Englisches Seminar
der Universität Zürich

November 1998

Lizentiatsprüfungen der Philosophischen Fakultät I der Universität Zürich:
Dreitägige Hausarbeit

Referentin: Prof. Dr. Gunnel Tottie

An Introduction to Government & Binding

Notes for an Imaginary Colloquium

Gerold Schneider

Büro 27-K-48
Institut für Informatik
Winterthurerstr. 190
CH - 8057 Zürich

Lerchenhalde 43
CH - 8046 Zürich
+41 1 / 371 85 06
+41 1 / 635 67 24

gschneid@ifi.unizh.ch

November 2-5, 1998

Colloquium Gerold Schneider: An Introduction to Government & Binding Theory

Colloquium Description: This colloquium is intended to give an introduction to one of the major formal grammar theories: *Government & Binding (GB)*, also known as *Principles & Parameters*. It is a natural continuation and extension to the syntax taught in *Introduction to Linguistics I* and *II*. The colloquium rather aims to give an overview than to dwell on technicalities. It will also discuss some fundamental problems of any grammar theory and highlight similarities and differences to other approaches. Chomsky's theories about language knowledge and acquisition will not be discussed in this course.

Course Requirements: The colloquium is situated at seminar level, i.e. you should have passed your Linguistik-Akzessprüfung, and it may be helpful if you have attended a seminar beforehand.

Course Structure: We will deal with one topic in each meeting, i.e. seven topics in total, as the chapters of these colloquium notes express. It will be a fully taught course, there will be no student presentations, but discussions and questions are an essential part of the colloquium. You are requested to read the corresponding chapter of these notes before the colloquium (except for the first chapter, of course), which we will work through carefully and discuss in class. I will recommend excerpts of additional literature for most chapters. Please ask me if you would like to get copies of these excerpts.

1 BACKGROUND OF GOVERNMENT & BINDING (GB) 4
1.1 LANGUAGE STRUCTURE 5
1.1.1 Structure inside Structure 5
1.1.2 Constituents 5
1.1.3 Phrase Structure or Rewrite Rules 6
1.1.4 Decisions about Linguistic Structure 6
1.1.5 Exploring the Elephant 7
1.1.6 Recursion 8
1.2 COMPETENCE VERSUS PERFORMANCE 8
1.2.1 Statements on Syntactic Performance 9
1.2.2 Using Performance Statements for (Automatic) Parsing 9
1.3 SYNTACTIC RULES AND THE LEXICON 10
1.3.1 Verb Subcategorisation 10
1.3.2 Eliminating Redundancy 11
1.3.3 Lexicalism 11
1.4 TRANSFORMATIONS, DEEP AND SURFACE STRUCTURE 12
1.4.1 Deep and Surface Structure 12
1.4.2 The T-Model 15
1.5 THE LANGUAGE FACULTY 16
1.5.1 Principles and Parameters 16
1.5.2 Language Acquisition 16
1.6 RECOMMENDED LITERATURE 16
2 X-BAR SYNTAX 18
2.1 THE HEAD OF A PHRASE 18
2.1.1 A Universal Rewrite Rule 19
2.1.2 Transitive and Intransitive Prepositions 19
2.2 INTERMEDIATE CATEGORIES 21
2.2.1 First Disadvantage: One-Pronominalisation 21
2.2.2 Second Disadvantage: Arguments and Adjuncts 22
2.3 THE X-BAR RULE SCHEMATA 25
3 FUNCTIONAL CATEGORIES 27
3.1 SOME DETECTIVE WORK 27
3.2 THE HEAD OF A SENTENCE 28
3.2.1 NP versus VP 28
3.2.2 V-Triple-Bar ? 28
3.2.3 What is S ? 29
3.2.4 Inflection as a Functional Category 30
3.2.5 INFL as the Head of the Sentence 31
3.2.6 The Subject under the IP assumption 32
3.3 COMPLEMENTIZERS 33
3.4 OTHER FUNCTIONAL CATEGORIES 34
4 MOVEMENT AND CASE THEORY 35
4.1 A DETAILED EXAMPLE OF MOVEMENT 35
4.1.1 Building Up the D-Structure 35
4.1.2 Moving to S-Structure 38
4.1.3 Chains 42
4.2 CASE THEORY 42
4.2.1 Restricting Movement 42
4.2.2 Abstract Case 43
4.2.3 Inherent Case 43
4.2.4 Structural Case 44
5 AN APPLICATION OF GB: V2 IN GERMANIC LANGUAGES 45
6 ASPECTS OF MINIMALISM 46
6.1 CHECKING THEORY 46
6.2 BARE PHRASE STRUCTURE 47
7 WHAT WE HAVE NOT COVERED 48
8 BIBLIOGRAPHY 49

1 Background of Government & Binding (GB)

RECOMMENDED READING for Chapter 1: Cook & Newson (1996: 1-49)

GB has several predecessors and a successor is beginning to emerge now. While some parts of the theory have remained stable right from the start, others have developed considerably. The most important predecessor is *Transformational Grammar*, which is presented first in Chomsky (1957) and Chomsky (1965). As the name suggests, **transformations** were and still are a most prominent part of the theory. Transformations basically transform related structures into each other. Transformations can e.g. express the relationship between the declarative sentence *You are a student* and the question *Are you a student ?*. We will hear more about transformations in 1.4 and chapter 4.

The successor that is emerging now is called *The Minimalist Theory* (Chomsky 1995), but it is not entirely clear yet in which direction it will develop. If we have enough time I would like to deal with two suggestions made in Minimalism; *Checking Theory* in subchapter 6.1, and *Bare Phrase Structure* in subchapter 6.2.

GB was first introduced in Chomsky (1981). **Noam Chomsky**, the inventor and main protagonist of GB, is notorious for writing articles and books that are difficult to read, and for inventing (and sometimes rebutting) a plethora of technical terms. While it may certainly be interesting to read „the original“, be warned about the extremely difficult style: Even experts admit they find it very difficult to follow! In a much milder form, the same warning applies to any introductory book to GB: Do not be put off by linguistic jargon and technical terms. Try to find their definition, you will usually discover that a complicated name hides a more simple and fairly intuitive concept. It has to be admitted, however, that many of the textbook discussions are rather technical, which I will try to avoid here, whenever possible.

1.1 Language Structure

One of the main aims of linguistics is to find and describe the structure that is hidden behind the utterances we make. Let us look at two examples:

(1) Peter loves Mary.

(2) *Loves Peter Mary.

We will intuitively accept (1) but reject (2). We will conclude from this, in accordance with all linguists, that an English sentence usually has Subject-Verb-Object (SVO) order. SVO, in this order, is the structure of an English sentence.

1.1.1 Structure inside Structure

If we see a sentence like

(3) Peter loves a clever girl.

should we suggest that English has Subject-Verb-Article-Adjective-Object structure?

Probably all linguists would still agree that (3) is also an example of the Subject-Verb-Object structure of English, and that its object consists of several words, or that the object has words connected to it on a secondary level, or something similar – in short, they would say that the object has a structure inside the structure of the sentence. Language is generally supposed to have hierarchical structure, i.e. subordinate structures within bigger structures.

1.1.2 Constituents

We can replace the single word *Mary* by a sequence that *consists of* words, so-called constituents, without rendering the sentence ungrammatical. Beside this well known **replacement test** linguists use a couple of additional tests to assess if a sequence of words makes up a constituent (Radford 1988: 69 ff.). Constituents are bracketed:

- Preposing: Only entire constituents can be preposed: *I can't stand [your older sister]* <=> *[Your older sister], I can't stand.*
- Postposing: Only entire constituents can be postposed: *He explained [all the problems that he had encountered] to her.*<=> *He explained to her [all the problems that he had encountered].*
- Conjunction: Only constituents of the same level can be conjoined: *He called [Mary] and [a clever girl].*

- Short answers: Only constituents can be asked for or constitute a short answer: *[Who] did you call ? – [A clever girl].*
- Pronominalisation: Only entire constituents can be replaced by pronouns: *Have you seen [the young student from Cambridge] ? – Yes, I have seen [him].*

1.1.3 Phrase Structure or Rewrite Rules

Both *Mary* and *a clever girl* are instances of a noun phrase (NP). The structure of a NP can consist of a proper noun (PN) like *Mary*, or of an article (ART) followed by a noun (N), which is in turn possibly followed by an adjective (ADJ). We can describe these facts in so-called **phrase structure rules** or **rewrite rules**, where „→" can be read as „consists of" (cf. e.g. Finegan (1994: 133))

(4) NP → PN

NP → ART (ADJ) N

If we ask ourselves how the structure of the whole sentence (S) could be expressed in this way, it will be natural to suggest the following rule, expressing the discussed Subject-Verb-Object order (where V stands for Verb):

(5) S → NP V NP

We may find, however, that this rule predicts ungrammatical sentences like

(6) *Peter laughs Mary.

Rule (5) would only allow transitive verbs. We have to formulate an additional rule

(7) S → NP V

1.1.4 Decisions about Linguistic Structure

It has generally been accepted, however, that it is more elegant to distinguish between sentences with *transitive* and *intransitive* verbs at a subordinate rewrite rule stage, not at the S-level, the topmost level in the hierarchy of our small rewrite rule grammar. Transitive and intransitive verbs are said to have too much in common to justify two completely separate sentential representation rules. We therefore use a uniform rewrite rule for the sentence by introducing a Verb Phrase (VP), which in turn contains a verb in case of an intransitive verb or a verb followed by a noun for a transitive verb.

(8) S → NP VP

VP → V

VP → V NP

While this grammar is more elegant, as it allows a uniform treatment of all verbs, the solution we have sketched with rules (5) and (7) cannot be said to be „wrong“ in any sense. We can at best find it less convincing and therefore discard it. We had to take a decision as to how the underlying structure of a sentence could look like. While we had good arguments for this decision and could possibly find more, nobody can be forced to agree, because we cannot physically see the structure of a sentence under a microscope. We will see later that there are cases where agreement among linguists is more difficult to achieve, that different solutions are being discussed.

1.1.5 Exploring the Elephant

Linguists are like people in a dark room who touch an elephant and try to describe it. The one touching the tail will find it hairy, the one touching the trunk will find it soft and skilful, the one touching a leg will describe it as rough-skinned and ponderous. Yet none can know how the „elephant“, i.e. the structure of language, really looks like or if it exists at all.

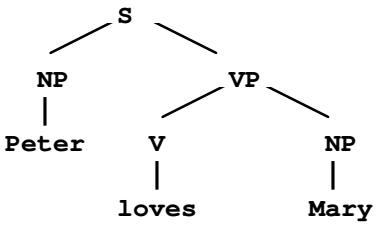
> The conventional wisdom that a language like English has a real existence led to the idea that we only have to discover the structure of *the* language. The question than became: *How* could we discover its structure and the structure of language in general? In later years it was thought, optimistically, that it was simply a problem of the blind men, each examining a different part of an elephant. It was thought that more work was needed on devising better test criteria and on how to interpret the results of the tests, and that these results would eventually converge on a consistent overall description of the elephant.
>
> But, as one of my students gasped when realizing the implications of the domain confusions, "There is no elephant!"
>
> (Yngve 1996: 46)

Doubting the very existence of language structure is of course the most radical formulation. By *domain confusion*, Yngve means that many linguists tend to present their often debatable findings as if they were proven scientific facts, e.g. structures we see under a microscope.

One of the characteristics of GB is indeed that everything is defined in terms of structure-relations, so-called **configurationally**.

GB Principle 1: Operations on sentences require a knowledge of the structural relationships of the words

We have seen that English has SVO order. Using the simple phrase structure grammar just developed in (8) we get syntax trees like



We can define *subject* to be the NP whose mother node is S, and *object* to be the NP whose mother is VP. GB in fact uses a further development of this simple definition, which does not necessarily apply to all languages.

I do not want to discuss these issues in detail, but it is important to stress that the assumptions that GB takes, the structures that GB suggests are often not be the only solutions possible. On the other hand, a formal grammar like GB is consistent and rigorous. Once a decision has been taken, it is followed consistently, and the assumed decisions are often applied to decide related questions.

1.1.6 Recursion

Let us return to our small rewrite grammar, which expresses one possible way to see structure in a sentence. If we have a sentence like

(9) Sandra knows that Peter loves Mary

our small grammar fails to express it. The VP is far more complex in this sentence. Instead of V+NP, as is usual for transitive verbs, it consists of V followed by *that* and the assertion of a fact that looks exactly like a sentence itself. We may indeed suggest that the VP of (9) can be expressed by rule (10):

(10) VP → V *„that“* S

This allows us to analyse sentences like

(11) He suspects that Sandra knows that Peter loves Mary.

If a rule (partly) rewrites as a previous rule or even as itself we speak of **recursion**. Recursion can potentially go on forever, which allows, at least theoretically, for endless structures. Recursion is one way in which, as it is said of language, we can form an infinite number of structures from a finite number of rules.

1.2 Competence versus Performance

The mental capacity that enables a speaker to utter well-formed sentences and to refute ungrammatical ones is called **competence**. It can be described by a rule system similar to (but of course much more complex than) the small rewrite grammar we have started to write in 1.1. But such a grammar makes no statements about **performance**, i.e. which

sentences and structures are actually found in language use, and how frequent they are. Although GB has developed in many ways from rewrite grammars, as we shall see, it is still and intentionally a competence-based grammar theory. Chomsky uses the terms **I-language** (internal language) for competence, and **E-language** (external language) for performance.

A major aim of *Corpus Linguistics* is to make statements about language performance. At the first sight, Corpus Linguistics and a formal grammar theory, such as GB, have few common interests. But they can benefit from each other, precisely because of their different approach, in order to complement each other.

1.2.1 Statements on Syntactic Performance

If we syntactically analyse a corpus by using the rigorous and consistent methods of a formal grammar such as GB, we can find out which structures and which rules or principles are used more frequently, and which are very rare in language use. It usually also happens – even with very sophisticated grammars – that some sentences cannot be correctly analysed, which indicates that the grammar is not (yet) complete. Corpus data helps to test a grammar, as well as the assumptions that a grammar makes.

1.2.2 Using Performance Statements for (Automatic) Parsing

Conversely, a grammar theory can greatly profit from the syntactical performance statements that we can obtain as just described. Most sentences are **ambiguous** in many ways, although we are rarely aware of it. The short sentence

(12) Time flies like an arrow.

describes at least the following four facts:

(12a) Time elapses too quickly.

(12b) Short-lived insects sympathise with an arrow.

(12c) Use your stop watch to measure the speed of insects, as if they were an arrow.

(12d) Use your stop watch to measure the speed of insects, as if you yourself were an arrow.

A very good automatic analyser would report all these readings and fail to decide which could be best. But if a corpus provides us with morphosyntactic performance data, in this case the frequency of the words in the sentence in their different possible lexical categories, reading (12a) will be recognised as the most likely reading.

1.3 Syntactic Rules and the Lexicon

1.3.1 Verb Subcategorisation

In a rewrite grammar fragment we might find the following rules

(13) VP → V

VP → V NP

There are different verb **subcategories**, e.g. transitive and intransitive. How can we make sure that intransitive verbs will only pick the former VP rule, and transitive verbs only the latter rule? One solution is to add an argument to the rule, as follows:

(14) VP → V(intrans)

VP → V(trans) NP

The information on the transitivity of a verb depends on the verb itself. There is no way to predict if, say an unknown verb in a foreign language we see in isolation, could be transitive or intransitive. This information is *idiosyncratic* for each verb, it therefore needs to come from the lexicon. Lexical entries can be made as rewrite rules, which would look as follows for the verbs *to love* and *to laugh*:

(15) V(intrans) → laugh

V(trans) → love

There are many different types of verbs, however, which select all kinds of complements. We have seen in 1.1.6 that the verb *know* takes a subordinating element (so-called **complementizer**), i.e. *that*, followed by a sentence.

(16) Sandra knows that Peter loves Mary

The verb *ask* takes a similar, but slightly different complement:

(17) *Sandra asks that Peter loves Mary

(18) Sandra asks if Peter loves Mary

But the verb *want* only takes a non-finite clause, a „sentence without inflection“ for GB as we shall see later.

(19) *Sandra wants that Peter loves Mary

(20) Sandra wants Peter to love Mary

Or, one more example, the verb *put* is ditransitive, but it requires a preposition, usually *on* or *into* for its second argument.

(21) *Sandra puts the book

(22) Sandra puts the book on the table

This diversity of verb requirements has led linguists to directly mark in the lexicon what each verb takes as complement. The (simplified) lexical entries for some of the above verbs look as follows in GB:

(23) love Verb, [_ NP]

know Verb, [_ that-S]

ask Verb, [_ if-S]

put Verb, [_ NP PP]

The underscore (_) refers to the subject, on which the verb usually places no categorial restrictions. The fact that verbs are grouped into different subcategories is called **subcategorisation**. When subcategorisation is expressed directly in the lexical entry, as in (23), we speak of **direct** or **strict subcategorisation**. The bracketed subcategorisations, e.g. [_ NP] are also referred to as **subcategorisation frame**, and the verb *love* **subcategorises for** an NP.

1.3.2 Eliminating Redundancy

Strict subcategorisation, as explained above, expresses information that is already contained in the phrase structure rules, it only helps to select the appropriate rule. It is indeed questionable if it is useful to store the same information in two places, to have **redundant** information. Because subcategorisation information depends on the individual verb we cannot eliminate it from the lexicon, but we can eliminate it in the phrase structure rule. We only have to make sure that lexical information is properly inserted into the remaining partial phrase structure rules. The so-called **Projection Principle** ensures this (cf. Cook & Newson 1996: 20)

GB Principle 2 (Projection Principle): The properties of lexical entries project onto the syntax of the sentence.

1.3.3 Lexicalism

Not only verbs, but also many lexical entries show some idiosyncratic behaviour, i.e. they specifically select certain other categories or lexical items. The noun *relation*, e.g., sometimes occurs alone, but it often subcategorises for a PP with the preposition *to,* or two nouns connected by *between*:

(24) a. The verb has a relation to the subject.

b.*The verb has a relation.

(25) a. The diplomatic relations between Turkey and Iran are quite bad.

b. The diplomatic relations are quite bad.

Such information belongs into the lexicon. Most formal grammars, including GB, have increasingly come to depend on the lexicon. The only remnant of phrase structure rules left in GB are a structural constraint called X-bar, which will be presented in chapter 2. Grammars that rely mainly on the lexicon are called **lexicalist**.

In parsing technology, which is a topic in computational linguistics or artificial intelligence, one distinguishes between *top-down* and *bottom-up* techniques. Top-down algorithms start at the topmost node, usually S, and try to apply the rewrite rules until they reach the lexical items, like a phrase structure rewrite grammar. Bottom-up algorithms start with the lexical items and build up the structure from them, just as a lexicalist grammar does.

1.4 Transformations, Deep and Surface Structure

Transformations, also called **movements**, are one of the first and most prominent parts of GB and its predecessor, *Transformational Grammar* (Chomsky(1957)). The basic idea of transformation is that functionally related sentences should be derived from the same structure. E.g. the question

(26) What are you seeing at the cinema ?

is related to the declarative sentence

(27) You are seeing „When Harry met Sally" at the cinema.

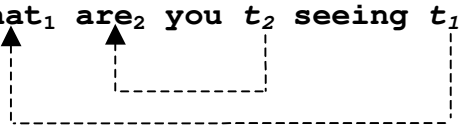
1.4.1 Deep and Surface Structure

GB distinguishes a **surface structure** (**S-structure**), the sentences as we see them in print, and an underlying **deep structure** (**D-structure**). Both (26) and (27) are surface structures, but they have an identical deep structure – except for the question word *what* instead of the movie title, of course. The surface structure of the declarative sentence (26) is (almost) identical to its deep structure. The deep structure of (26) is probably (but we will revise this slightly in a minute):

(26 D) You are seeing what at the cinema ?

In the transformation from the deep structure to the surface structure the question word *what* is fronted (as in *topicalisation*) and the auxiliary verb *are* moves as well. All the moved elements leave so-called **traces** (indicated by ***t***) in their places of origin in order to maintain the structural relationship to the deep structure.

(26 S) What$_1$ are$_2$ you t_2 seeing t_1 at the cinema ?



After we will have seen in chapter 2 how the assumed structure of a sentence looks like in GB, we will see in detail how this transformation works in chapter 4. Let us remember for the moment:

GB Principle 3: D-structure expresses the underlying form before movement.

GB Principle 4: S-structure describes the related form of the sentence after movement, including traces (*t*) of the original D-structure positions of the items.

In *Transformational Grammar* there were many different transformations, the one for questions presented above was just one of them. This plethora of transformations turned out to be very difficult to handle, especially computationally. It was very difficult to calculate where a transformation should move to, and the different transformations interfered with each other. For these technical reasons, only one general transformation was adopted, so-called **move-alpha** (which could basically move anything anywhere), whose power is restricted by many **constraints**, i.e. restrictions, which ensure that only the desired transformations can take place. Such constraints restrict what can be moved, and where the elements can move to.

Two such constraints state, e.g., that movement can only be towards the front of the sentence, and only to positions that are still empty and suitable for the moved element (so-called **landing sites**).

GB Principle 5: Elements can only be moved towards the beginning of a sentence.

GB Principle 6: Elements can only be moved to suitable landing sites, which need to be unoccupied.

Another GB principle says that transformations are **structure-preserving**, which means that the deep and the surface structure of a sentence have the same structure.

GB Principle 7: Move-alpha is structure-preserving.

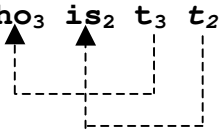
Although we do not know the exact structure of GB sentences yet, from the fact that transformations are structure-preserving we can already predict that (26 D) and the almost identical (27D) must have at least two empty positions at the beginning of the sentence in order to take up the moved elements. Let us provisionally call them *E1* and *E2*.

(26 D) *E1 E2* You are seeing what at the cinema ?

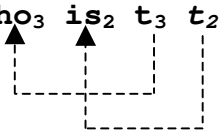
(27 D) *E1 E2* You are seeing „When Harry met Sally" at the cinema.

If the constraints on movements should make sure that only appropriate movements should take place we can also safely assume that – at the very least in this kind of sentence – there are only two initial and suitable empty positions at D-structure, i.e. positions that can serve as landing sites at S-structure. If more than two landing sites were available, more surface structures without any obvious difference to (26 S) and (27 S) would be allowed, and questions with more than one fronted question word would also be allowed. (28 S) or (29 S) move the question word about the movie-spectator's identity to the front, but (30S) tries to move both the question word about the movie-spectator's identity and the question word for the title to the front, which is ungrammatical.

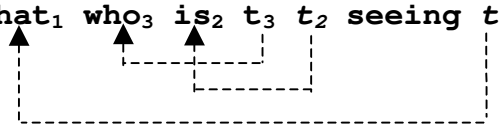
(28 S) Who$_3$ is$_2$ t$_3$ *t$_2$* seeing „When Harry met Sally" at the cinema ?



(29 S) Who$_3$ is$_2$ t$_3$ *t$_2$* seeing what at the cinema ?



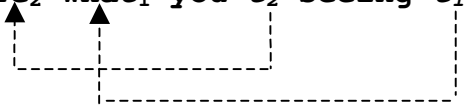
(30 S)*What$_1$ who$_3$ is$_2$ t$_3$ *t$_2$* seeing *t$_1$* at the cinema ?



The ungrammaticality of (30 S) also indicates that, as just said, every position can only be occupied once.

We can also assume that the auxiliary *are* can only move to *E2*, while the question word *what* can only move to *E1*. Otherwise the following sentence would be grammatical:

(26 S)*Are$_2$ what$_1$ you *t$_2$* seeing *t$_1$* at the cinema ?



The landing sites need to be somehow *compatible* to the elements that move to them. We will come back to this issue at the beginning of chapter 4.

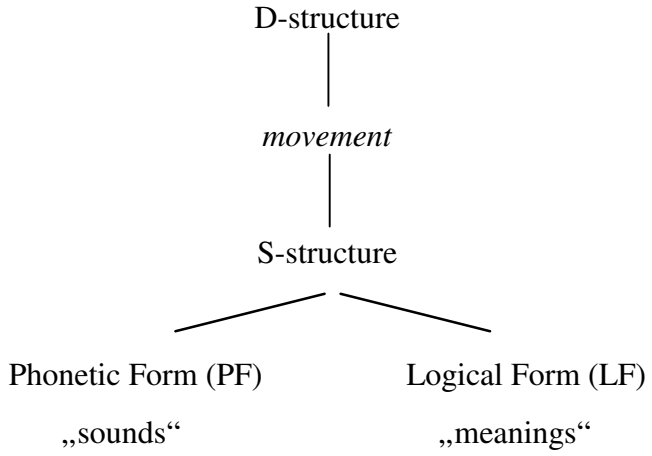
1.4.2 The T-Model

We have now seen the relationship between D-structure and S-structure. Traces and structure-preservation ensure that the D-structure is always fully recoverable from the S-structure.

GB Principle 8: D-structure is always recoverable from S-structure.

We have a model of syntax now. Chomsky stresses the **autonomy of syntax**, i.e. that syntax has properties of its own, which are completely unrelated to morphology, phonology or semantics. But in order to make up an all-encompassing language-model, as GB aims to do, we have to include them somewhere in the model. We shall see that morphology is taken care of during movement, and that it receives a novel treatment under *Checking theory* in Minimalism. As for phonology and semantics, GB provides a special structure level for each. Phonological structure, called **phonetic form (PF)** is of course derived from S-structure. This part of GB has in fact received very little attention until now. It is more surprising to see that the semantic structure, called **logical form (LF)**, should also be derived from S-structure, and no longer from D-structure, as was suggested originally in Transformational Grammar. This is possible because (due to GB Principle 8 above) still contains all the information of the D-structure. I will not deal with LF here, but I can recommend Hornstein (1995) for those interested, although more than a superficial acquaintance with GB is assumed to profit from reading it.

Thus the language model of GB, also called the T-model because of its shape of a reversed T, looks as follows:



1.5 The Language Faculty

1.5.1 Principles and Parameters

Principles and Parameters, in short P&P, is another name for Government and Binding Theory, one which is even, as Chomsky himself points out, more appropriate. While Government and Binding are small subparts of the theory, which we will probably not even discuss in our colloquium, the term *principles and parameters* conveys the credo of Chomsky. His aim is to describe a universal grammar theory (UG), which describes all languages. The core is common to all languages – these are the *principles*. The individual languages only differ in *parameters*. One such parameter is e.g. the order within constituents.

1.5.2 Language Acquisition

Chomsky has far-reaching claims with his principles and parameters model. The principles should be the language universals everybody is born with, while the parameters are set during language acquisition. I will not deal with this topic here, but quote one passage from Chomsky (1995):

> We assume that the system described by UG is a real component of the mind/brain, put to use in the complex circumstances of ordinary life. The validity of this assumption is hardly in question. (Chomsky 1995: 19)

1.6 Recommended Literature

While this colloquium can give you a short overview, it is warmly recommended to read one of the following books in addition to following the course. Many of the issues are covered in greater detail in these books, and it usually helps to hear the same thing twice, from a different viewpoint.

I will recommend a chapter from one or two of these books for each chapter or subchapter. My notes do not closely follow the recommended chapter, however, but they try to take up some of the points in a less technical fashion.

CHOMSKY, NOAM, 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press. The original. Only for the adventurous. You have been warned.

COOK, VIVIAN and MARK NEWSON, 1996. *Chomsky's Universal Grammar*. Oxford: Blackwell. A book that I can recommend without reservations. It is fairly manageable to read and understand, yet never simplistic.

The summaries and definitions in grey boxes, as well as the quick reference list at the end of the book are especially helpful for reference.

COWPER, ELISABETH, 1992. *A Concise Introduction to Syntactic Theory.* Chicago: The University of Chicago Press. As the title promises, very concise and quite short. Rather technical but very well structured, recommendable especially for its logical flow of arguments. The book with which I have made my first steps into GB.

CULICOVER, PETER, 1997. *Principles and Parameters: An Introduction to Syntactic Theory.* Oxford: OUP. Not my favourite book. It is unusual that it discusses movement in detail before introducing X-bar grammar. I often had serious difficulties to follow the rushed arguments and dense style. I do not know GB well enough yet to profit from this book.

HAEGEMAN, LILIANE, 1994. *Introduction to Government & Binding Theory*, 2nd ed. Oxford: Blackwell. The most comprehensive book on this list. More than just an introduction, quite technical. In my opinion, suitable as somebody's second book on GB.

RADFORD, ANDREW, 1988. *Transformational Grammar: A First Course.* Cambridge: CUP. A classic. Well suitable as an introduction, also to syntax in general. Some fundamental issues which are more or less taken for granted today, e.g. identifying constituents or why one needs intermediate categories, are discussed at the length they deserve. As the title says, the version of GB presented is unfortunately a bit outdated.

RADFORD, ANDREW, 1997. *Syntactic Theory and the Structure of English. A Minimalist Approach.* Cambridge: CUP. Has good chances to become a follow-up classic. It naturally integrates some Minimalist suggestions, especially Checking theory. The last two chapters, however, describe quite baroque extensions to the theory which are perhaps too unwieldy to convince.

2 X-bar Syntax

RECOMMENDED READING for Chapter 2: Cowper (1992: 19-33), Radford (1988: 167-187)

So far we have employed phrase structure rewrite rules to make a small English grammar. We have heard in 1.3 that a part of them may be replaced by lexical requirements projected up to the syntax, but this does not change anything about how the structure of a sentence will eventually look.

2.1 The Head of a Phrase

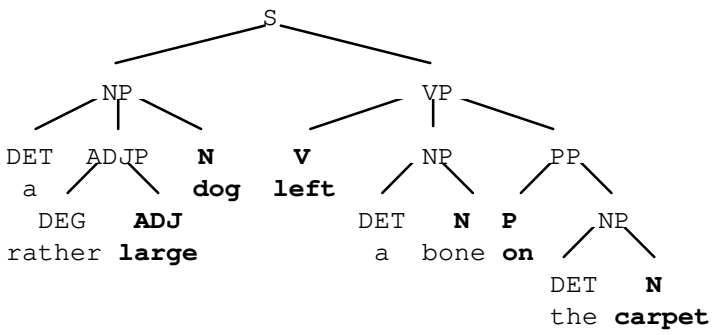
A typical pre-GB phrase structure rule grammar (taken from Cowper (1992: 20), which I will follow for much of this subchapter) may have looked as follows (elements in brackets are optional):

(31) **S → NP VP**
VP → V (NP) (ADVP / PP)
NP → (DET) (ADJP) N
PP → P NP
ADJP → (DEG) ADJ

Except for the categories S and PP, all the rules have only one obligatory element, which is the word that gives the phrase its name, N for NP, V for VP, etc. It is obviously the most prominent element in the phrase. It is therefore called the **head** of the phrase. We will discuss if the NP needs to be compulsory in PPs below, and we will discuss the special status of S in chapter 3.

The small grammar (31) allows us to analyse sentences like (32); heads are in boldface in the syntactic tree:

(32) A rather large dog left a bone on the carpet



In addition to generally being obligatory, the head of a category has other privileged characteristics. The semantic, syntactic and morphological properties of the head usually become the properties of the whole phrase. If we consider the following examples

(33) a. the tall man
b. the happy woman
c. the brown dog
d. the intelligent dog

(34) a. the brown dog
b. the red car
c. the white sugar
d. the blue sky

all nouns in (33) refer to animate nouns, therefore the whole NP can be said to be animate. On the other hand, although all the adjectives in (34) are colour adjectives, the NPs can hardly be said to be colour NPs or something similar.

2.1.1 A Universal Rewrite Rule

Except for the category S, every rewrite rule of grammar (31) states that a phrase consists of at least an element of the same category as the head, and possibly more. The fact that every phrase is headed by a category of the same type as the whole phrase is one of the cornerstones of X-bar theory:

GB Principle 9 (X-bar): a phrase always contains a head of the same type.

We can suggest a universal rewrite rule, some kind of a skeleton for all rewrite rules, a first version of X-bar theory. X stands for any category, e.g. N or V or whatever.

(35) XP → ... X ... [where ... contains complements etc.]

Rule skeleton (35) will still show considerable variation in the individual rewrite rules that will be fashioned for the grammar. We have seen that, except for PP, all non-head elements are optional. Let us take a closer look at PPs:

2.1.2 Transitive and Intransitive Prepositions

It is obvious that many prepositions require a noun complement, as the following ungrammatical sentences illustrate:

(36) *She ran into.

(37) *He cut the bread with.

However, there is a category of words that are very closely related to prepositions, the so-called verbal particles, which do not require prepositions.

(38) a. He turned the lamp on.
b. She put the book down.
c. Sue put her hat on.

Most of the words that can be verbal particles also belong to the lexical category of prepositions. It has been suggested that verbal particles are indeed prepositions, although prepositions that do not take complements. Verbs without complements are called intransitive verbs. It is therefore natural to suggest that there are *intransitive prepositions* (formerly called verbal particles) and *transitive prepositions* (the former classical prepositions). As with verbs, some prepositions are always transitive, others always intransitive, many can be both:

(39) Intransitive only:
a. She threw the book away.
b. *She threw the book away the window.

(40) Transitive only:
a. *She ran into.
b. She ran into the house.

(41) Both transitive and intransitive:
a. Sue put her hat on.
b. Sue put her hat on the shelf.

If we accept the hypothesis that verbal particles are prepositions[1], we can also use a phrase structure rule in which the NP is optional, and the complement requirements for prepositions will be listed in the lexical entries of each preposition, as we have suggested for verbs in 1.3. All the complements are now projected up from the lexicon, the only non-head elements left in the phrase structure rules are entirely optional, they only refer to optional extensions, modifiers and adjuncts. The universal rewrite rule from (35) can now be reformulated:

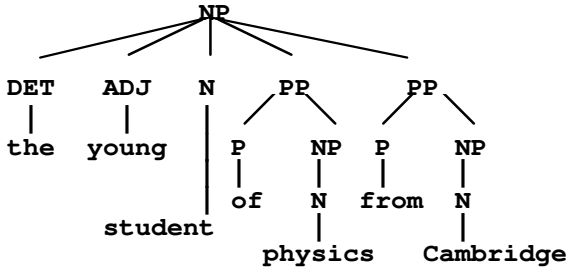
(42) XP → (...) X (...) [where all (...) are optional or projected up from the lexicon]

[1] See Cowper (1992: 23 ff) for additional argumets, but also counter-arguments to this hypothesis.

2.2 Intermediate Categories

The phrase categories we have used until now have two levels of structure: the level of the entire phrase (XP), and the level consisting of the head (X), complements projected up from the lexicon and optional elements. Let us take a look at the representation for the following sentence:

(43) the young student of physics from Cambridge.

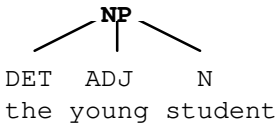


The NP splits up many times at the same level, without any possible distinction between the different elements. This is a so-called **flat structure**. It has two disadvantages, however, as I shall explain now.

2.2.1 First Disadvantage: *One*-Pronominalisation

We have seen in 1.1.2 that one of the tests used to find constituents is the pronominalisation test. Only entire constituents can be pronominalised.

(44) The young student



In (44), the only constituent, therefore the only candidate for pronominalisation, is the NP.

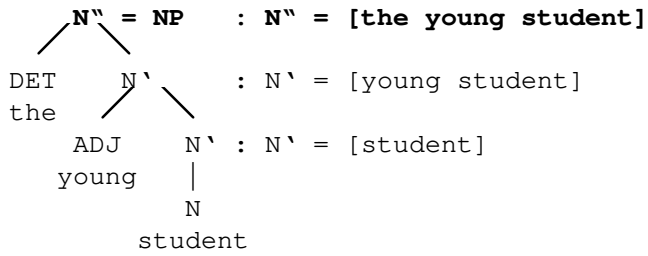
(45) I like the [young student].
I like [him].
***I like the [him].**
***I like the young [him].**

There is one pronoun, however, which allows to replace these elements – the pronoun *one*:

(46) I like the [young student].
I like [one] as well.
You mean this [one] !
I prefer the old [one].

Either we have to give up the pronominalisation test, or we are forced to admit that the pronominalised elements are also constituents. In the representation (44) they are not. GB has chosen to keep the pronominalisation test and to assert that pronominalised elements in (46) are indeed constituents. The pronominalised elements are *[student], [young student]* and *[the young student].* We know that *[the young student]* is an NP, but *[student]* and *[young student]*have to be something between N and NP. We have three N-levels now: We will call them **N0** or **N**, which is the lexical category, **N1** or **N‘**, which is the new intermediate level, and **N2** or **N‘‘**, which is the former XP.

(47) the young student



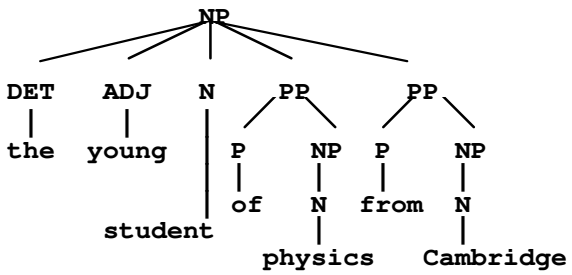
As we can see, the intermediate category N‘ is allowed to repeat itself several times, e.g. when several adjectives precede a noun. We have seen in 1.1.6 that this is called **recursion**. The recursion of N‘ in (47) corresponds to the following rewrite grammar rule:

N‘ → ADJ N‘

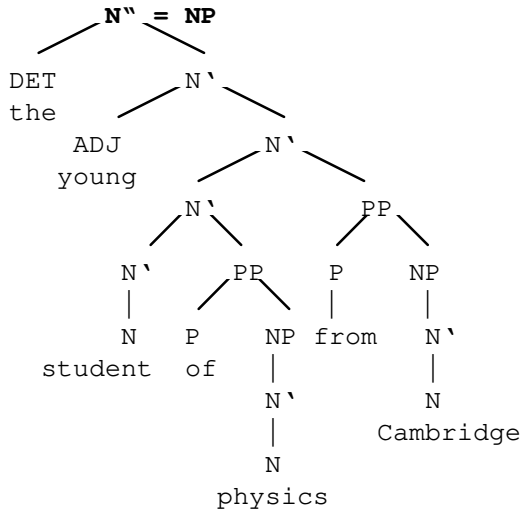
There are some notational variants for N‘ and N‘‘. The original notation was to write a bar, or two bars respectively, on top of the N, or whichever other category. Again, the universal phrase category label is X. The pronunciation for X‘ is still **X-bar**, for X“ it is **X-double-bar** – hence the name **X-bar syntax**.

2.2.2 Second Disadvantage: Arguments and Adjuncts

The two PPs at the end of the NP in (43) (repeated below for clarity) are both at the same level, so they cannot be distinguished.



If we adopt the new notation, as suggested below, they are at different levels, which means that they can be *one*-pronominalised. But as they both have an N‘ as mother node and sister node they can still not be distinguished from each other:



We should expect, therefore, that they have equal status, that they can be deleted and replaced at will.

(48) a. the young student
b. the young student of physics
c. the young student from Cambridge
d. the young student of physics from Cambridge
e.*the young student from Cambridge of physics

For some reason, the PP *of physics* has to stay closer to the noun than the PP *from Cambridge*. We have seen in 1.3.3 that some nouns tend to subcategorise for PPs, as if they were *transitive*. They are called *relational nouns*. In all of them, the subcategorised PP has to occur closer to the noun than additional PPs. Here are some examples for the nouns *chairman*, *piece* and *relation*.

(49) a. the chairman of the board in the blue suit
b.*the chairman in the blue suit of the board
(50) a. the piece of cake on a plate
b.*the piece on a plate of cake
(51) a. the relation of verbs to nouns in our discussion
b.*the relation in our discussion of verbs to nouns

A chairman always has to be a chairman of something, a piece is a piece of something, a relation is always a relation of something to something or a relation between two things, much as whenever you kick, you kick something, or when you love, you love somebody – transitive verbs, transitive prepositions and relational nouns all subcategorise for complements. These subcategorised elements are also called **arguments**. They are usually obligatory, but can be dropped in many cases:

(52) a. He has spent hours kicking (the ball)
b. this relation (to nouns) is a difficult issue
c. Loving (somebody) sets you free.

The optional additional elements which are neither heads nor complements (all complements are projected up from the lexicon) are called **adjuncts**. They can be modifiers, PPs etc. They have much less influence on the semantics of the head than arguments, and they are always further away from the head than arguments, as we have seen for nouns in (49). An example for verbs:

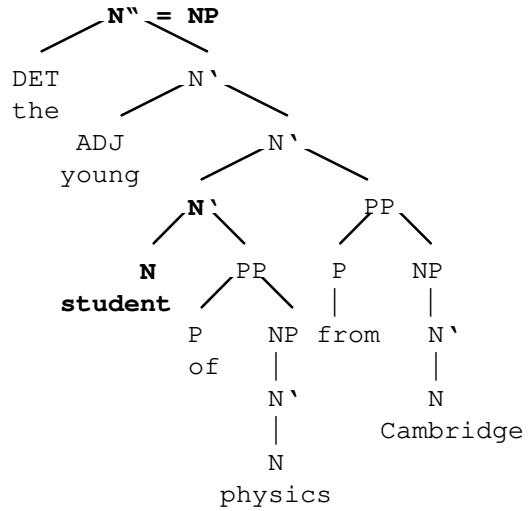
(53) a. She depends on her friends for good company.
b.*She depends for good company on her friends.

How can we distinguish between arguments and adjuncts in X-bar theory ? The X-bar suggestion is that adjuncts are attached *under recursion to X',* i.e. with X' both as a mother node and sister node, while arguments are attached when X' is rewritten to X0:

(54) X-bar Adjunct Rule: X' → Y X' [where Y is the adjunct]

(55) X-bar Argument Rule: X' → (Y)(Z) X [where Y and Z are the arguments, as subcategorised for by the lexical entry of the head, and projected up to the syntax]

When we apply these rules, we get the following representation for *(43) the young student of physics from Cambridge.*(with the crucial differences in boldface):



While recursion, and therefore the admission of new adjuncts, may go on forever, argument positions may only be occupied once. We can easily imagine sentences with many adjective and PP adjuncts:

(56) The young handsome clever bright student from Cambridge with a raincoat standing under the porch

But a recursion of arguments, i.e. complements, is impossible, or it conveys a different fact from what we intended:

(57) a. *She loves Peter Mary.
b. *Cats live on meat on milk.
c. *A student of physics of maths
d. *The relation of verbs of adjectives to nouns

An attentive reader may wonder why the adjunct rule (54) has an obligatory adjunct Y, instead of bracketed and hence optional one. The problem would be that then the rule could enter endless recursion without ever taking up an adjunct. Because the rule is recursive it is the rather rule itself which is optional, instead of its elements. Recursion will only take place if more adjuncts need to be placed in the syntactic structure. In an adjunct-free sentence, the adjunct rule will never apply.

2.3 The X-bar Rule Schemata

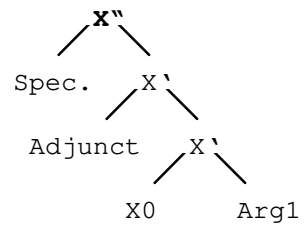
For every phrase X, only following three rewrite schemata, skeletons of former rewrite rules, remain:

GB Principle 10(X-bar): X-bar Rule-schemata:

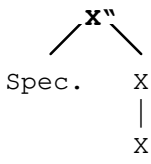
Specifier Rule:	**X`` → (Specifier) X`**
Optional Adjunct Rule:	**X` → Adjunct X`**
Argument Rule:	**X` → (Arg1) (Arg2) X**

Specifiers are articles in NPs, in ADJPs or VPs they may be adverbs. Nothing has been said about *the order* of the constituents in the rules. It is assumed that this should be a language-dependent parameter. In English, e.g., specifiers come before X‘, but arguments after X. An English X-bar phrase can e.g. take one of the following shapes:

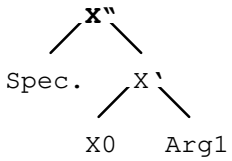
one adjunction, one argument:



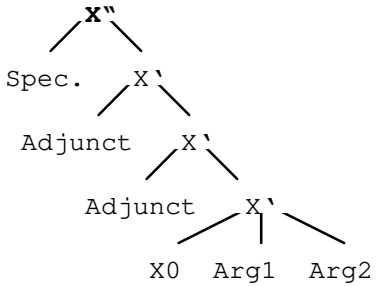
no adjunction, no argument:



no adjunction, one argument:



two adjunctions, two arguments:



etc. etc.

Adjuncts are inserted by recursion of the adjunct rule, as required by the sentence to be analysed. Arguments are present or not as required by the lexical entry of the head. Arguments in turn are usually phrases, the head of the phrase introduced by this argument has new arguments, etc. In this fashion, the representation for an entire sentence is built up. X-bar depends heavily on the lexicon and the projection principle we have met in 1.3.2. X" are also called **maximal projections** to indicate the lexicalist character of GB. The rule schemata that are left in GB cannot be called rules any longer, they are too schematic, and their whole purpose it to direct the lexical projections. They restrict the possible lexical projections. For this reason, X-bar is also called a **structural constraint**. X-bar is the only structural constraint on the syntax in GB, everything else is lexical projection.

When not needed, the intermediate level is often eliminated in syntactic representations for the sake of legibility, and the terms XP and X" are used interchangeably.

Specifier nodes often remain empty, and as we shall see, they often make good landing sites for elements under movement. We shall see in chapter 4 that mainly heads and specifiers can move. We will also see that specifiers move to specifier positions, and heads to head positions.

3 Functional Categories

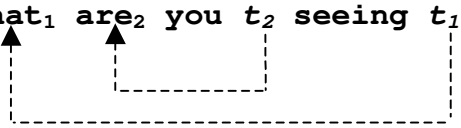
3.1 Some Detective Work

NO READING SUGGESTIONS for Subchapter 3.1

At the current stage of our progress in understanding GB, we have a couple of mysteries and indications which we would like to solve:

- In 1.4.1 we have seen that movement will always be structure-preserving, i.e. that D-structure (deep structure) and S-structure (surface structure) of a given sentence are the same, except that movement has taken place in the surface structure, as in (26), repeated for clarity:

(26 S) What$_1$ are$_2$ you t_2 seeing t_1 at the cinema ?



We have postulated that because movement is structure-preserving, the two empty positions at the beginning of the sentence to which movement occurs, have to be provided for at D-structure. How can we account for these empty positions ?

- In chapter 2 we have not been able to include the topmost rewrite rule into the X-bar scheme. We have still used this rule:

  S → NP VP

  But we would like to use a uniform X-bar representation for all constituents. What could be the head of a sentence?

- Arguments in English are positioned after the head, and therefore after the specifier, according to X-bar theory. If the NP or the VP under the hitherto S-node could be conceived to be arguments of something, this something would be positioned on the left of the subject and the main verb – at the very beginning of the sentence, exactly where we expect empty positions at D-structure to be found!

Such argumentation is of course not properly scientific, but it may give us an indication of where we will be going. Let us try to solve the questions of what the S category is, and what its head could be.

3.2 The Head of a Sentence

RECOMMENDED READING for Subchapter 3.2: Cowper (1992: 64-70)

3.2.1 NP versus VP

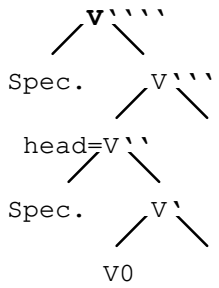
Assuming that a sentence consists of an NP and a VP, which could rather be its head? Logicians and computational linguists assume that the verb is the top head of a semantic representation. They take the verb to be predicate, and the noun to be its argument. The sentence *Peter laughs* is e.g. represented in predicate logic as:

(58) laugh(Peter).

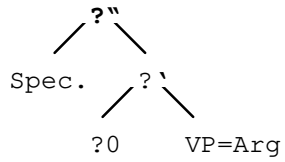
Many highly inflectional languages, e.g. the Romance languages Latin or Italian, drop the subject if it is a personal pronoun. This phenomenon is called **pro-drop**. The identity of the subject is already well-marked on the verb by the inflection, so why repeat it? GB assumes that every sentence has a subject, in a pro-drop language it is an empty category, however. We will see that GB uses many empty categories. But the empty category could indicate that subjects are perhaps less suitable as heads than the VP, or a part of it. From English we know the famous existential *it* and *there*. They act as mere syntactic dummies and have no referential meaning, but they have subject function. Their existence also speaks against selecting the NP as the head.

3.2.2 V-Triple-Bar ?

If we chose VP, i.e. V“ to be the head of S, we would run into new difficulties, however. Usually, a zero-bar constituent, i.e. X0, is a head. If X“ was used as a head we would be forced to assume V-triple-bar and V-quadruple-bar in order to abide to X-bar conventions, in which maximal projection is two bar-levels higher than a head. This would seem a bold step.



As we wanted to find a head for S in order to include it into the X-bar syntax I will not pursue this idea. If we want to abide to usual X-bar syntax, VP will rather be the argument of a new unknown head represented by ‚?‘ in the following X-bar schema:



We would of course run into the same problem, i.e. needing N-triple-bar, if we chose the subject NP to be the head of a sentence.

3.2.3 What is S ?

We remember from 1.1.6 that some verbs take clausal complements. An example sentence and the suggested rewrite rule for its VP were:

(59) Sandra knows that Peter loves Mary

(60) VP → V *„that"* S

Let us consider the following sentences (taken from Cowper 1992: 67), where (that)+S is bracketed:

(61) a.*Sue wanted [that John would leave early].
b. Sue wanted [John to leave early].

(62) a. Alan insisted [that Joe had washed the dishes]..
b.*Alan insisted [Joe to wash the dishes].

(63) a. George expected [that Sue would buy the car].
b. George expected [Susan to buy the car].

Apparently, some verbs select finite clausal complements, others non-finite ones, some accept both. All non-finite subordinate clauses inherit information from the matrix clause, be it the time frame or even the subject, as in (64b), a so-called *control-structure*, in which Sue is also the subject of the subordinate clause. We will see an example of a complex control structure in 4.1.3.

(64) a. Sue wanted [John to leave early].
b. Sue_1 wanted [t_1 to leave early].

The non-finite subordinate clauses would be ungrammatical on their own. What is it actually that makes up a sentence? What is the „essence" of sentencehood? Let us consider these sentences:

(65) *Peter love Mary.

(66) *Any clever student to go to Cambridge.

Both sentences have a (subject) NP and a VP, but something is missing – the inflection. Until now we have had the impression that GB does not integrate

morphology or that it takes morphology for granted. That is not the case, however. For GB, loosely speaking, a sentence minimally consists of an NP, inflection (called INFL) and a VP.

3.2.4 Inflection as a Functional Category

In a theory which is as consistently structurally oriented as GB, even morphological markers like INFL have a category of their own, a so-called functional category, opposed to the more usual lexical categories. Other well-known functional category are DET (determiner, article) and prepositions. Functional categories are closed lexical categories and have no referential meaning, i.e. they do not denote any real object in the outside world. While prepositions and determiners may at the first sight seem better candidates for categories than morphological inflection, there are counter arguments. First, there are languages in which prepositions or determiners are morphological. As for prepositions, highly inflectional languages like Latin or Finnish cover a wide range of English prepositions by means of case-inflection. The English preposition *of* as opposed to the genitive *'s* illustrates that case-inflection and prepositions are virtually identical in function. As for articles, the Scandinavian languages are especially interesting, as they sometimes attach a morphological article to the noun, but when adjectives precede the noun the article is realised as a lexical category in front of the noun, as is usual in English:

(67) Mannen stannar <=> Den stora Mann stannar
man-the stops the big man stops

Secondly, isolating languages do not mark tense, number, aspect etc. morphologically, but (sometimes !) by functional lexical categories.

Pre-X-bar versions of GB have often used one of the following top rewrite rule:

(68) S → NP AUX VP

(69) S → NP INFL VP

Especially in Germanic languages, such a rule seems intuitive as a part of the verb, often an auxiliary, but always the inflected part of the verb, appears at the second position in the sentence.[2] Germanic languages are often referred to as verb-second languages. We will hear more about verb-second in the discussion of my article in chapter 5. The distance between the inflected auxiliary and the main verb is often considerable, as these English and German examples illustrate (AUX underlined, *Verb* in italics):

(70) I have often and repeatedly, but never successfully *warned* you.

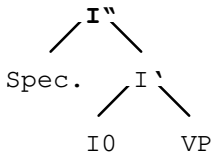
[2] It is arguable if verb-second also applies for yes/no-questions like *Are you reading ?*. One can either say that yes/no-questions are verb-first, or that the first position remains empty, and the verb is still at the second position.

(71) Das Volk hat die Abstimmung über diese sehr umstrittene Initiative an diesem Wochenende mit deutlichem Mehr *abgelehnt*.

In short, as inflection and the second position in the clause, which is either filled by an auxiliary or the main verb are so closely tied together it is well justifiable to have a structural category for an essentially morphological process, at least for Germanic languages. It was of course quite bold to postulate typically Germanic characteristics in a theory of Universal Grammar (UG) by rules (68) or (69).

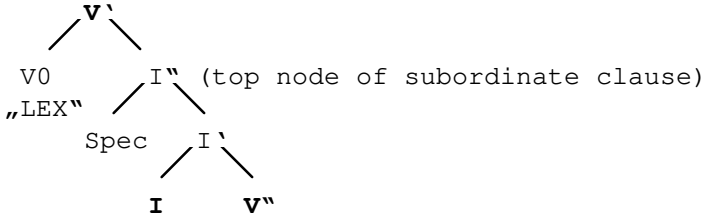
3.2.5 INFL as the Head of the Sentence

INFL, later abbreviated to **I** (I will use both interchangeably) was indeed to Germanic at its second position in the S rewrite rule. But an S rewrite rule does not fit the X-bar scheme anyway. Therefore, it was suggested that I should be the new head of VP, IP therefore the new top-node of the sentence. I is the „unknown" head we have been looking for in 3.2.2:



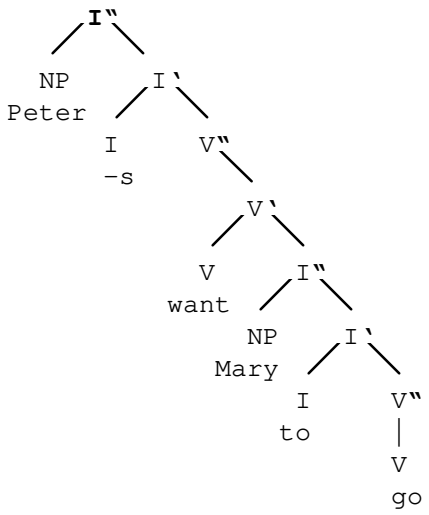
Unlike in the quite Germanic claim of rules (68) or (69), the claim is now only that any sentence has to be finite; that finiteness, i.e. inflection for tense, aspect, person, number etc. is the touchstone of sentencehood.

For those who are still unconvinced if INFL should be the top head of a sentence there is also a more technical argument. We have seen in 3.2.3 that some of the verbs requiring clausal complements take finite clauses, while others take non-finite clauses. The lexical entry (say *„lex"*)subcategorises for a sentence with finite or non-finite INFL respectively. But subcategorisations only have a locally limited validity: „subcategorization statements ... refer only to the heads of categories in the immediate context" (Cowper 1992: 66). INFL carries the information whether the subordinate clause is finite or non-finite, if INFL was a non-head or a head far below in the syntactic structure, this information would not be accessible to the verb that introduces the subordinate clause. Under the assumption that IP is the top node INFL is easily accessible:



A non-finite INFL contains the infinitive marker to, as the following example illustrates:

(72) Peter wants Mary to go

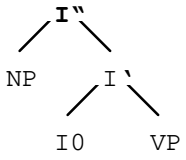


We shall see in chapter 4 that want actually moves up to I (and remains there) in order to take up the inflectional affix and to assure the usual verb-second position.

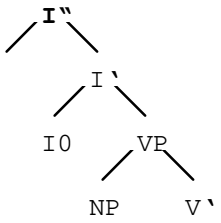
3.2.6 The Subject under the IP assumption

If we accept the suggestion that IP is the new topmost phrase of every sentence, where can we place the subject NP, formerly the sister of VP in the old $S \rightarrow NP\ VP$ rule?

First we need not forget that we only want to speak about the D-structure place of the subject NP, because, as we shall see in chapter 4, it may move to other positions. Until few years ago, subject NPs were suggested to reside in the empty specifier position of IP, as just seen in (72):



More recently, however, the subject NP has been suggested to originate from inside the VP, from where it moves up to I-Specifier even in most declarative sentences.



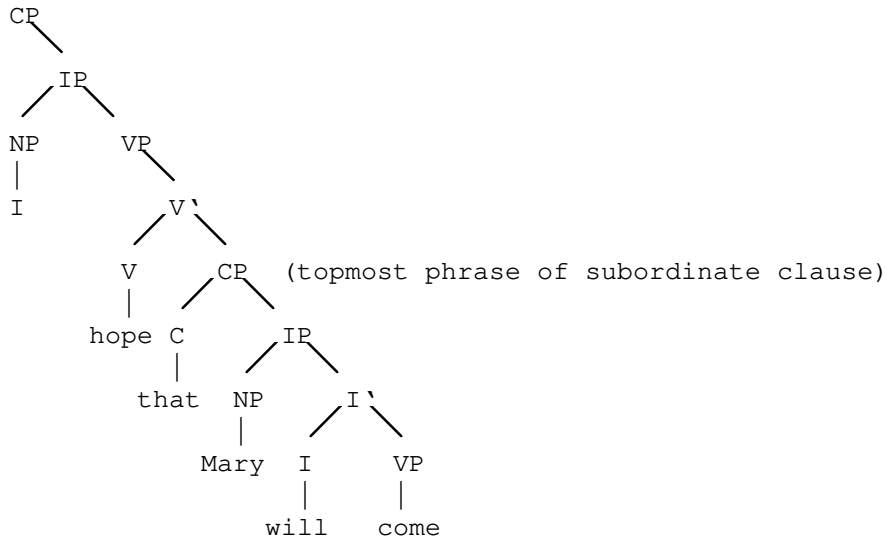
This hypothesis is known as the *subject-within-VP hypothesis* which we shall see in 4.

3.3 Complementizers

NO READING RECOMMENDATIONS for Subchapter 3.3

After allowing the functional category I to be a head, many other categories followed. Complementizers are words like *that, if* or *whether*, which introduce finite subordinate clauses. A complementizer phrase (CP) introduces an IP, it is therefore a new topmost phrase. Let us look at an example sentence – all unnecessary intermediate levels are eliminated, and it is the D-structure representation:

(73) I hope that Mary will come.



This sentence structure illustrates that sentences start with a CP, which usually consists of at least an IP, which at least consists of a VP, which consists of whatever the lexicon selects, e.g. an object noun or (as in (73)) a subordinate clause. The cascade from CP to IP to VP is not prescribed by X-bar, it follows entirely from the lexicon. A complementizer (even an empty one as in (73)) naturally subcategorises for an IP. An inflection I naturally subcategorises for a verb. A verb has more idiosyncratic subcategorisational requirements, as discussed. This is the stage of the theory which I will mainly use.

3.4 Other Functional Categories

NO READING RECOMMENDATIONS for Subchapter 3.4

Recently, many linguists have started to use DP (determiner phrase), which heads and subcategorises for an NP. There have also been suggestions to split up IP into several functional phrases. IP marks both tense, and agreement to subject, and agreement to object, so that e.g. splits into tense-phrase, agreement-subject-phrase and agreement-object-phrase have been suggested. I did not want to discuss these considerably technical issues in an introductory course, however.

It is indeed questionable if functional categories, many of them bordering on the line between syntax and morphology across languages, should be heads. Some linguists prefer to treat them as so-called markers, which are headed by the lexical category they mark. The authors of a rival theory to GB named Head-Driven Phrase Structure Grammar (HPSG) criticise that many of the arguments brought forth in support of functional heads „depend on GB-internal assumptions“ (Pollard & Sag 1994: 44) like X-bar theory.

> On our account, a marker is a word that is 'functional' or 'grammatical' as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous). A marker, so-called because it *marks* the constituent in which it occurs, combines with another element that heads that constituent. In addition to the complementizers *that* and *for*, other examples of markers include the comparative words *than* and *as*, the case-marking post-clitics of Japanese and Korean, and perhaps nonpredicative adpositions in (the vast majority of) languages where adpositions stranding does not occur.
>
> (Pollard & Sag 1994. 44-5)

4 Movement and Case Theory

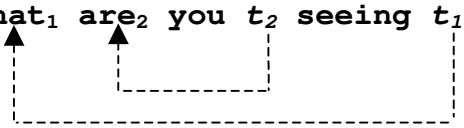
4.1 A Detailed Example of Movement

RECOMMENDED READING for Subchapter 4.1: Radford (1997: 216-230)

In the introduction to movement in 1.4 I have shown how the S-structure of a question is derived from its D-structure, the latter being almost identical to both the D-structure and the S-structure of the corresponding declarative sentence. I have said that GB postulates that movement is structure-preserving, i.e. that the positions to which elements can move have to be provided for at D-structure (I have called them *E1* and *E2* for convenience, but this is not GB terminology). Let consider the examples again, the D-structure (26D) and its S-structure (26S):

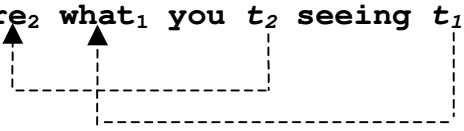
(26 D) ***E1 E2*** **You are seeing what at the cinema ?**

(26 S) What$_1$ are$_2$ you t_2 seeing t_1 at the cinema ?



We have also seen that the places where elements move to, the so-called landing sites, have to be compatible to the element under movement. The following S-structure is ruled out:

(26 S)*Are$_2$ what$_1$ you t_2 seeing t_1 at the cinema ?



The simplest idea to ensure such compatibility is to postulate that elements should only move to positions where they typically occur in X-bar, i.e. specifiers to specifier positions, head to head positions etc. This will indeed be the case.

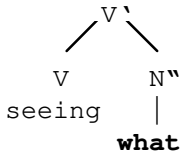
4.1.1 Building Up the D-Structure

Since we know the principles of X-bar now, we can imagine how the D-structure of (26) must look like.

(26 D) ***E1 E2*** **You are seeing what at the cinema ?**

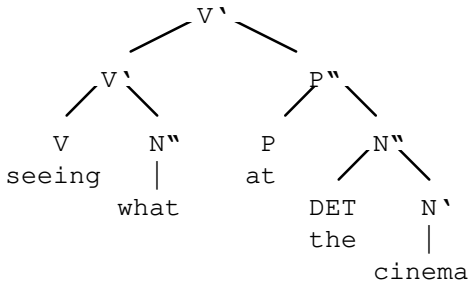
Let us start with the VP *seeing what at the cinema*. *What* has to be the subcategorised object, i.e. an argument of *seeing*. The X-bar argument rule has to apply:

(74) X-bar Argument Rule: X' → (Y)(Z) X [where Y and Z are the arguments, as subcategorised for by the lexical entry of the head, and projected up to the syntax]

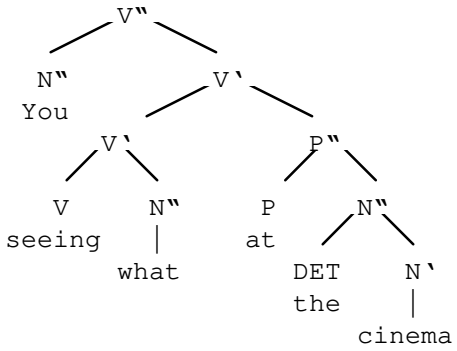


But *at the cinema* will be an optional adjunct, the X-bar adjunct rule is therefore used next:

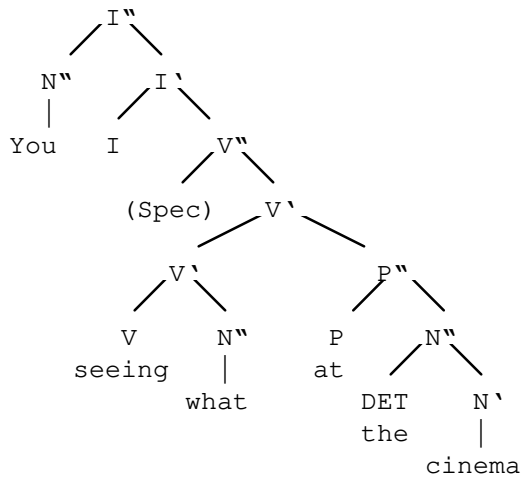
(75) X-bar Adjunct Rule: X' → Y X' [where Y is the adjunct]



The VP is almost complete now. If we want to follow the novel subject-within-VP hypothesis, we attach the subject NP to the VP-specifier position in the D-structure:

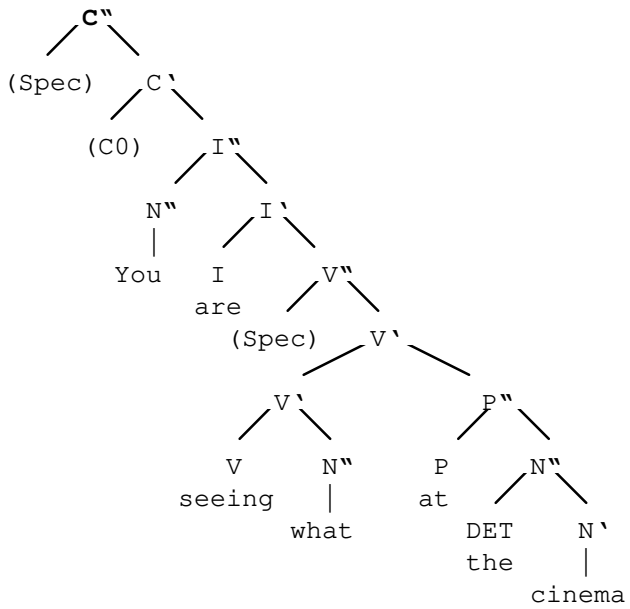


In the more classical GB assumption, the D-structure subject NP resides in the IP-specifier position, leaving the VP-specifier empty:



Many textbooks still use the old version, but at least Cook & Newson (1996) and Radford (1997) already use the new subject-within-VP hypothesis. There are no far-reaching consequences for the theory if we adopt one assumption or the other.

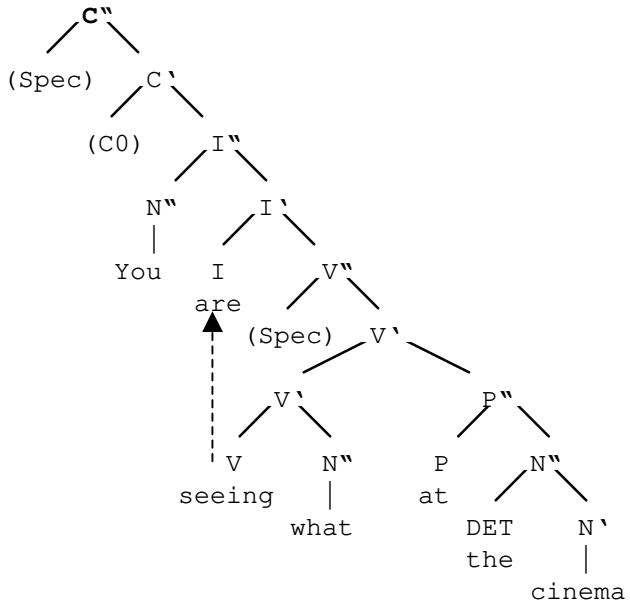
The inflected auxiliary *are* is still missing in this tree. The most natural place to put it is of course I0, the inflection-bearing head of the sentence. We are almost finished now. We only have to account for the empty positions (which I have called *E1* and *E2*). Movement is structure-preserving, and subordinated clauses, which are introduced by *that* or another complementizer, have the same D-structure as matrix clauses. Accordingly, the CP category has to be provided for in D-structure:



Except for one simplification which I shall immediately explain, this is indeed the correct D-structure, from which the S-structure can be derived, as we shall see now.

4.1.2 Moving to S-Structure

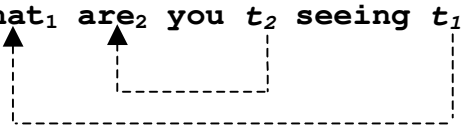
The one simplification I have made is that I have written an inflected auxiliary verb at I0 in the D-structure. This is not entirely correct. Only the inflection resides in I at D-structure. I is a functional category without lexical content, so what we find in I at D-structure is in fact a set of inflectional features, in our example [PERS: 3], [NUM: sing], [TENSE: pres], [ASPECT: progressive]. V0 only contains the base form of the verb. The part of the verb which carries the inflection, i.e. an auxiliary, or, if the main verb is inflected, the main verb itself. Let us express this first movement in the following representation:



If movement is only upwards how can we make sure that the part of the verb remaining in V is the progressive gerund *seeing* in our example, and not e.g. the base form *see* or the past participle *seen*? As GB postulates that movement should be upwards only, how can the features [TENSE: pres], [ASPECT: progressive] reach down to V0? The answer is subcategorisation. I subcategorises for a verb head which meets the required features. We will come back to this issue in Checking Theory in 6.1.

After this one movement we still have to account for the fact that the question word what and the inflected verb move to the beginning of the sentence, as predicted:

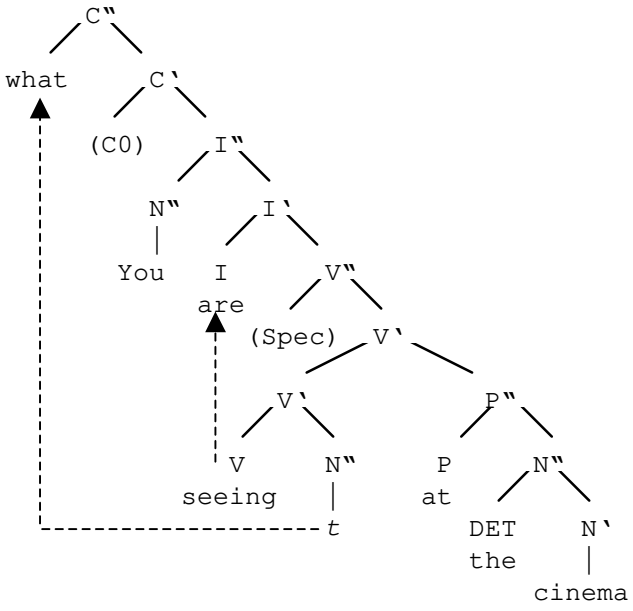
(26 S) What$_1$ are$_2$ you t_2 seeing t_1 at the cinema ?



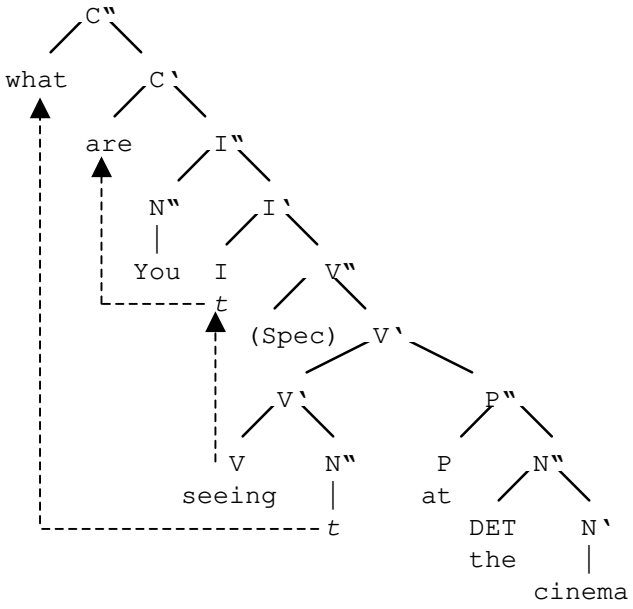
We can see that the D-structure representation for our example (and in fact every other sentence) provides two empty positions at the beginning of the sentence, under the CP node, C-specifier and C0. The two elements to be moved are quite different in character: The question word *what* is rather a maximal projection (X") than a single

word, as it corresponds to a N“ in an declarative sentence. The auxiliary verb is a head, i.e. X0.

We have seen in our discussion about the head of the sentence in 3.2.2 that allowing maximal projections to appear in head positions would lead to very considerable problems for X-bar theory, as the maximal projection X“ would have to continue projecting up to X‘‘‘ and X‘‘‘‘. We would run into these problems if we allowed the N“ what to select C0 as a landing site. But C-specifier can easily admit maximal projections. Although object NPs do not originate in specifier positions at D-structure, NPs are often found in specifier positions. The subject NP, e.g., originates either in the IP-specifier, or in the VP-specifier (if we follow the subject-within-NP hypothesis). The C-specifier position is therefore a perfect landing site for the question word *what*:



The auxiliary verb originates in a head position, it is therefore natural that it will also move to a head position. The only free head position available upward in the structure is C0:



Done! The S-structure is now complete, and we have acquired a fair impression of how movement works.

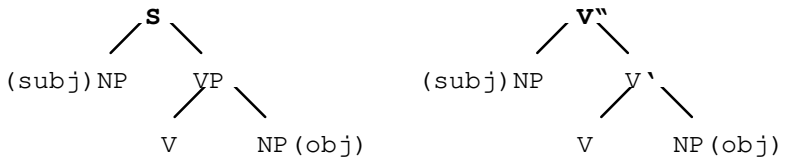
Let us take quick look at the corresponding declarative sentence, i.e.

(27 D) *E1 E2* You are seeing „When Harry met Sally" at the cinema.

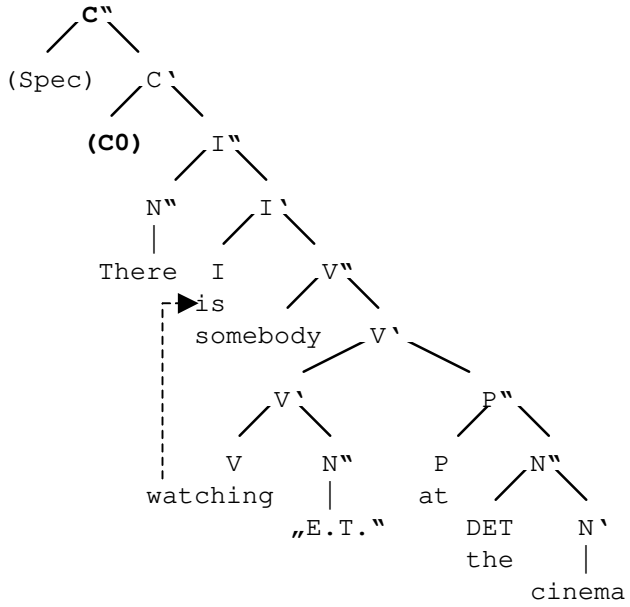
If we assume the classical analysis, in which the subject NP is in I-specifier at D-structure, no NP movement takes place between D- and S-structure. The only movement necessary is the one of the inflected part of the verb to I0. If we assume the subject-within-VP analysis, the subject NP will usually move up to I-specifier at S-structure. The advantages of this hypothesis are:

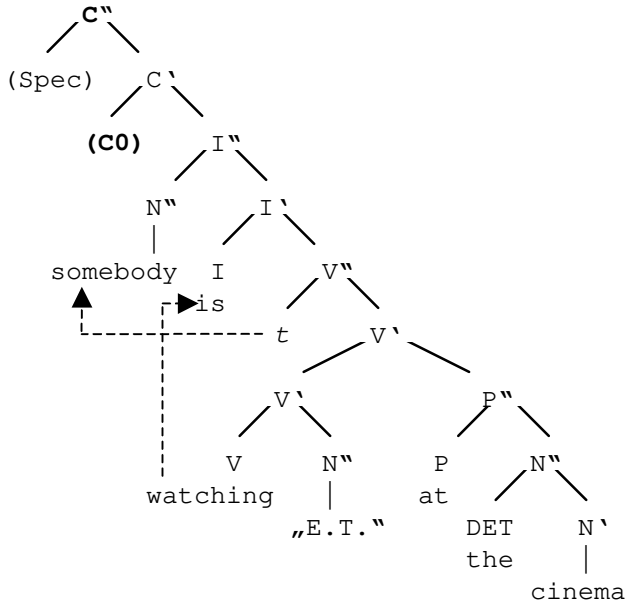
- all lexical elements reside inside the VP at D-structure, so that the entire verb, not only its inflected part, dominates the rest of the sentence, including VP. This view is closer to semantic representations in predicate logic, where both subject and object are arguments of the verb predicate, e.g. *Peter loves Mary* is *loves(Peter, Mary).*
- Verbs place strong restrictions on their objects and lesser restrictions on their subjects. We have seen in 1.3.1 that verbs select the categories of their complements but not of their subjects. Nevertheless, verbs can make semantic restrictions on their subjects. E.g. only animate NPs may love or laugh. (This is one of the topics of theta-theory, which I will not cover). A head has full access to its arguments (i.e. complements) and limited access to its adjuncts and its specifier. The subject-within-VP hypothesis suits nicely in this pattern.

- The ancient rewrite rule shape dominated by the S-node is mirrored in this assumption.



- The structural relation between sentences with or without expletive *it* or *there* can be expressed. Since elements can only move to empty positions, movement is simply blocked by an expletive element on the landing site:





We will look at an application of movement from comparative linguistics in chapter 5.

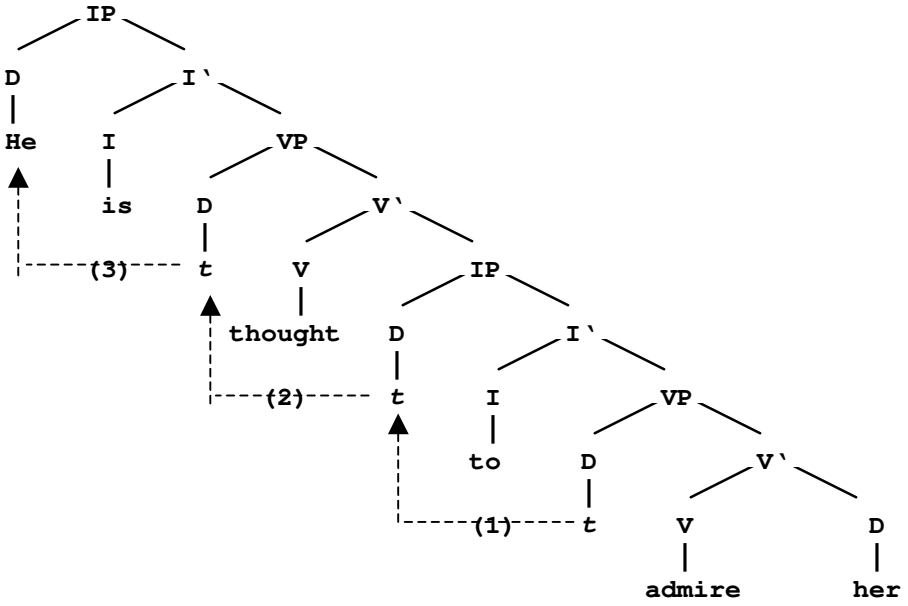
4.1.3 Chains

A moved element can easily move again. An auxiliary in a question typically moves up from V0 to I0 to C0. We have seen in (64) that in *control* structures (the example is repeated below), a subordinate clause and a matrix clause share the same subject.

(76) Sue$_1$ wanted [t_1 to leave early].

In complex control structures, an element may move through a big number of intermediate landing sites until it reaches its S-structure position. Because all the intermediate positions may need to access to the element that has moved to it and away again (e.g. when building the semantic representation LF) the moving element leaves traces in all positions.

(77) He is thought to admire her.



4.2 Case Theory

RECOMMENDED READING for Subchapter 4.2: Cook & Newson (1996: 222-233)

4.2.1 Restricting Movement

Movement has to be restricted in several ways. We have already seen that it is only allowed to proceed upwards, that landing sites need to be compatible etc. Still, these restrictions are not restrictive enough yet. For example, they could not prevent elements from moving up randomly to the empty positions in CP.

Movement has to be restricted so much that it only happens when something triggers it. A trigger may e.g. be that the sentence is a question. Then the empty CP positions ask to be filled, and the appropriate elements, i.e. an auxiliary verb (in

English, in many other languages like French, German etc. the whole verb can move in questions) and the question word are allowed to move up. There are other such types of sentences that trigger certain transformations, e.g. passivisation.

But what about the movement from VP-spec to IP-spec which is supposed to take place even in declarative sentences ?

4.2.2 Abstract Case

Unlike inflectional languages, English has a very poor case system. But since GB aims to be a part of a Universal Grammar, a theory of Case was adopted, irrespective of whether Case is covert and abstract like in English or overtly marked.

We have seen that morphological processes of inflection are dealt with by stipulating a functional category INFL, which finally ended up being the head of the sentence. Case is realised in a different way. It is firmly assumed to be a morphological process. We have seen that I places morphological restrictions on V by means of **features,** such as [NUM: pl] or [-INFL]. the feature for e.g. accusative case is then [CASE: acc], etc. Selection of Case is at least partly idiosyncratic. Consider the German verbs *helfen* and *unterstützen.* They have virtually the same function and semantics, but the former takes a dative object, while the latter takes an accusative object:

(78) Ich helfe Dir.

(79) Ich unterstütze Dich.

GB distinguishes between **inherent** and **structural** Case. Inherent Case is subcategorised for by a lexical entry, structural Case is given to certain positions in the sentence structure. While most verbs take accusative objects, less take dative objects. Dative is an inherent Case.

4.2.3 Inherent Case

The lexical entry for *helfen* is therefore, at least as far categorial and Case selection is concerned:

(80) helfen Verb, [_ NP[CASE:dat]]

Mainly verbs and prepositions subcategorise for cases, but there may be other categories. Genitive case is e.g. often selected by a noun. In *Peter's house* the noun house can subcategorise for a proper noun.

Inherent Case is assigned at D-structure. If an element moves it takes its Case along, but it may not move to positions requiring a different Case.

4.2.4 Structural Case

There is a GB principle which ensures that every S-structure NP has Case, the so-called Case Filter (*phonetically realised* simply means at S-structure for our purposes):

GB Principle 11 (Case-Filter): Every phonetically realised NP must be assign (abstract) Case.

Those NPs which have not received inherent Case at D-structure will receive structural Case at S-structure. V and P structurally assign accusative Case. But how should nominative Case be assigned?

The claim that every sentence has an inflected verbal element is closely connected to the claim that most sentences have a nominative subject. It is therefore natural to suggest that INFL assigns structural nominative Case. If a subject has not received inherent Case at D-structure, it will receive structural nominative Case at S-structure.

If we accept the subject-within-VP hypothesis, the case filter forces the subject to move up: It will only receive Case in a position to which I0 has access, i.e. IP-specifier (or I0, but a maximal projection cannot move there as we have seen). I have mentioned briefly that subcategorisation selection only has a limited domain of validity. A head (and all lexical entries are heads) can only subcategorise for arguments, and sometimes specifiers. But the subject cannot receive Case when it remains in the VP-specifier position.

Some subjects have already received inherent Case at D-structure and will not take structural Case by I0. Consider the German sentences

(81) Mir ist kalt hier.

(82) Es ist mir kalt hier.

Kalt assigns inherent dative Case to the pronoun, which it keeps when it moves up to the subject position in (81). The expletive *es* in (82), on the other hand, which did not receive any Case at D-structure, takes structural nominative Case at S-structure.

It seems, however, that accusative may also be inherent rather than structural Case sometimes. Consider another German example:

(83) Mich friert.

(84) Es friert mich.

Frieren seems to assign inherent Case in much the same way as *kalt* did.

5 An Application of GB: V2 in Germanic Languages

RECOMMENDED READING for Chapter 5: Cook & Newson (1996: 215-222)

In this session we will discuss an article of mine about verb-second phenomena, and how they can be elegantly described in GB (Schneider (1995)). I will hand out the article separately. The part where I have implemented the theory in a computational framework under Lexical-Functional Grammar is not relevant for our discussion, and you may skip it.

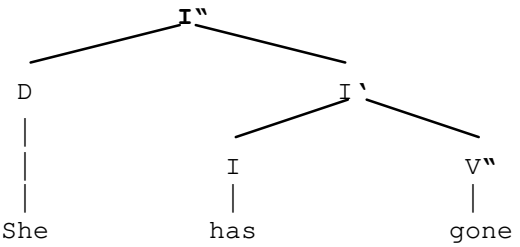
6 Aspects of Minimalism

Due to the most unnatural time constraints during the composition of these notes for my colloquium, I have not been able to prepare this session in detail yet. I will hand out an extended and more easily understandable version of this chapter in the course of the term.

6.1 Checking Theory

RECOMMENDED READING for Subchapter 6.1: Radford (1997: 170-198)

Checking Theory is an extension and formalisation of the rule that a head can subcategorise for features and categories of its arguments and, to a lesser degree, its specifiers. It is an assertion of the lexicalist character of GB, since features subcategorised for by lexical heads. Checking Theory makes sure that the necessary features agree. It postulates that every lexical head has head features (its own features) and requires other features to be met by its specifier and its complement.



PERS:3 GEND:fem NUM:sg CASE:nom	TENSE:pres	+N	*head-features* (self-features)
	PERS:3 GEND:fem NUM:sg CASE:nom		*specifier-features* (required on spec.)
	+N		*complement-features* (required on compl.)

The feature [+N] is a cross-categorial generalisation, comprising nouns, adjectives and participles.

Checking Theory is also a concession to other formal grammar theories, especially unification-based grammars as computational linguists use them. There, features and feature checking have always been fundamental.

6.2 Bare Phrase Structure

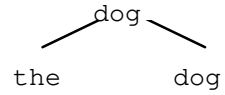
RECOMMENDED READING For Subchapter 6.2: Cook & Newson (1996: 338-344)

Minimalism gives up X-bar theory and suggests Bare Phrase Structure instead. The differences are not fundamental, however. Bare Phrase Structure is more process-oriented than X-bar grammar. An operation called Merge takes two adjacent words in the sentence and decides which should be the head.

The minimalist programme is focused on the process of building up syntactical structures.

> In GB, D-structure was presented as a complete structure and not much was said about the internal process of how it was formed ... However, in getting rid of D- and S-structure, minimalism places more emphasis on the internal workings of the structure formation process ... part of this process involves building individual trees for lexical items (...) and then combining these at some point to form a larger tree. How and when they [=individual trees] are combined will have repercussions on the eventual SD [=syntactic description] formed
>
> (Cook & Newson 1996: 323)

The operation that combines lexical items or partial trees with others is called *Merge* by Chomsky, "an operation that forms larger units out of those already constructed" (Chomsky 1995: 396). Aiming at minimalism, *Merge* is always a binary relation, i.e. it combines two partial trees to form something new. The new tree is labelled with the lexical entry of the head.



Bare Phrase Structure makes no distinction between different bar-levels. The mother node *dog* and the daughter node *dog* have the same status. They are both treated as lexical heads. Bar-status (i.e. whether a constituent is X0, X' or X") is no longer a fundamental principle, but a derivable property:

> First, the X' schema stipulates that XP must dominate X' and X' must dominate X. Second, it stipulates that the complement is adjoined to the head and that the specifier is adjoined to the result. Suppose that we simply abandon the notion that phrasal categories are distinguished by their level; on such a view heads and maximal projections are all of the same category, and the are distinguished by where they appear in the structure. An X that does not branch is a head; an X that does not have an X above it and that has a path of nodes only of category X down to a head of category X is a maximal projection.
>
> (Culicover 1997: 355-6)

7 What We Have not Covered

This last session will be a rag-bag overview of what we have not covered in the colloquium, perhaps we will also have more questions about the previous sessions.

There are some important parts of the theory I have chosen not to include are. This list should provide a short overview, as well as an explanation (or apology) why I have not included these topics in the colloquium.

- **Theta-theory**: Theta-Theory assigns thematic roles like *agent, experiencer* or *goal* to the participants in a predication. I have not discussed it here because it is partly more semantically than syntactically oriented, because the status and number of thematical roles is still fiercely debated in the theory of thematic cases itself. Cowper (1992: 64) admits that „much work in this area has to be done".
- **Government**: Government is a structural relation, which gives rise to many technical argumentations, but which could never be satisfactorily defined:

  > A ... GB problem is the concept of government, which, despite having considerable empirical motivation, is none the less an 'arbitrary syntactic relation' (Lasnik, 1993, p.3). Moreover, there are many different notions of government, some working better than others for certain phenomena; unfortunately, no version is perfect for all purposes.
  >
  > (Cook & Newson 1996: 315)

  Minimalism completely gives up government.
- **Binding Theory**: Binding Theory describes the references of anaphora and pronouns, certainly an interesting topic, but too specific for a short introductory course.
- **Barriers**: Barriers restricts long-distance dependencies. Even in introductory books on GB, it is usually presented at the end. Quite a thorough understanding of GB is necessary to profit from its discussion.

8 Bibliography

CHOMSKY, NOAM, 1957. *Syntactic Structures.* The Hague: Mouton.

CHOMSKY, NOAM, 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

CHOMSKY, NOAM, 1986. *Knowledge of Language. Its Nature, Origin and Use.* New York: Praeger.

CHOMSKY, NOAM, 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

COOK, VIVIAN and MARK NEWSON, 1996. *Chomsky's Universal Grammar.* Oxford: Blackwell.

COWPER, ELISABETH, 1992. *A Concise Introduction to Syntactic Theory.* Chicago: The University of Chicago Press

CULICOVER, PETER, 1997. *Principles and Parameters: An Introduction to Syntactic Theory.* Oxford: OUP.

FINEGAN, EDWARD, 1994. Language. Its Structure and Use. 2nd ed. Fort Worth, TX: Harcourt Brace

HAEGEMAN, LILIANE, 1994. *Introduction to Government & Binding Theory*, 2nd ed. Oxford: Blackwell.

HORNSTEIN, NORBERT, 1995. *Logical Form. From GB to Minimalism*. Oxford: Blackwell.

POLLARD, CARL and IVAN A. SAG, 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

RADFORD, ANDREW, 1988. *Transformational Grammar: A First Course.* Camdridge: CUP.

RADFORD, ANDREW, 1997. *Syntactic Theory and the Structure of English. A Minimalist Approach.* Cambridge: CUP.

SCHNEIDER, GEROLD, 1995. „Using GB Principles to DescribeDanish, English and German in LFG". In: Proceedings of the First Swiss-Estonian Strudent Workshop, September 1995. Tartu: University of Tartu, Department of General Linguistics and Department of Computer Science.

YNGVE, VICTOR, 1996. *From Grammar to Science: New Foundations for General Linguistics*. Amsterdam: Benjamin.