2000. *The higher functional field: Evidence from Northern Italian Dialects.* Oxford: Oxford University Press.

Pollard, Carl J. 1984. *Generalized phrase structure grammars, head grammars, and natural language*: Stanford University dissertation.

Pollard, Carl J. 1988. Categorial grammar and phrase structure grammar: An excursion on the syntax-semantics frontier. In Richard Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial grammars and natural language structures*, 391–415. Dordrecht/Boston/London: D. Reidel Publishing Company.

Pollard, Carl J. 1996a. The nature of constraint-based grammar. http://lingo.stanford.edu/sag/L221a/pollard-96.txt.

Pollard, Carl J. 1996b. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin/New York, NY: Mouton de Gruyter. Veröffentlichte Version eines Ms. von 1990.

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

Pollard, Carl J. & Andrew M. Moshier. 1990. Unifying partial descriptions of sets. In Philip P. Hanson (ed.), *Information, language and cognition* (Vancouver Studies in Cognitive Science 1), 285–322. Vancouver: University of British Columbia Press.

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

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

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

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Popowich, Fred & Carl Vogel. 1991. A logic based implementation of Head-Driven Phrase Structure Grammar. In Charles Grant Brown & Gregers Koch (eds.), *Natural language understanding and logic programming, iii. the 3rd internat. workshop, stockholm, sweden, 23–25 jan., 1991*, 227–246. Amsterdam: Elsevier, North-Holland.

Porzel, Robert, Vanessa Micelli, Hidir Aras & Hans-Peter Zorn. 2006. Tying the knot: Ground entities, descriptions and information objects for construction-based information extraction. In *Proceedings of the ontolex workshop at lrec, may 2006. genoa, italy*, 35–40.

Postal, Paul M. 2004. *Skeptical linguistic essays.* Oxford University Press. http://www.nyu.edu/gsas/dept/lingu/people/faculty/postal/papers/skeptical/.

Postal, Paul M. 2009. The incoherence of chomsky's 'biolinguistic' ontology. *Biolinguistics* 3(1). 104–123.

Postal, Paul M. & Geoffrey K. Pullum. 1986. Misgovernment. *Linguistic Inquiry* 17(1). 104–110.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. RuCCS Technical Report 2 Center for Cognitive Science, Rutgers University, Piscataway, N.J., and Computer Science Department, University of Colorado, Boulder. http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF.

Przepiórkowski, Adam. 1999a. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*: Eberhard-Karls-Universität Tübingen dissertation. http://w210.ub.uni-tuebingen.de/dbt/volltexte/1999/72/.

Przepiórkowski, Adam. 1999b. On case assignment and “adjuncts as complements”. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford, CA: CSLI Publications.

Przepiórkowski, Adam & Anna Kupść. 2006. HPSG for Slavicists. *Glossos* 8. 1–68. http://www.seelrc.org/glossos/issues/8/.

Przepiórkowski, Adam, Anna Kupść, Małgorzata Marciniak & Agnieszka Mykowiecka. 2002. *Formalny opis języka polskiego: Teoria i implementacja*. Warsaw: Akademicka Oficyna Wydawnicza EXIT.

Pullum, Geoffrey K. 1977. Word order universals and grammatical relations. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations* (Syntax and Semantics 8), 249–277. New York, San Francisco, London: Academic Press.

Pullum, Geoffrey K. 1982. Free word order and phrase structure rules. In James Pustejovsky & Peter Sells (eds.), *Proceedings of the 12th anual meeting of the northeast linguistic society*, 209–220. Amherst: Graduate Linguistics Student Association.

Pullum, Geoffrey K. 1983. How many possible human languages are there? *Linguistic Inquiry* 14(3). 447–467.

Pullum, Geoffrey K. 1985. Assuming some version of X-bar Theory. In *Papers from the 21st Annual Meeting of the Chicago Linguistic Society*, 323–353.

Pullum, Geoffrey K. 1986. Footloose and context-free. *Natural Language and Linguistic Theory* 4(3). 409–414.

Pullum, Geoffrey K. 1988. Citation etiquette beyond thunderdome. *Natural Language and Linguistic Theory* 6(4). 579–588.

Pullum, Geoffrey K. 1989a. Formal linguistics meets the boojum. *Natural Language and Linguistic Theory* 7(1). 137–143. http://dx.doi.org/10.1007/BF00141350.

Pullum, Geoffrey K. 1989b. The incident of the node vortex problem. *Natural Language and Linguistic Theory* 7(3). 473–479.

Pullum, Geoffrey K. 1991. *The great eskimo vocabulary hoax and other irreverent essays on the study of language*. Chicago, IL: The University of Chicago Press.

Pullum, Geoffrey K. 1996. Learnability, hyperlearning, and the poverty of the stimulus. In J. Johnson, M.L. Juge & J.L. Moxley (eds.), *Proceedings of the 22nd annual meeting of the berkeley linguistics society: General session and parasession on the role of learnability in grammatical theory*, 498–513. Berkeley, CA: Berkeley Linguistic Society. http://users.ecs.soton.ac.uk/harnad/Papers/Py104/pullum.learn.html.

Pullum, Geoffrey K. 2003. Learnability: Mathematical aspects. In William J. Frawley (ed.), *Oxford international encyclopedia of linguistics*, 431–434. Oxford: Oxford University

Press 2nd edn. http://www.kornai.com/MatLing/.

Pullum, Geoffrey K. 2007. The evolution of model-theoretic frameworks in linguistics. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10 – proceedings of the esslli 2007 mts@10 workshop, august 13–17*, 1–10. Dublin: Trinity College Dublin. http://cs.earlham.edu/esslli07mts/.

Pullum, Geoffrey K. 2009. Response to Anderson. *Language* 85(2). 245–247.

Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind and Language* 28(4). 492–521. http://www.lel.ed.ac.uk/~gpullum/Metatheory.pdf.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between generative-enumerative and model-theoretic syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor (eds.), *Logical aspects of computational linguistics: 4th international conference* (Lecture Notes in Computer Science 2099), 17–43. Berlin/Heidelberg/New York, NY: Springer Verlag.

Pullum, Geoffrey K. & Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19(1–2). 9–50.

Pullum, Geoffrey K. & Barbara C. Scholz. 2010. Recursion and the infinitude claim. In Harry van der Hulst (ed.), *Recursion in human language* (Studies in Generative Grammar 104), 113–138. Berlin/New York, NY: Mouton de Gruyter. http://ling.ed.ac.uk/~gpullum/bcscholz/Infinitude.pdf.

Pulman, Stephen G. 1985. A parser that doesn't. In Maghi King (ed.), *Proceedings of the 2nd european meeting of the association for computational linguistics*, 128–135. Geneva: Association for Computational Linguistics. http://aclweb.org/anthology/E/E85/.

Pulvermüller, Friedemann. 2003. *The neuroscience of language: On brain circuits of words and serial order*. Cambridge, UK: Cambridge University Press.

Pulvermüller, Friedemann. 2010. Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain & Language* 112(3). 167–179.

Pulvermüller, Friedemann, Bert Cappelle & Yury Shtyrov. 2013. Brain basis of meaning, words, constructions, and grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar* (Oxford Handbooks), Oxford: Oxford University Press.

Quaglia, Stefano. 2014. On the syntax of some apparent spatial particles in Italian. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 503–523. Stanford, CA: CSLI Publications.

Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax*. Cambridge, MA: Blackwell Publishing Ltd.

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

Rákosi, György, Tibor Laczkó & Gábor Csernyi. 2011. On english phrasal verbs and their Hungarian counterparts: from the perspective of a computational linguistic project. *Argumentum* 7(80–89).

Rambow, Owen. 1994. *Formal and computational aspects of natural language syntax*: University of Pennsylvania dissertation.

Ramchand, Gillian. 2005. Post-davidsonianism. *Theoretical Linguistics* 31(3). 359–373.

Randriamasimanana, Charles. 2006. Simple sentences in malagasy. In Henry Y. Chang, Lillian M. Huang & Dah ah Ho (eds.), *Streams converging into an ocean: Festschrift in honor of professor paul jen-kuei li on his 70th birthday*, Taipei, Taiwan: Institute of Linguistics, Academia Sinica.

Raposo, Eduardo & Juan Uriagereka. 1990. Long-distance case assignment. *Linguistic Inquiry* 21(4). 505–537.

Rappaport, Malka. 1983. On the nature of derived nominals. In Lori S. Levin, Malka Rappaport & Annie Zaenen (eds.), *Papers in Lexical Functional Grammar*, 113–42. Indiana University Linguistics Club.

Rauh, Gisa. 2013. Linguistic categories and the syntax-semantics interface: Evaluating competing approaches. Ms. Universität Wuppertal.

Reape, Mike. 1991. Word order variation in Germanic and parsing. DYANA Report Deliverable R1.1.C University of Edinburgh.

Reape, Mike. 1992. *A formal theory of word order: A case study in West Germanic*: University of Edinburgh dissertation.

Reape, Mike. 1994. Domain union and word order variation in German. In Nerbonne et al. (1994) 151–198.

Reis, Marga. 1974. Syntaktische Hauptsatzprivilegien und das Problem der deutschen Wortstellung. *Zeitschrift für Germanistische Linguistik* 2(3). 299–327.

Reis, Marga. 1980. On justifying topological frames: 'positional field' and the order of nonverbal constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine* 22/23. 59–85.

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im deutschen – vorschläge zur syntaktischen, semantischen und pragmatischen fundierung* (Studien zur deutschen Grammatik 15), 171–211. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Remberger, Eva-Maria. 2009. Null subjects, expletives and locatives in Sardinian. In Georg A. Kaiser & Eva-Maria Remberger (eds.), *Proceedings of the workshop* Null-subjects, expletives, and locatives in Romance (Arbeitspapier 123), 231–261. Fachbereich Sprachwissenschaft, Universität Konstanz.

Resnik, Philip. 1992. Probabilistic Tree-Adjoining Grammar as a framework for statistical natural language processing. In Antonio Zampoili (ed.), *14th international conference on computational linguistics (COLING '92), august23–28*, 418–424. Nantes, France: Association for Computational Linguistics.

Reyle, Uwe. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Joual of Semantics* 10(2). 123–179.

Richards, Marc. 2015a. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science), 803–839. Berlin/New York, NY: Mouton de Gruyter 2nd edn.

Richards, Marc. 2015b. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science), Berlin/New York, NY: Mouton de Gruyter 2nd edn. Entwurf vom 02.12.2009.

Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-Driven Phrase Structure Grammar*: Eberhard-Karls-Universität Tübingen Phil. Dissertation (2000). http://w210.ub.uni-tuebingen.de/dbt/volltexte/2004/1203/.

Richter, Frank. 2007. Closer to the truth: A new model theory for HPSG. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10 – proceedings of the esslli 2007 mts@10 workshop, august 13–17*, 101–110. Dublin: Trinity College Dublin. http://cs.earlham.edu/esslli07mts/.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of lexical resource semantics. In Arnold Beckmann & Norbert Preining (eds.), *Esslli 2003 – course material i* (Collegium Logicum 5), 87–143. Kurt Gödel Society Wien.

Richter, Frank & Manfred Sailer. 2009. Phraseological clauses as constructions in HPSG. In Stefan Müller (ed.), *Proceedings of the 16th international conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 297–317. Stanford, CA: CSLI Publications.

Riehemann, Susanne. 1993. *Word formation in lexical type hierarchies: A case study of* bar*-adjectives in german*. Eberhard-Karls-Universität Tübingen MA thesis. Also published as SfS-Report-02-93, Seminar für Sprachwissenschaft, University of Tübingen.

Riehemann, Susanne Z. 1998. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2(1). 49–77. http://doors.stanford.edu/~sr/morphology.ps.

Riezler, Stefan, Tracy Holloway King, Ronald M. Kaplan, Richard Crouch, John T. Maxwell III & Mark Johnson. 2002. Parsing the Wall Street Journal using a Lexical-Functional Grammar and discriminative estimation techniques. In Pierre Isabelle (ed.), *40th annual meeting of the association for computational linguistics. proceedings of the conference*, 271–278. University of Pennsylvania, Philadelphia: Association for Computational Linguistics. http://www.aclweb.org/.

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

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17(3).

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

Rizzi, Luigi. 2009a. The discovery of language invariance and variation, and its relevance for the cognitive sciences. *The Behavioral and Brain Sciences* 32(5). 467–468.

Rizzi, Luigi. 2009b. Language variation and universals: Some notes on n. evans and s. c. levinson (2009) 'the myth of language universals: Language diversity and its importance for cognitive science'. In Paola Cotticelli-Kurras & Alessandra Tomaselli (eds.), *La grammatica tra storia e teoria. studi in onore di giorgio graffi*, Alessandra: Edizioni dell'Orso.

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

Rogers, James. 1994. Obtaining trees from their descriptions: An application to Tree-Adjoining Grammars. *Computational Intelligence* 10(4). 401–421.

Rogers, James. 1997. 'Grammarless' phrase structure grammar. *Linguistics and Philosophy* 20. 721–746.

Rogers, James. 1998. *A descriptive approach to language-theoretic complexity* (Studies in Logic, Language and Information). Stanford, CA: CSLI Publications.

Rohrer, Christian. 1996. Fakultativ kohärente Infinitkonstruktionen im Deutschen und deren Behandlung in der Lexikalisch Funktionalen Grammatik. In Gisela Harras & Manfred Bierwisch (eds.), *Wenn die semantik arbeitet. klaus baumgärtner zum 65. geburtstag*, 89–108. Tübingen: Max Niemeyer Verlag.

Rohrer, Christian & Martin Forst. 2006. Improving coverage and parsing quality of a large-scale LFG for German. In *Proceedings of the language resources and evaluation conference (lrec-2006)*, Genoa, Italy.

Ross, John Robert. 1967. *Constraints on variables in syntax*: MIT dissertation. http://www.eric.ed.gov/. Reproduced by the Indiana University Linguistics Club.

Ross, John Robert. 1986. *Infinite syntax!* Norwood, New Jersey: Ablex Publishing Corporation.

Rothkegel, Annely. 1976. *Valenzgrammatik* (Linguistische Arbeiten 19). Sonderforschungsbereich Elektronische Sprachforschung, Universität des Saarlandes.

Sabel, Joachim. 1999. Das Passiv im Deutschen. Derivationale Ökonomie vs. optionale Bewegung. *Linguistische Berichte* 177. 87–112.

Sáfár, Éva & John Glauert. 2010. Sign Language HPSG. In *Proceedings of the 4th workshop on the representation and processing of sign languages: Corpora and sign language technologies, lrec 2010, 22–23 may 2010, malta*, 204–207.

Sáfár, Éva & Ian Marshall. 2002. Sign language translation via DRT and HPSG. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent text processing: Third international conference, cicling 2002 mexico city, mexico, february 17–23, 2002 proceedings* (Lecture Notes in Computer Science 2276), 58–68. Berlin/Heidelberg/New York, NY: Springer Verlag.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484. http://lingo.stanford.edu/sag/papers/rel-pap.pdf.

Sag, Ivan A. 2000. Another argument against *Wh*-trace. Jorge Hankamer Webfest. http://ling.ucsc.edu/Jorge/sag.html.

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545. http://lingo.stanford.edu/sag/papers/xcons.pdf.

Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications. http://lingo.stanford.edu/sag/papers/theo-syno.pdf.

Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing sign-based construction grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based construction grammar* (CSLI Lecture Notes 193), 1–29. Stanford, CA: CSLI Publications.

Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: The complex noun phrase constraint. In Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki & Suwon Yoon (eds.), *Proceedings of the 43rd annual meeting of the chicago linguistic society*, 215–229. Chicago, IL: Chicago Linguistic Society.

Sag, Ivan A. & Carl J. Pollard. 1991. An integrated theory of complement control. *Language* 67(1). 63–113.

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

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

Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*: Eberhard-Karls-Universität Tübingen Dissertation. http://w210.ub.uni-tuebingen.de/dbt/volltexte/2003/916/.

Samarin, William J. 1984. Socioprogrammed linguistics. *The Behavioral and Brain Sciences* 7(2). 206–207.

Sampson, Geoffrey. 1989. Language acquisition: Growth or learning? *Philosophical Papers* 18(3). 203–240.

Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43. 605–645.

Sarkar, Anoop & Aravind K. Joshi. 1996. Coordination in Tree Adjoining Grammars: Formalization and implementation. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 610–615. Copenhagen, Denmark: Association for Computational Linguistics.

Sato, Yo. 2008. *Implementing head-Driven linearisation grammar*: King's College London dissertation.

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

Savin, Harris B. & Ellen Perchonock. 1965. Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 4(5). 348–353.

Schein, Barry. 1993. *Plurals and events* (Current Studies in Linguistics 23). Cambridge, MA: MIT Press.

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.

Schluter, Natalie & Josef van Genabith. 2009. Dependency parsing resources for french: Converting acquired lexical functional grammar f-structure annotations and parsing f-structures directly. In Kristiina Jokinen & Eckhard Bick (eds.), *Nodalida 2009 conference proceedings*, 166–173.

Schmidt, Paul, Sibylle Rieder & Axel Theofilidis. 1996a. Final documentation of the German LS-GRAM lingware. Deliverable DC-WP6e (German) IAI Saarbrücken.

Schmidt, Paul, Axel Theofilidis, Sibylle Rieder & Thierry Declerck. 1996b. Lean formalisms, linguistic theory, and applications: Grammar development in ALEP. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96. 16th international conference on computational linguistics (COLING96). copenhagen, denmark, august 5–9, 1996*, 286–291. Copenhagen, Denmark: Association for Computational Linguistics. doi:http://dx.doi.org/10.3115/992628.992679.

Scholz, Barbara C. & Geoffrey K. Pullum. 2002. Searching for arguments to support linguistic nativism. *The Linguistic Review* 19(1–2). 185–223.

Schubert, K. 1987. *Metataxis: Contrastive dependency syntax for machine translation.* Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

Schumacher, Helmut, Jacqueline Kubczak, Renate Schmidt & Vera de Ruiter. 2004. *VALBU – Valenzwörterbuch deutscher Verben.* Tübingen: Gunter Narr Verlag.

Schütz, Jörg. 1996. The ALEP formalism in a nutshell. http://www.iai.uni-sb.de/docs/alep-nutshell.pdf.

Seiss, Melanie & Rachel Nordlinger. 2012. An electronic dictionary and translation system for murrinh-patha. *The EUROCALL Review: Proceedings of the EUROCALL 2011 Conference* 20(1). 135–138.

Sengupta, Probal & B. B. Chaudhuri. 1997. A delayed syntactic-encoding-based LFG parsing strategy for an Indian language—Bangla. *Computational Linguistics* 23(2). 345–351.

Seuren, Pieter A. M. 1984. The Bioprogram Hypothesis: Facts and fancy. *The Behavioral and Brain Sciences* 7(2). 208–209.

Seuren, Pieter A. M. 2004. *Chomsky's Minimalism.* Oxford: Oxford University Press.

Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343.

Shieber, Stuart M. 1986. *An introduction to unification-based approaches to grammar* (CSLI Lecture Notes 4). Stanford, CA: CSLI Publications.

Shieber, Stuart M. & Mark Johnson. 1993. Variations on incremental interpretation. *Journal of Psycholinguistic Research* 22(2). 287–318.

Shieber, Stuart M., Hans Uszkoreit, Fernando Pereira, Jane Robinson & Mabry Tyson. 1983. The formalism and implementation of PATR-II. In *Research on interactive acquisition and use of knowledge*, 39–79. Menlo Park, CA: Artificial Intelligence Center, SRI International.

Shtyrov, Y., E. Pihko & F. Pulvermüller. 2005. Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? *Neuroimage* 27(1). 37–47.

Siegel, Melanie. 2000. Hpsg analysis of Japanese. In Wahlster (2000) 264–279.

Siegel, Melanie & Emily M. Bender. 2002. Efficient deep processing of Japanese. In *Proceedings of the 3rd workshop on asian language resources and international standardization at the 19th international conference on computational linguistics. taipei, taiwan*, http://www.aclweb.org/anthology-new/W/W02/W02-1210.pdf.

Simov, Kiril, Petya Osenova, Alexander Simov & Milen Kouylekov. 2004. Design and implementation of the bulgarian hpsg-based treebank. *Research on Language and Computation* 2(4). 495–522.

Simpson, Jane. 1983. Resultatives. In Lori S. Levin, Malka Rappaport & Annie Zaenen (eds.), *Papers in Lexical Functional Grammar*, Indiana University Linguistics Club. Reprint: Simpson (2005b).

Simpson, Jane. 2005a. Depictives in English and Warlpiri. In Nikolaus P. Himmelmann & Eva Schultze-Berndt (eds.), *Secondary predication and adverbial modification. the typology of depictives*, 69–106. Oxford: Oxford University Press.

Simpson, Jane. 2005b. Resultatives. In Miriam Butt & Tracy Holloway King (eds.), *Lexical semantics in LFG*, 149–161. Stanford, CA: CSLI Publications.

Singleton, Jenny L. & Elissa L. Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407.

Smith, Carlota S. 1970. Jespersen's "move and change" class and causative verbs in English. In Bert Peeters (ed.), *The lexicon–encyclopedia interface*, 101–109. Amsterdam: Elsevier.

Smith, Carlota S. 1972. On causative verbs and derived nominals in English. *Linguistic Inquiry* 3(1). 136–138.

Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77(2). 324–342.

Son, Minjeong. 2007. Directionality and resultativity: The cross-linguistic correlation revisited. *Tromsø University Working Papers on Language & Linguistics* 34. 126–164. http://hdl.handle.net/10037/3191.

Son, Minjeong & Peter Svenonius. 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. In Natasha Abner & Jason Bishop (eds.), *Proceedings of the 27th west coast conference on formal linguistics*, 388–396. Somerville, MA: Cascadilla Proceedings Project.

Song, Sanghoun, Jong-Bok Kim, Francis Bond & Jaehyung Yang. 2010. Development of the Korean Resource Grammar: Towards grammar customization. In *Proceedings of the 8th workshop on asian language resources, beijing, china, 21–22 august 2010*, 144–152. Asian Federation for Natural Language Processing.

Sorace, Antonella. 2003. Near-nativeness. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition* (Blackwell Handbooks in Linguistics), 130–151. Blackwell Publishing Ltd.

Stabler, Edward P. 1987. Restricting logic grammars with Government-Binding Theory. *Computational Linguistics* 13(1–2). 1–10.

Stabler, Edward P. 1991. Avoid the pedestrian's paradox. In Robert C. Berwick, Steven P. Abney & Carol Tenny (eds.), *Principle-based parsing: Computation and psycholinguistics*, 199–237. Dordrecht: Kluwer Academic Publishers.

Stabler, Edward P. 1992. *The logical approach to syntax: Foundations, specifications, and implementations of theories of government and binding* (ACL-MIT Press Series in Natural Language Processing). Cambridge, MA: MIT Press.

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. 2010. After Governement and Binding Theory. In Johan F. A. K. van Benthem & G. B. Alice ter Meulen (eds.), *Handbook of logic and language*, 395–414. Cambridge, MA: MIT Press 2nd edn. http://www.linguistics.ucla.edu/people/stabler/afterGB.pdf.

Stabler, Edward P. 2011. Computational perspectives on Minimalism. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic Minimalism* (Oxford Handbooks in Linguistics), chap. 27, 616–641. Oxford: Oxford University Press. http://www.linguistics.ucla.edu/people/stabler/Stabler10-Min.pdf.

Starosta, Stanley & Hirosato Nomura. 1986. Lexicase parsing: A lexicon-driven approach to syntactic analysis. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 127–132. University of Bonn: Association for Computational Linguistics. doi:10.3115/991365.991400. http://dx.doi.org/10.3115/991365.991400.

von Stechow, Arnim. 1979. Deutsche Wortstellung und Montague-Grammatik. In Jürgen M. Meisel & Martin D. Pam (eds.), *Linear order and generative theory*, 317–490. Amsterdam/Philadelphia: John Benjamins Publishing Co.

von Stechow, Arnim. 1989. Distinguo. Eine Antwort auf Dieter Wunderlich. *Linguistische Berichte* 122. 330–339.

von Stechow, Arnim. 1996. The different readings of *wieder* "again": A structural account. *Journal of Semantics* 13(2). 87–138.

von Stechow, Arnim & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen wissens. ein lehrbuch der generativen grammatik*. Opladen/Wiesbaden: Westdeutscher Verlag.

Steedman, Mark J. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61(3). 523–568.

Steedman, Mark J. 1989. Constituency and coordination in a combinatory grammar. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 201–231. Chicago/London: The University of Chicago Press.

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

Steedman, Mark J. 1997. *Surface structure and interpretation* (Linguistic Inquiry Monographs 30). Cambridge, MA/London, England: MIT Press.

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

Steedman, Mark J. 2011. Romantics and revolutionaries. *Linguistic Issues in Language Technology* 6(11). http://elanguage.net/journals/index.php/lilt/issue/view/330. Special Issue on Interaction of Linguistics and Computational Linguistics.

Steedman, Mark J. & Jason Baldridge. 2006. Combinatory Categorial Grammar. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 610–621. Oxford: Elsevier 2nd edn.

Steels, Luc. 2003. Evolving grounded communication for robots. *Trends in Cognitive Science* 7(7). 308–312.

Steels, Luc (ed.). 2011. *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11). Amsterdam/Philadelphia: John Benjamins Publishing Co.

Steels, Luc (ed.). 2012. *Computational issues in Fluid Construction Grammar* (Lecture Notes in Computer Science 7249). Berlin/Heidelberg/New York, NY: Springer Verlag.

Steels, Luc. 2013. Fluid Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar* (Oxford Handbooks), Oxford: Oxford University Press.

Steels, Luc & Joachim De Beule. 2006. A (very) brief introduction to Fluid Construction Grammar. Paper presented at the Third International Workshop on Scalable Natural Language Understanding (ScaNaLU 2006) June 8, 2006, following HLT/NAACL, New York City. http://arti.vub.ac.be/~joachim/acl-ny-06-3.pdf.

Steels, Luc & Remi van Trijp. 2011. How to make construction grammars fluid and robust. In Steels (2011) 301–330.

Stefanowitsch, Anatol. 2008. Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19(3). 513–531.

Stefanowitsch, Anatol & Kerstin Fischer (eds.). 2008. *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik* (Stauffenburg Linguistik 47). Tübingen: Stauffenburg Verlag.

Stefanowitsch, Anatol & Stephan Th. Gries. 2009. Corpora and grammar. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics. an international handbook*, vol. 29.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), chap. 43, 933–952. Berlin: Mouton de Gruyter.

Steinitz, Renate. 1969. *Adverbial-Syntax* (studia grammatica X). Berlin: Akademie Verlag. Unter Mitarbeit von Ewald Lang.

Sternefeld, Wolfgang. 1985a. Deutsch ohne grammatische Funktionen: Ein Beitrag zur Rektions- und Bindungstheorie. *Linguistische Berichte* 99. 394–439.

Sternefeld, Wolfgang. 1985b. On case and binding theory. In Jindřich Toman (ed.), *Studies in german grammar* (Studies in Generative Grammar 21), 231–285. Dordrecht: Holland, Cinnaminson: U.S.A.: Foris Publications.

Sternefeld, Wolfgang. 1995. Voice phrases and their specifiers. *FAS Papers in Linguistics* 3. 48–85.

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

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

Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag verbaler Präfixe und Partikeln* (studia grammatica 39). Berlin: Akademie Verlag.

Stowell, Timothy. 1981. *Origins of phrase structure*: MIT dissertation. http://hdl.handle.net/1721.1/15626.

Suchsland, Peter. 1997. *Syntax-Theorie. Ein zusammengefaßter Zugang* (Konzepte der Sprach- und Literaturwissenschaft 55). Tübingen: Max Niemeyer Verlag. Deutsche Bearbeitung von Borsley (1991) durch Peter Suchsland.

Sulger, Sebastian. 2009. Irish clefting and information-structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2009 conference*, 562–582. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/14/.

Sulger, Sebastian. 2010. Analytic and synthetic verb forms in irish – an agreement-based implementation in LFG. In Manfred Pinkal, Ines Rehbein, Sabine Schulte im Walde & Angelika Storrer (eds.), *Semantic approaches in natural language processing: Proceedings of the conference on natural language processing 2010*, 169–173. Saarbrücken: Saarland University Press (universaar).

Svenononius, Peter. 2004. Slavic prefixes inside and outside VP. *Nordlyd. Special Issue on Slavic Prefixes* 32(2). 205–253.

Takami, Ken-ichi. 1988. Preposition stranding: Arguments against syntactic analyses and an alternative functional explanation. *Lingua* 76(4). 299–335.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634. http://www.bcs.rochester.edu/people/mtan/publications/1995Tanenhaus_Sci.pdf.

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.

ten Hacken, Pius. 2007. *Chomskyan linguistics and its competitors*. London: Equinox Publishing Ltd.

Tesnière, Lucien. 1959. *Eléments de syntaxe structurale*. Paris: Librairie C. Klincksieck.

Tesnière, Lucien. 1980. *Grundzüge der strukturalen Syntax*. Stuttgart: Klett-Cotta. Translated by Ulrich Engel.

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

Thiersch, Craig L. 1978. *Topics in German syntax*: M.I.T. Dissertation.

Thompson, Henry S. 1982. Handling metarules in a parser for GPSG. D.A.I. Research 175 University of Edinburgh.

Timberlake, Alan. 1982. The impersonal passive in Lithuanian. In *Proceedings of the eighth annual meeting of the berkeley linguistics society*, 508–524. Berkeley: University of California.

Tomasello, Michael. 1995. Language is not an instinct. *Cognitive Development* 10(1). 131–156.

Tomasello, Michael. 2000. Do young children have adult syntactic competence? *Cognition* 74(3). 209–253. http://www.ehess.fr/centres/lscp/persons/anne/master/tomasello2000Cognition.pdf.

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

Tomasello, Michael. 2005. Beyond formalities: The case of language acquisition. *The Linguistic Review* 22(2–4). 183–197. http://www.eva.mpg.de/psycho/pdf/Publications_2005_PDF/Beyond_formalities_05.pdf.

Tomasello, Michael. 2006a. Acquiring linguistic constructions. In Deanna Kuhn & Robert Siegler (eds.), *Handbook of child psychology*, vol. 2, New York: John Wiley & Sons, Inc. 6th edn. http://email.eva.mpg.de/~tomas/pdf/tomasello_HoCP2005.pdf.

Tomasello, Michael. 2006b. Construction grammar for kids. *Constructions* Special Volume 1. http://www.constructions-online.de/articles/specvol1/689.

Tomasello, Michael. 2006c. Konstruktionsgrammatik und früher Erstspracherwerb. In Fischer & Stefanowitsch (2006) 19–37.

Tomasello, Michael. 2009. Universal Grammar is dead. *The Behavioral and Brain Sciences* 32(5). 470–471.

Tomasello, Michael, Nameera Akhtar, Kelly Dodsen & Laura Rekau. 1997. Differential productivity in young children's use of nouns and verbs. *Journal of Child Language* 24(2). 373–387.

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and sharing intentions: the origins of cultural cognition. *The Behavioral and Brain Sciences* 28(5). 675–735.

Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge, MA: M.I.T. Dissertation.

Tseng, Jesse (ed.). 2000. *Aspekte eines HPSG-Fragments des Deutschen* (Arbeitspapiere des SFB 340 No. 156). Eberhard-Karls-Universität Tübingen. http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/156/156abs.html.

Tseng, Jesse L. 2003. LKB grammar implementation: French and beyond. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the esslli 2003 workshop "ideas and strategies for multilingual grammar development"*, 91–97. Vienna, Austria. http://w3.erss.univ-tlse2.fr/textes/pagespersos/tseng/Pubs/mgd03.pdf.

Tseng, Jesse L. 2007. English prepositional passive constructions. In Stefan Müller (ed.), *Proceedings of the 14th international conference on Head-Driven Phrase Structure Grammar*, 271–286. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/.

Umemoto, Hiroshi. 2006. Implementing a Japanese semantic parser based on glue approach. In *Proceedings of the 20th pacific asia conference on language, information and computation*, 418–425. http://dspace.wul.waseda.ac.jp/dspace/handle/2065/28974.

Uszkoreit, Hans. 1986a. Categorial Unification Grammars. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 187–194. University of Bonn: Association for Computational Linguistics. http://aclweb.org/anthology-new/C/C86/C86-1045.pdf.

Uszkoreit, Hans. 1986b. Linear precedence in discontinuous constituents: Complex fronting in German. Report No. CSLI-86-47 Center for the Study of Language and Information Stanford, CA.

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

Uszkoreit, Hans. 1990. Extraposition and adjunct attachment in Categorial Unification Grammar. In Werner Bahner, Joachim Schildt & Dieter Viehweger (eds.), *Proceedings of the fourteenth international congress of linguists, Berlin/GDR, August 10–15, 1987*, vol. 3, 2331–2336. Berlin: Akademie-Verlag.

Uszkoreit, Hans, Rolf Backofen, Stephan Busemann, Abdel Kader Diagne, Elizabeth A. Hinkelman, Walter Kasper, Bernd Kiefer, Hans-Ulrich Krieger, Klaus Netter, Günter Neumann, Stephan Oepen & Stephen P. Spackman. 1994. DISCO—an HPSG-based NLP system and its application for appointment scheduling. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 436–440. Kyoto, Japan: Association for Computational Linguistics.

Uszkoreit, Hans, Rolf Backofen, Jo Calder, Joanne Capstick, Luca Dini, Jochen Dörre, Gregor Erbach, Dominique Estival, Suresh Manandhar, Anne-Marie Mineur & Stephan Oepen. 1996. The EAGLES formalisms working group – final report expert advisory group on language engineering standards. Technical Report LRE 61–100. http://www.coli.uni-sb.de/publikationen/softcopies/Uszkoreit:1996:EFW.pdf.

Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40(1–2). 21–81.

Van Langendonck, Willy. 1994. Determiners as heads? *Cognitive Linguistics* 5. 243–259.

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Steels (2011) 115–145.

van Trijp, Remi. 2013. A comparison between Fluid Construction Grammar and Sign-Based Construction Grammar. *Constructions and Frames* 5(1). 88–116.

van Trijp, Remi. 2014. Long-distance dependencies without filler–gaps: A cognitive-functional alternative in Fluid Construction Grammar. *Language and Cognition* 1–29.

Van Valin, Robert D. Jr. (ed.). 1993. *Advances in Role and Reference Grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Van Valin, Robert D. Jr. 1998. The acquisition of *wh*-questions and the mechanisms of language acquisition. In *The new psychology of language: Cognitive and functional approaches to language structure*, 221–249. Hillsdale, NJ: Lawrence Erlbaum.

Vancoppenolle, Jean, Eric Tabbert, Gerlof Bouma & Manfred Stede. 2011. A german grammar for generation in OpenCCG. In Hanna Hedeland, Thomas Schmidt & Kai Wörner (eds.), *Multilingual resources and multilingual applications: Proceedings of the conference of the german society for computational linguistics and language technology (GSCL) 2011* (Arbeiten zur Mehrsprachigkeit/Working Papers in Multilingualism, Folge B/Series B 96), 145–150. Universität Hamburg. http://www.corpora.uni-hamburg.de/gscl2011/downloads/AZM96.pdf.

Vargha-Khadem, Faraneh, Kate E. Watkins, Katie Alcock, Paul Fletcher & Richard Passingham. 1995. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. In *Proceedings of the national academy of sciences of the united states of america*, vol. 92, 930–933.

Vasishth, Shravan & Richard L. Lewis. 2006. Human language processing: Symbolic models. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, vol. 5, 410–419. Oxford: Elsevier Science Publisher B.V. (North-Holland) 2nd edn. http://www.ling.uni-potsdam.de/~vasishth/Papers/vasishthlewisElsevier.pdf.

Vasishth, Shravan, Katja Suckow, Richard L. Lewis & Sabine Kern. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes* 25(4). 533–567.

Vater, Heinz. 2010. Strukturalismus und generative Grammatik in Deutschland. In Hans-Harald Müller, Marcel Lepper & Andreas Gardt (eds.), *Strukturalismus in deutschland: Literatur- und sprachwissenschaft 1910–1975*, Göttingen: Wallstein Verlag.

Veenstra, Mettina Jolanda Arnoldina. 1998. *Formalizing the minimalist program*: Groningen dissertation.

Verhagen, Arie. 2010. What do you think is the proper place of recursion? conceptual and empirical issues. In Harry van der Hulst (ed.), *Recursion in human language* (Studies in Generative Grammar 104), 93–110. Berlin/New York, NY: Mouton de Gruyter.

Verspoor, Cornelia Maria. 1997. *Contextually-dependent lexical semantics*: University of Edinburgh dissertation. ftp://ftp.cogsci.ed.ac.uk/pub/kversp/thesis.ps.gz.

Vierhuff, Tilman, Bernd Hildebrandt & Hans-Jürgen Eikmeyer. 2003. Effiziente Verarbeitung deutscher Konstituentenstellung mit der Combinatorial Categorial Grammar. *Linguistische Berichte* 194. 213–237.

Vijay-Shanker, K. & Aravind K. Joshi. 1988. Feature structures based tree adjoining grammars. In Dénes Vargha (ed.), *Proceedings of COLING 88*, vol. 1, University of Budapest: Association for Computational Linguistics. http://www.aclweb.org/anthology-new/C/C88/C88-2147.pdf.

Villavicencio, Aline. 2002. The acquisition of a unification-based generalised categorial grammar. UCAM-CL-TR-533 University of Cambridge Computer Laboratory.

Vogel, Ralf. 2001. Case conflict in German free relative constructions. an optimality theoretic treatment. In Gereon Müller & Wolfgang Sternefeld (eds.), *Competition in syntax*, 341–375. Berlin: Mouton de Gruyter.

Vogel, Ralf & Markus Steinbach. 1998. The dative – an oblique case. *Linguistische Berichte* 173. 65–91.

Volk, Martin. 1988. *Parsing German with GPSG: The problem of separable-prefix verbs*. University of Georgia MA thesis. http://www.ifi.unizh.ch/CL/volk/SyntaxVorl/GPSG.thesis.html.

Wada, Hajime & Nicholas Asher. 1986. BUILDRS: an implementation of DR theory and LFG. In Makoto Nagao (ed.), *Proceedings of COLING 86*, 540–545. University of Bonn: Association for Computational Linguistics.

Wahlster, Wolfgang (ed.). 2000. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin/Heidelberg/New York, NY: Springer Verlag.

Walther, Markus. 1999. *Deklarative prosodische morphologie: Constraint-basierte analysen und computermodelle zum finnischen und tigrinya* (Linguistische Arbeiten 399). Tübingen: Max Niemeyer Verlag.

Webelhuth, Gert. 1985. German is configurational. *The Linguistic Review* 4(3). 203–246.

Webelhuth, Gert. 1990. Diagnostics for structure. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers* (Linguistik Aktuell/Linguistics Today 5), 41–75. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Webelhuth, Gert. 1995. X-bar theory and case theory. In Gert Webelhuth (ed.), *Government and binding theory and the minimalist program. principles and parameters in syntactic theory* (Generative Syntax), 15–95. Oxford, UK & Cambrigde, USA: Blackwell Publishing Ltd.

Webelhuth, Gert. 2011. Paradigmenwechsel rückwärts: die Renaissance der grammatischen Konstruktion. In Stefan Engelberg, Anke Holler & Kristel Proost (eds.), *Sprachliches wissen zwischen lexikon und grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), Berlin/New York, NY: de Gruyter.

Weber, Heinz J. 1997. *Dependenzgrammatik. Ein interaktives Arbeitsbuch* (Narr Studienbücher). Tübingen: Gunter Narr Verlag 2nd edn.

Wechsler, Stephen Mark. 1991. *Argument structure and linking*: Stanford University dissertation.

Wechsler, Stephen Mark. 1995. *The semantic basis of argument structure* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.

Wechsler, Stephen Mark. 1997. Resultative predicates and control. In Ralph C. Blight & Michelle J. Moosally (eds.), *Texas linguistic forum 38: The syntax and semantics of predication. proceedings of the 1997 texas linguistics society conference*, 307–321. Austin, Texas: University of Texas Department of Linguistics.

Wechsler, Stephen Mark. 2005. What is right and wrong about little v. In *Grammar and beyond—essays in honour of lars hellan*, 179–195. Oslo, Norway: Novus Press.

Wechsler, Stephen Mark. 2008a. A diachronic account of English deverbal nominals. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th west coast conference on formal linguistics*, 498–506. Somerville, MA: Cascadilla Proceedings Project.

Wechsler, Stephen Mark. 2008b. Dualist syntax. In Stefan Müller (ed.), *Proceedings of the 15th international conference on Head-Driven Phrase Structure Grammar*, 294–304. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2008/.

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

Wegener, Heide. 1985. *Der Dativ im heutigen Deutsch* (Studien zur deutschen Grammatik 28). Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Weir, Morton W. 1964. Developmental changes in problem-solving strategies. *Psychological Review* 71(6). 473–490.

Weissgerber, Monika. 1983. *Valenz und Kongruenzbeziehungen: Ein Modell zur Vereindeutigung von Verben in der maschinellen Analyse und Übersetzung*. Frankfurt a. M.: Peter

Lang.

Weisweber, Wilhelm. 1987. Ein Dominanz-Chart-Parser für generalisierte Phrasenstrukturgrammatiken. KIT-Report 45 Technische Universität Berlin Berlin.

Weisweber, Wilhelm & Susanne Preuss. 1992. Direct parsing with metarules. In Antonio Zampoili (ed.), *14th international conference on computational linguistics (COLING '92), august23–28*, 1111–1115. Nantes, France: Association for Computational Linguistics. http://acl.ldc.upenn.edu/C/C92/C92-4174.pdf.

Welke, Klaus. 1988. *Einführung in die valenz- und kasustheorie*. Leipzig: Bibliographisches Institut.

Welke, Klaus. 2009. Konstruktionsvererbung, Valenzvererbung und die Reichweite von Konstruktionen. *Zeitschrift für Germanistische Linguistik* 37(3). 514–543.

Welke, Klaus. 2011. *Valenzgrammatik des Deutschen: Eine Einführung* (De Gruyter Studium). Berlin/New York, NY: de Gruyter.

Wells, Rulon S. 1947. Immediate constituents. *Language* 23(2). 81–117.

Wetta, Andrew C. 2011. A construction-based cross-linguistic analysis of V2 word order. In Stefan Müller (ed.), *Proceedings of the 18th international conference on Head-Driven Phrase Structure Grammar, university of washington*, 248–268. Stanford, CA: CSLI Publications.

Wexler, Kenneth. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106(1–4). 23–79.

Wexler, Kenneth & Peter W. Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA/London: MIT Press.

Wharton, R. M. 1974. Approximate language identification. *Information and Control* 26(3). 236–255.

White, Mike & Jason Baldridge. 2003. Adapting chart realization to CCG. In Ehud Reiter, Helmut Horacek & Kees van Deemter (eds.), *Proceedings of the 9th european workshop on natural language generation (enlg-2003) at eacl 2003*, 119–126.

Wijnen, Frank, Masja Kempen & Steven Gillis. 2001. Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language* 28(3). 629–660.

Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Kristine Bentzen & Þorbjörg Hróarsdóttir. 2007. Rethinking Scandinavian verb movement. *Journal of Comparative Germanic Linguistics* 10(3). 203–233.

Wilcock, Graham. 2001. Towards a discourse-oriented representation of information structure in HPSG. In *13th nordic conference on computational linguistics, uppsala, sweden*, www.ling.helsinki.fi/~gwilcock/Pubs/2001/Nodalida-01.pdf.

Wilder, Chris. 1991. Small clauses and related objects. *Groninger Arbeiten zur Germanistischen Linguistik* 34. 215–236.

Williams, Edwin. 1984. Grammatical relations. *Linguistic Inquiry* 15(4). 639–673.

Winkler, Susanne. 1997. *Focus and secondary predication* (Studies in Generative Grammar 43). Berlin, New York: Mouton de Gruyter.

Wittenberg, Eva, Ray S. Jackendoff, Gina Kuperberg, Martin Paczynski, Jesse Snedeker & Heike Wiese. 2014. The processing and representation of light verb constructions. In Asaf Bachrach, Isabelle Roy & Linnaea Stockall (eds.), *Structuring the argument* (Language Faculty and Beyond 10), Amsterdam/Philadelphia: John Benjamins Publishing Co.

Wittenberg, Eva & Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3). 393–413.

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

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

Wunderlich, Dieter. 1989. Arnim von Stechow, das Nichts und die Lexikalisten. *Linguistische Berichte* 122. 321–333.

Wunderlich, Dieter. 1992. Cause and the structure of verbs. Arbeiten des SFB 282 No. 36 Heinrich Heine Uni/BUGH Düsseldorf/Wuppertal.

Wunderlich, Dieter. 1997. Argument extension by lexical adjunction. *Journal of Semantics* 14(2). 95–142.

Wunderlich, Dieter. 2004. Why assume UG? *Studies in Language* 28(3). 615–641.

Wunderlich, Dieter. 2008. Spekulationen zum Anfang von Sprache. *Zeitschrift für Sprachwissenschaft* 27(2). 229–265.

Wurmbrand, Susanne. 2003a. *Infinitives. restructuring and clause structure* (Studies in Generative Grammar 55). Berlin/New York, NY: Mouton de Gruyter.

Wurmbrand, Susanne. 2003b. Long passive (corpus search results). http://wurmbrand.uconn.edu/research/files/long-passive.pdf.

XTAG Research Group. 2001. A Lexicalized Tree Adjoining Grammar for English. Tech. rep. Institute for Research in Cognitive Science Philadelphia. ftp://ftp.cis.upenn.edu/pub/xtag/release-2.24.2001/tech-report.pdf.

Yamada, Hiroyasu & Yuji Matsumoto. 2003. Statistical dependency analysis with support vector machines. In Gertjan van Noord (ed.), *Proceedings of the 8th international workshop on parsing technologies (iwpt 03)*, Nancy.

Yamada, Jeni. 1981. Evidence for the independence of language and cognition: Case study of a "hyperlinguistic" adolescent. UCLA Working Papers in Cognitive Linguistics 3 University of California, Los Angeles.

Yampol, Todd & Lauri Karttunen. 1990. An efficient implementation of PATR for Categorial Unification Grammar. In Hans Karlgren (ed.), *COLING-90. papers presented to the 13th international conference on computational linguistics*, 419–424. Helsinki.

Yang, Charles D. 2004. Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8(10). 451–456. doi:10.1016/j.tics.2004.08.006. http://www.isrl.uiuc.edu/~amag/langev/paper/yang04UG.html.

Yasukawa, Hidekl. 1984. LFG system in prolog. In Yorick Wilks (ed.), *Proceedings of the 10th international conference on computational linguistics and 22nd annual meeting of the association for computational linguistics*, 358–361. Stanford University, California: Association for Computational Linguistics.

Yip, Moira, Joan Maling & Ray S. Jackendoff. 1987. Case in tiers. *Language* 63(2). 217–250.

Yoshinaga, Naoki, Yusuke Miyao, Kentaro Torisawa & Jun'ichi Tsujii. 2001. Resource sharing amongst hpsg and ltag communities by a method of grammar conversion between FB-LTAG and HPSG. In *Proceedings of acl/eacl workshop on sharing tools and resources for research and education*, 39–46. Toulouse, France.

Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In Dalrymple et al. (1995) 215–239.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483. http://www2.parc.com/istl/members/zaenen/publications/Passive.pdf.

Zappa, Frank. 1986. Does humor belong in music? EMI Music Germany GmbH & Co.KG.

Zucchi, Alessandro. 1993. *The language of propositions and events: Issues in the syntax and the semantics of nominalization* (Studies in Linguistics and Philosophy 51). Berlin/Heidelberg/New York, NY: Springer Verlag.

Zwart, C. Jan-Wouter. 1994. Dutch is head-initial. *The Linguistic Review* 11(3–4). 377–406.

Zweigenbaum, Pierre. 1991. Un analyseur pour grammaires lexicales-fonctionnelles. *TA Informations* 32(2). 19–34. http://www.limsi.fr/~pz/Publications/biblio-pierre-pardate/.

Zwicky, Arnold M., Joyce Friedman, Barbara C. Hall & Donald E. Walker. 1965. The MITRE syntactic analysis procedure for Transformational Grammars. In *Proceedings – fall joint computer conference*, 317–326. http://doi.ieeecomputersociety.org/10.1109/AFIPS.1965.108.

Name index

Language index

Subject index