Constraints in Syntax

Gereon Müller

Universität Leipzig

Lecture Notes

Version 11th April 2006

Contents

Chapter 1

Phrase Structure and Derivations

1.1 Background

Task:
Determine whether a given string of words is a grammatical (well-formed) sentence of language L_i or not.

Approach:
Throughout, a derivational, incremental approach to grammar will be adopted (following Chomsky (1995; 2000; 2001a)).

(1) *The syntactic component of a grammar*:
  a. Lexicon: Set of lexical items (LIs)
  b. Operations: Merge, Move
  c. Constraints: derivational, representational, global, transderivational, translocal

Categories:
LIs can be grouped together if they have a similar syntactic distribution and similar morphological properties: word classes, categories. There are *lexical* (typically open-class) and *functional* (typically closed-class) categories (both are lexical items).

(2) *Lexical categories*:
  a. N (noun): *Mary*, *man*, *book*, *idea*, ...
  b. V (verb): *snore*, *meet*, *kiss*, *give*, *believe*, ...
  c. A (adjective): *green*, *nice*, *fond*, ...
  d. P (preposition): *of*, *under*, *for*, *after*, ...

(3) *Functional categories*:

a. D (determiner): *the*, *a*, *every*, *some*, *three*, *she*, *it*, Ø ...
b. T (tense): *did*, *can*, *will*, Ø, ...
c. C (complementizer): *that*, *whether*, *if*, *for*, Ø

(4) *The make-up of LIs*:
a. Phonological representation
b. Semantic denotation: This includes the *argument structure* or Θ-*grid* of a predicate (see Heim & Kratzer (1998))
c. Morpho-syntactic features (incl. categorial features, Case features, ϕ-features (number, person, gender), tense features, and selectional features)

Basic assumption:
Syntactic structures arise by iterative application of simple operations to LIs and categories formed by these operations – a building block system that makes syntactic structures grow. This procedure is called a *derivation*. Every intermediate structure thus formed is a *representation*. A sentence is a final representation that results when the derivation terminates. Chomsky (1995; 2000; 2001a) assumes that before a derivation starts, all the LIs that will be used in the derivation are assembled in a *lexical array* (LA) (sometimes, the notion of *numeration* is used – a numeration is a lexical array in which one and the same LI can be selected more than once).

(5) *Lexical array* (LA):
The lexical array is a set of LIs that are selected before the syntactic derivation starts.

(6) *Inclusiveness Condition*:
Material that is not part of the lexical array (other LIs, additional features) is inaccessible throughout a derivation.

1.2 Merge

The first basic operation that drives derivations is the operation *Merge*, which applies to two categories and yields a complex, structured object (a *phrase marker* or *tree* that is necessarily binary branching). If Merge uses a LI from the LA, the LI has irrevocably left that array.

(7) *Merge*:
Merge(α,β) yields $[_\gamma\ \alpha\ \beta]$.

(8) *Head*:

a. If γ has been created by Merge of α and β, α is the *head* of γ (and γ is a *projection* of α) if α selects β.
b. A head and its projection share morpho-syntactic features.

Note:
A LI may require one or more other categories with certain properties in its projection in the syntax. Requirements of this type can be encoded in *selectional features* on a LI: [*F*], where [F] is the feature that the LI is looking for. [F] is typically a categorial feature (in which case the selection requirement is often called a *subcategorization* requirement); but it may also be some other morpho-syntactic feature. Selectional features have a peculiar status: They do not seem to be interpretable semantically (at the level of Logical Form).

(9) *Interpretability Condition*:
Features on LIs that are uninterpretable at level R_i must be removed at level R_{i-1}.

(10) *Levels of grammar*:
Lexicon (plus Morphology) → Syntax → Phonological Form, Logical Form

Assumption:
Selectional features must be deleted in the course of a syntactic derivation. Merge deletes a selectional feature under identity with the corresponding feature on a selected item; if uninterpretable, the corresponding feature is also deleted.

(11) *Economy Constraint on Merge*:
Merge can only apply if it deletes a selectional feature of a LI.

(12) *The Scope of Feature Deletion*:
Deletion of selectional features implies deletion of matching uninterpretable features.

(13) *Crucial concepts of phrase structure*:
a. β is a *complement* of α iff β has been merged with a LI α that selects it.
b. β is a *specifier* of α iff β has been merged with a non-LI α that selects it.
c. α and β are *sisters* iff they have been merged.
d. γ *immediately dominates* α iff Merge has applied to α yield γ. (In that case, α is a *daughter* of γ.)

e. γ *dominates* α iff (a) or (b) holds:
(i) γ immediately dominates α.
(ii) γ immediately dominates δ, and δ dominates α.
(α is a *term* or *constituent* of γ iff γ dominates α.)

f. α *c-commands* β iff (a) or (b) holds:
(i) β is a sister of α.
(ii) β is is dominated by a sister of α.

g. γ is an XP category iff it is the maximal projection of α.

h. γ is an X′ category iff it is a non-maximal, non-minimal projection of α.
(Note: The status of a projection as an XP, X′, or X (= LI) may change during the derivation.)

Note:
As such, the operation Merge does not say anything about linearization. Linear precedence is handled separately, by linear precedence statements (see Gazdar, Klein, Pullum & Sag (1985)).

(14) *Linear precedence statements* (English):
a. A head precedes its complement: X CompX.
b. A head follows its specifier(s): SpecX X.

(15) *Linearization Constraint on Merge*:
The output of Merge in language L_i must conform to the linear precedence statements of L_i.

1.3 Verb Phrases

(16) *Verb types*:
a. intransitive verbs: *sleep*, *snore*, *arrive*
argument structure: $[\Theta_1]$
b. transitive verbs: *kiss*, *like*, *believe*
argument structure: $[\Theta_1 > \Theta_2]$
c. ditransitive verbs: *give*, *send*, *show*
argument structure: $[\Theta_1 > \Theta_2 > \Theta_3]$

Note:
Θ-roles are sometimes given names, such as “Agent”, “Patient”, “Goal”, “Experiencer”.

Terminology:
Linguistic expressions that realize Θ-roles of the argument structures of verbs are called *arguments*. The argument that realizes the highest Θ-role of a (non-trivial) argument structure is sometimes called the *external* argument; other arguments can be referred to as *internal* arguments. Linguistic expressions that have argument structures (i.e., that need to combine with arguments in the syntax) are called *predicates*.

Intended state of affairs:
The highest Θ-role ends up on the highest argument in VP, the lowest Θ-role on the lowest VP-internal argument.

Question:
How can this be ensured by the Merge operation?

(17) *Linking Principle*:
The semantically-based argument structure corresponds to a reverse hierarchy of syntactically accessible selectional features.

Assumption:
The *Economy Constraint on Merge* is revised appropriately:

(18) *Economy Constraint on Merge* (revised):
Merge can only apply if it deletes the highest-ranked selectional feature of a LI.

(19) *Selection requirements and argument structures of verbs*:

a. arrive: [*D*] $\leftarrow [\Theta_1]$
b. like: [*D*] > [*D*] $\leftarrow [\Theta_1 > \Theta_2]$
c. believe: [*C*] > [*D*] $\leftarrow [\Theta_1 > \Theta_2]$

Note:
believe (and even *kiss*) can also impose other selection requirements based on the same argument structure.

Assumption:
Words like the verbs in (19) enter LAs after having passed the morphological component, as fully inflected word forms.

(20) is a derivation of a sentence with a transitive verb. Categorial features are represented as indices on constituents in the derivation (*labelled bracketing*). The LIs of the LA are simplified.

(Gender agreement is not morphologically visible in English, which has an impoverished inflectional system. Gender marking on a verb may become visible in languages like Russian – *On čital/ona čitala knigu*, 'He/she read a book'.)

(20) *A LA of a VP with an intransitive verb*:
  a. arrived: { [V], [3pers,–pl,+masc], [+fin,+past], [*D*] }
  b. he: { [D], [3pers,–pl,+masc], [nom] }

(21) *A derivation of a VP with an intransitive verb*:
  Merge ($[_{\text{D}}$ he] , $[_{\text{V}}$ arrived]) → $[_{\text{VP}}$ $[_{\text{V}}$ arrived] $[_{\text{DP}}$ he]]

Note 1:
Merge is possible because it deletes the selectional, uninterpretable [*D*] feature of V; it does not delete the categorial [D] feature of D, which is interpretable.

Note 2:
The DP is a complement of V (CompV), not a specifier (SpecV). Hence, given the linear precedence statements that hold in English, the DP must show up to the right of V in VP. Since this is not the surface word order, it can be concluded that further derivational steps are necessary in (21). As it stands, the same prediction is made for verbs like *snore*. [It has been argued that there are two classes of intransitive verbs – *unaccusative* intransitive verbs, and *unergative* intransitive verbs. The former are then assumed to be merged in the complement position of V, as in (21), whereas the latter are merged in the position otherwise reserved for the external argument in transitive constructions. Chomsky (1995, 315) suggests that "unergative verbs are hidden transitives", with an unexpressed DP in the complement position.]

(22) *A LA of a VP with a transitive verb*:
  a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
  b. she: { [D], [3pers,–pl,+fem], [nom] }
  c. him: { [D], [3pers,–pl,+masc], [acc] }

(23) *A derivation of a VP with a transitive verb*:
  a. Merge ($[_{\text{D}}$ him], $[_{\text{V}}$ likes]) → $[_{\text{VP}}$ $[_{\text{V}}$ likes] $[_{\text{DP}}$ him]]
  b. Merge ($[_{\text{D}}$ she], $[_{\text{VP}}$ $[_{\text{V}}$ likes] $[_{\text{DP}}$ him]]) → $[_{\text{VP}}$ $[_{\text{DP}}$ she] $[_{\text{V}'}$ $[_{\text{V}}$ likes] $[_{\text{DP}}$ him]]]

Problem:
Is there anything that would preclude the reverse application of Merge operations?

(24) *An unwanted derivation for a VP with a transitive verb*:

a. Merge ([$_{D}$ she], [$_{V}$ likes]) → [$_{VP}$ [$_{V}$ likes] [$_{DP}$ she]]
b. Merge ([$_{D}$ him], [$_{VP}$ [$_{V}$ likes] [$_{DP}$ she]]) → [$_{VP}$ [$_{DP}$ him] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ she]]]
(Intended meaning: 'He likes her')

Solution:
Transitive verbs have Case features that play a role similar to selectional features. There are two types of object Cases – *structural* Case and *inherent* Case. ("Case" is written with a capital "C" in order to indicate that the case is *abstract*, not necessarily morphologically visible.) Inherent Case is tied to a certain selectional feature (hence, to a certain Θ-role in the argument structure); structural Case is not. The distinction is particularly relevant in languages like German. The accusative Case of a regular transitive verb is structural. It does not have to go hand in hand with a specific selectional feature (although it usually does); see (25-a) (where *ihn* satisfies a selectional [*D*] feature of V) vs. (25-b) (where *ihn* does not).

(25) *Structural accusative Case in German*:

a. dass [$_{VP}$ sie ihn mag]
that she$_{nom}$ him$_{acc}$ likes
b. dass [$_{VP}$ sie [ihn gehen] laesst
that she$_{nom}$ him$_{acc}$ go lets

On the other hand, the dative Case of a simple transitive verb in German is not regular; it is tied to a certain selectional feature:

(26) *Inherent dative Case in German*:

a. dass [$_{VP}$ sie ihm hilft]
that she$_{nom}$ him$_{dat}$ helps
b. *dass [$_{VP}$ sie [ihm gehen] laesst]
that she$_{nom}$ him$_{acc}$ go lets

Another difference between inherent and structural Cases emerges in *passive* constructions in German: Inherent Case is maintained, structural Case disappears.

(27) *Structural vs. inherent Case in passive constructions*:

a. dass er/*ihn gemocht wird
   that he_{nom}/him_{acc} liked Aux
b. dass *er/ihm geholfen wird
   that he_{nom}/him_{dat} helped Aux

(28) *Features of transitive verbs in German*:
a. mag: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
b. hilft: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*,*dat*] > [*D*] }

Note:
Like other selectional features, [*Case*] features are uninterpretable; so are [Case] features (in contrast to categorial features). Given the Interpretability Condition, they must be removed from the derivation. Assumption: The removal must proceed as quickly as possible. As soon as V can delete its [*acc*] feature, and DP's [acc] feature, by Merge, it must do so. This rules out the derivation in (24).

(29) *The Timing of Feature Deletion*:
Uninterpretable features are deleted as soon as possible.

1.4 Determiner Phrases and Noun Phrases

(30) *A LA of a simple DP in English*:
a. a: { [D], [3pers,–pl,+fem], [nom], [*N*] }
b. woman: { [N], [3pers,–pl,+fem], [nom] }

(31) *A derivation of a simple DP in English*:
Merge ($[_{D}$ a], $[_{N}$ woman]) → $[_{DP}$ $[_{D}$ a] $[_{NP}$ woman]]

Note:
Since D and N both have (uninterpretable) [Case] features, not (uninterpretable) selectional [*Case*] features, there is no deletion of [Case] features involved in the Merge operation; only the [*N*] feature of D is deleted.

(32) *A LA of a more complex DP in German*:
a. manchen: { [D], [3pers,+pl,–masc,–fem], [+dat], [*N*] } (some)
b. des: { [D], [3pers,–pl,+masc], [+gen], [*N*] } (the)
c. Interpretationen: { [N], [3pers,+pl,–masc,–fem], [+dat], [*D*] } (interpretations)
d. Vorschlags: { [N], [3pers,–pl,+masc], [+gen], } (proposal)

(33) *A derivation of a more complex DP in German*:

a. Merge ([$_{D}$ des] [$_{N}$ Vorschlags]) → [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]
b. Merge ([$_{N}$ Interpretationen], [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]) →
[$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]]
c. Merge ([$_{D}$ manchen], [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]]) → [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]]]

(34) *Phrase markers that cannot be generated*:
a. *[$_{DP}$ [$_{D}$ manchen] [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{NP}$ Vorschlags] [$_{D}$ des]]]]
(*Linearization Constraint on Merge)
b. *[$_{DP}$ [$_{D}$ manchen] [$_{NP}$ [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]] [$_{N}$ Interpretationen]]
(*Linearization Constraint on Merge)
c. *[$_{DP}$ [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ des] [$_{NP}$ Vorschlags]]] [$_{D}$ manchen]]
(*Linearization Constraint on Merge)
d. *[$_{DP}$ [$_{NP}$ [$_{DP}$ [$_{NP}$ Vorschlags] [$_{D}$ des]] [$_{N}$ Interpretationen]] [$_{D}$ manchen]]
(*Linearization Constraint on Merge)
e. *[$_{DP}$ [$_{D}$ manchen] [$_{DP}$ [$_{D}$ des] [$_{NP}$ [$_{N}$ Interpretationen] [$_{NP}$ Vorschlags]]]]
(*Economy Constraint on Merge)
f. *[$_{NP}$ [$_{N}$ Vorschlags] [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ manchen] [$_{DP}$ des]]]]
(*Economy Constraint on Merge)

Problem:
Is there anything that would preclude the application of Merge operations in (35)?

(35) *An unwanted derivation for a more complex DP in German*:
a. Merge ([$_{D}$ manchen], [$_{N}$ Vorschlags]) → [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ Vorschlags]]
b. Merge ([$_{N}$ Interpretationen], [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ Vorschlags]]) →
[$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ Vorschlags]]]
c. Merge ([$_{D}$ des], [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ Vorschlags]]]) → [$_{DP}$ [$_{D}$ des] [$_{NP}$ [$_{N}$ Interpretationen] [$_{DP}$ [$_{D}$ manchen] [$_{NP}$ Vorschlags]]]]

Solution:
This derivation violates an agreement constraint on Merge that holds in German:

(36) *Agreement Constraint on Merge with D*:
D can only be merged with N if the two items have identical Case and ϕ-features.

Note:
So far, the assumption is that personal pronouns like *he* are D elements, hence, DPs when they occur as arguments of predicates. Expressions like *a woman* are also DPs. In line with this, selectional features of verbs so far only included [*D*], not [*N*].

Problem:
What about proper names (like *John*, *Madonna*) or indefinite common nouns in the plural (so-called "bare plurals", like *books*)?

(37) *Proper names and bare plural common noun phrases*:
- a. She likes John
- b. She likes books about Madonna

Assumption:
[*N*] cannot be a selectional feature on a predicate; but [*D*] can be. The NPs in (37) are selected by empty D elements. If there is no empty D selecting, e.g., *John* in the initial LA, (37-a) cannot be generated by a derivation.

(38) *LA of (37-a)*:
- a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
- b. she: { [D], [3pers,–pl,+fem], [nom] }
- c. John: { [N], [3pers,–pl,+masc], [acc] }
- d. Ø: { [D], [3pers,–pl,+masc], [acc], [*N*] }

(39) *Derivation of (37-a)*:
- a. Merge ([$_{\text{D}}$ Ø], [$_{\text{N}}$ John]) → [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]
- b. Merge ([$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]], [$_{\text{V}}$ likes]) → [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]
- c. Merge ([$_{\text{D}}$ she], [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]) → [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V'}}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]

(40) *LA of (37-b)*:

a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. Ø: { [D], [3pers,–pl,+fem], [acc], [*N*] }
d. Madonna: { [N], [3pers,–pl,+fem], [acc] }
e. books: { [N], [3pers,+pl,–fem,–masc], [acc], [*P:about*] }
f. about: { [P], [*D*], [*acc*] }

Note:
The selectional feature [*P:about*] on the noun *books* is optional (as is the feature [*D*] on the noun *Interpretationen* in the German example above). This is due to a general difference between V and N: Whereas arguments of V are typically obligatory, arguments of N are optional.

(41) a. (i) $[_{VP}$ Mary describes John]
(ii) *$[_{VP}$ Mary describes]
(iii) *$[_{VP}$ Describes John]
b. (i) $[_{DP}$ Mary $[_{D}$'s] descriptions of John]
(ii) $[_{DP}$ Ø descriptions of John]
(iii) $[_{DP}$ Mary $[_{D}$'s] descriptions]

(42) *Derivation of (37-b)*:
a. Merge ($[_{D}$ Ø], $[_{N}$ Madonna]) → $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]
b. Merge ($[_{P}$ about], $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]) →
$[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]
c. Merge ($[_{N}$ books], $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]) →
$[_{NP}$ $[_{N}$ books] $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]]
d. Merge ($[_{V}$ likes], $[_{NP}$ $[_{N}$ books] $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]]) →
$[_{VP}$ $[_{V}$ likes] $[_{NP}$ $[_{N}$ books] $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]]]
e. Merge ($[_{D}$ she],
$[_{VP}$ $[_{V}$ likes] $[_{NP}$ $[_{N}$ books] $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]]]) →
$[_{VP}$ $[_{DP}$ she] $[_{V'}$ $[_{V}$ likes] $[_{NP}$ $[_{N}$ books] $[_{PP}$ $[_{P}$ about] $[_{DP}$ $[_{D}$ Ø] $[_{NP}$ Madonna]]]]]]

1.5 Tense Phrases

1.5.1 T

Problem:
Sentences like those in (43) cannot yet be generated. Auxiliaries and modal verbs do not form a LI together with the main verb (adverbs can intervene).

(43) a. John will (probably) arrive
b. She has (probably) taken the book
c. Mary can (probably) read it

Assumption:
We can assume that auxiliaries and modals have a selectional feature [*V*]; so they merge with VPs. The categorial feature of *will* and *can* can be assumed to be [T] (for "Tense"), not [V].

(44) *LA of (43-a)*:
a. arrive: { [V], [–fin,–part], [*D*] }
b. John: { [N], [3pers,–pl,+masc], [nom] }
c. Ø: { [D], [3pers,–pl,+masc], [nom], [*N*] }
d. will: { [T], [3pers,–pl,+masc], [+fin,+fut], [*V*, *–fin*,*–part*] }

(45) *An unsuccessful derivation of (43-a)*:
a. Merge ([$_{D}$ Ø], [$_{N}$ John]) → [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]
b. Merge ([$_{V}$ arrive], [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]] → [$_{VP}$ [$_{V}$ arrive] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]
c. Merge ([$_{T}$ will], [$_{VP}$ [$_{V}$ arrive] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]) →
[$_{TP}$ [$_{T}$ will] [$_{VP}$ [$_{V}$ arrive] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]]

(46) *LA of (43-b)*:
a. taken: { [V], [–fin,+part], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. the: { [D], [3pers,–pl,–fem,–masc], [acc], [*N*] }
d. book: { [N], [3pers,–pl,–fem,–masc], [acc] }
e. has: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*–fin*,*+part*] }

(47) *An unsuccessful derivation of (43-b)*:
a. Merge ([$_{D}$ the], [$_{N}$ book])→ [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]
b. Merge ([$_{V}$ taken], [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]) →
[$_{VP}$ [$_{V}$ taken] [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]]

c. Merge ([D she], [VP [V taken] [DP [D the] [NP book]]]) →
[VP [DP she] [V′ [V taken] [DP [D the] [NP book]]]]
d. Merge ([T has], [VP [DP she] [V′ [V taken] [DP [D the] [NP book]]]]) →
[TP [T has] [VP [DP she] [V′ [V taken] [DP [D the] [NP book]]]]]

(48) *LA of (43-c)*:
a. read: { [V], [–fin,–part], [*D*] > [*D*], [*acc*] }
b. Mary: { [N], [3pers,–pl,+fem], [nom] }
c. Ø: { [D], [3pers,–pl,+fem], [nom], [*N*] }
d. it: { [D], [3pers,–pl,–fem,–masc], [acc] }
e. can: { [T], [3pers,–pl,+masc], [+fin,–past], [*V*, *–fin*,*–part*] }

(49) *An unsuccessful derivation of (43-c)*:
a. Merge ([D Ø], [N Mary]) → [DP [D Ø] [NP Mary]]
b. Merge ([V read], [D it]) → [VP [V read] [DP it]]
c. Merge ([DP [D Ø] [NP Mary]], [VP [V read] [DP it]]) →
[VP [DP [D Ø] [NP Mary]] [V′ [V read] [DP it]]]
d. Merge ([T can], [VP [DP [D Ø] [NP Mary]] [V′ [V read] [DP it]]]) →
[TP [T can] [VP [DP [D Ø] [NP Mary]] [VP [V read] [DP it]]]]

Note:
The output of the first Merge operation in the derivation in (49) is not immediately an input for the second operation. In fact, these two operations can apply in either order. Trees that have been created and that then wait for further use in a derivation can be said to belong to the *workspace* of a derivation.

1.5.2 Move

Solution:
There is displacement of the external argument to the specifier of T (SpecT). This position is the *subject* position.
(Strictly speaking, this is the only sense in which one can talk about subjects in this approach: Subjects are XPs in SpecT, period. Informally, however, one often uses the notion of "subject" when referring to external arguments, or to DPs bearing nominative Case.)
Technically, displacement is brought about by the second basic operation that drives derivations, viz., the operation *Move*. Move is a composite op-

eration based on Merge. Consequently, all constraints on Merge (Economy Constraint, Linearization Constraint) also hold for Move.

(50) *Move*:
Move(α,β) = Merge(α,β), with the following additional requirements:
   a. The selectional [*F*] feature of the head α can only be deleted under identity with a corresponding feature on the non-head β in specifier position.
   b. The non-head β is taken from the same phrase marker (not from the LA or from the current workspace).
   c. A trace is left in the original position.

Note on (50-a):
[*F*] features triggering Move are sometimes referred to as "EPP" features (because of a constraint called "Extended Projection Principle" from Chomsky (1982), which inter alia required the filling of SpecT). As an alternative to explicitly restricting [*F*]-deletion to identity with a corresponding feature in a specifier position, one can postulate that the [*F*]-features in question are always lowest-ranked in hierarchies of selectional features: ... > [*F*]. In what follows, this latter assumption will be adopted throughout.
(The only potential difference might arise in modification contexts; see below.)

Note on (50-b):
Pure Merge is called *external Merge*, whereas Move is called *internal Merge* in recent work by Chomsky (2001b).

Note on (50-c):
There are two possibilities as to what a trace is.
(i) A trace is a empty category "t" The relation between a moved item and its trace is indicated by co-indexing.
(ii) A trace is an exact *copy* of the moved item that is unpronounced at the level of Phonological Form (the *copy theory* of movement).
For present purposes, (i) can be assumed (which may require a minimal modification of the Inclusiveness Condition).

(51) *Extended feature sets of T elements*:
   a. will: { [T], [3pers,–pl,+masc], [+fin,+fut], [*V*, *–fin*,*–part*] > [*D*] }
   b. has: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*–fin*,*+part*] > [*D*] }

c. can: { [T], [3pers,–pl,+masc], [+fin,–past], [*V*, *–fin*,*–part*] > [*D*] }

Consequence:
The derivations in (45), (47), and (49) are incomplete.

(52) *A successful derivation of (43-a)*:

...

c. Merge ([$_{\text{T}}$ will], [$_{\text{VP}}$ [$_{\text{V}}$ arrive] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]) →
[$_{\text{TP}}$ [$_{\text{T}}$ will] [$_{\text{VP}}$ [$_{\text{V}}$ arrive] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]

d. Move ([$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]], [$_{\text{TP}}$ [$_{\text{T}}$ will] [$_{\text{VP}}$ [$_{\text{V}}$ arrive] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]] [$_{\text{T}'}$ [$_{\text{T}}$ will] [$_{\text{VP}}$ [$_{\text{V}}$ arrive] t_1]]]

(53) *A successful derivation of (43-b)*:

...

d. Merge ([$_{\text{T}}$ has], [$_{\text{VP}}$ [$_{\text{D}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]) →
[$_{\text{TP}}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]]

e. Move ([$_{\text{DP}}$ she],
[$_{\text{TP}}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]]
)
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ t_1 [$_{\text{V}'}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]]]

(54) *A successful derivation of (43-c)*:

...

d. Merge ([$_{\text{T}}$ can], [$_{\text{VP}}$ [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ Mary]] [$_{\text{V}'}$ [$_{\text{V}}$ read] [$_{\text{DP}}$ it]]]) →
[$_{\text{TP}}$ [$_{\text{T}}$ can] [$_{\text{VP}}$ [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ Mary]] [$_{\text{V}'}$ [$_{\text{V}}$ read] [$_{\text{DP}}$ it]]]]

e. Move ([$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ Mary]], [$_{\text{TP}}$ [$_{\text{T}}$ can] [$_{\text{VP}}$ [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ Mary]] [$_{\text{V}'}$ [$_{\text{V}}$ read] [$_{\text{DP}}$ it]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ Mary]] [$_{\text{T}'}$ [$_{\text{T}}$ can] [$_{\text{VP}}$ t_1 [$_{\text{V}'}$ [$_{\text{V}}$ read] [$_{\text{DP}}$ it]]]]]

Problem:
Why can only the DP *she* move to SpecT, and not the DP *the book*, in a derivation based on the LA in (46) (repeated here as (55), with the revised entry for *has*)? Compare (43-b) with (56).

(55) *LA of (43-b)*:

a. taken: { [V], [–fin,+part], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. the: { [D], [3pers,–pl,–fem,–masc], [acc], [*N*] }
d. book: { [N], [3pers,–pl,–fem,–masc], [acc] }
e. has: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*–fin*,*+part*] > [*D*] }

(56) *An unwanted derivation of (43-b)*:

...

d. Merge ([$_{\text{T}}$ has], [$_{\text{VP}}$ [$_{\text{D}}$ she] [$_{\text{V'}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]) →
[$_{\text{TP}}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V'}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]]

e. Move ([DP [$_{\text{D}}$ the] [$_{\text{NP}}$ book]],
[$_{\text{TP}}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V'}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]] [$_{\text{T'}}$ [$_{\text{T}}$ has] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V'}}$ [$_{\text{V}}$ taken] t_2]]]]

Solution:
There is an agreement constraint on Merge operations applying to T that is similar to the constraint on Merge operations applying to D given above.

(57) *Agreement Constraint on Merge with T*:
T can only be merged with D if the two items have identical ϕ-features.

Consequence:
T and DP in SpecT must have identical ϕ-features. Given the LA in (55), this is possible if *she* moves to SpecT, but not if *the book* moves to SpecT.

A new problem:
But what if *the* and *book* had ϕ-features matching those of T in the LA to begin with? Could the accusative DP then move to the subject position SpecT?

Solution 1:
The [*D*] feature on T is replaced with a [*nom*] feature on T, as in (58-b) (vs. (58-a), assumed so far). On this view, T selects a nominative DP rather than a DP. Under present assumptions, this is in fact the only way to remove the uninterpretable [nom] feature on a DP; if T does not bear [*nom*], additional assumptions have to be made to ensure [nom] deletion on DP.

(58) *Two theories about the features of T*:

a. has: { [T], ... [*V*,*–fin*,*+part*] > [*D*] }
b. has: { [T], ... [*V*,*–fin*,*+part*] > [*nom*] }

Solution 2:
There is a constraint on Move that permits only movement of the higher (c-commanding) item in cases of ambiguity (superiority, minimality).

Conclusion:
In what follows, solution 1 will be adopted for the sake of simplicity.

1.5.3 Empty T

The situation so far:
(i) Movement to SpecT is required for clauses involving auxiliaries and modals to predict the correct word order.
(ii) Movement to SpecT does not seem to be required for clauses that contain only a transitive main verb to predict the correct word order; see derivation (23), whose output is repeated here as (59-a).
(iii) Movement to SpecT does seem to be required for clauses that contain only an intransitive main verb to predict the correct word order; see derivation (21), whose output is repeated here as (59-b).

(59) *Word order in the VP*:

a. [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]]
b. [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ he]]

Assumption:
All clauses have a T; but T can be empty.

(60) *An extension of the LA of a VP with a transitive verb in (22)*:

a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. him: { [D], [3pers,–pl,+masc], [acc] }
d. Ø: { [T], [3pers,–pl,+fem], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }

(61) *An extension of the derivation of a VP with a transitive verb in (23)*:

a. Merge ([$_{D}$ him], [$_{V}$ likes]) → [$_{VP}$ [$_{V}$ likes] [$_{DP}$ him]]
b. Merge ([$_{D}$ she], [$_{VP}$ [$_{V}$ likes] [$_{DP}$ him]]) → [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]]

c. Merge ([$_{T}$ Ø], [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]])
→ [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]]]
d. Move ([$_{DP}$ she], [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]]])
→ [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ him]]]]]

(62) *An extension of the LA of a VP with an intransitive verb in (20)*:
a. arrived: { [V], [3pers,–pl,+masc], [+fin,+past], [*D*] }
b. he: { [D], [3pers,–pl,+masc], [nom] }
c. Ø: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*+fin*,*+past*] > [*nom*] }

(63) *An extension of the derivation of a VP with a transitive verb in (21)*:
a. Merge ([$_{D}$ he] , [$_{V}$ arrived]) → [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ he]]
b. Merge ([$_{T}$ Ø], [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ he]]) → [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ he]]]
c. Move ([$_{DP}$ he], [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ he]]])
→ [$_{TP}$ [$_{DP_1}$ he] [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{V}$ arrived] [$_{DP}$ t_1]]]

Question:
How does it follow that the subject DP agrees with finite V in English with respect to ϕ-features?

Answer:
(i) T must always have the same ϕ-features as a DP in SpecT, because of the Agreement Constraint on Merge with T.
(ii) This is sufficient for auxiliary and modal contexts, where the finite verb is in T. For contexts in which the main verb is finite, it follows from the following assumption:

(64) *Selection by empty T*:
An empty T selects a V with ϕ-features that correspond to its own ϕ-features.

1.5.4 Empirical Evidence

Observation:
There is independent evidence from constituency tests that the external argument DP must leave the VP. (However, none of these tests is absolutely reliable.)

Displacement test:

If a sequence of words may undergo displacement as a group, they must form an XP.

(65) *Displacement*:
Mary wanted to read the book, and [$_{VP_1}$ read the book] she did t_1

Deletion test:
If a sequence of words may undergo deletion as a group, they must form an XP.

(66) *Deletion*:
Mary never wanted to read the book, but she did [$_{VP}$ –]

Pro-form test:
If a sequence of words may be replaced with a pro-form, it must form an XP.

(67) *Pro-form*:
Mary can read the book, and [$_{VP}$ so] can John

Pseudo-Cleft test:
If a sequence of words may show up to the right of the inflected form of *be* in pseudo-cleft sentences, they must form an XP.

(68) *Pseudo-Cleft*:
What Mary did was [$_{VP}$ read the book]

Note:
From this, one might conclude that external argument DPs are in fact never part of the VP. But:

Observation:
There is independent evidence that an external argument DP is merged within VP.

Quantifier floating in French (Sportiche (1988)):
In quantifier floating constructions, only a part of an external argument DP moves to SpecT, stranding DP-material in SpecV.

(69) *Quantifier floating in French*:

a. [$_{DP_1}$ Tous [$_{DP_2}$ les garçons]] ont [$_{VP}$ t_1 lu ce livre]
all the boys have read the book

b. [$_{DP_2}$ Les garçons]] ont [$_{VP}$ [$_{DP_1}$ tous t_2] lu ce livre]
the boys have all read the book

1.5.5 Passive

(70) *Active vs. passive sentences*:

a. John took the book
b. John has taken the book
c. The book was taken (by John)

Assumption:
Passivization is an operation in the lexicon affecting V: In English, passivization involves (i) the deletion of the highest-ranked Θ-role of the argument structure (hence, of the lowest-ranked selectional feature); and (ii) the deletion of an [*acc*] feature (if one is present). Passivization is indicated by [pas] on the passivized V. (The optional *by*-phrase is notoriously difficult to handle formally, due to its unclear status as an argument or modifier, and will be ignored here.)

(71) *Active vs. passive forms of V*:

a. took: { [V], [3pers,–pl,+masc], [+fin,+past], [*D*] > [*D*], [*acc*] }
b. taken: { [V], [–fin,+part], [*D*] > [*D*], [*acc*] }
c. taken: { [V], [–fin,+part], [pas], [*D*] }

(72) *A LA of (70-c)*:

a. taken: { [V], [–fin,+part], [pas], [*D*] }
b. the: { [D], [3pers,–pl,–fem,–masc], [nom], [*N*] }
c. book: { [N], [3pers,–pl,–fem,–masc], [nom] }
d. was: { [T], [3pers,–pl,–fem,–masc], [+fin,+past], [*V*,*–fin*,*+part*,*pas*] > [*nom*] }

(73) *Derivation of (70-c)*:

a. Merge ([$_{\text{D}}$ the], [$_{\text{N}}$ book])→ [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]
b. Merge ([$_{\text{V}}$ taken], [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]) →
[$_{\text{VP}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]
c. Merge ([$_{\text{T}}$ was], [$_{\text{VP}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]])
→ [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{VP}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]]
d. Move ([$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]],
[$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{VP}}$ [$_{\text{V}}$ taken] [$_{\text{DP}}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ [$_{\text{D}}$ the] [$_{\text{NP}}$ book]] [$_{\text{T}'}$ [$_{\text{T}}$ was] [$_{\text{VP}}$ [$_{\text{V}}$ taken] t_1]]]

Note 1:
The auxiliary *has* does not have the feature [*pas*]; hence, it can never

select a passive VP.

Note 2:
There are other instances of *was*; compare (74-abc). In these cases, *was* lacks the feature [*pas*], and either does have the feature [*V*] (but rather [*N*], [*P*] or [*A*]), or it has [*V*] plus [*ing*].

(74) *Other instances of be*:
   a. John was standing in the shadow
   b. Mary was a nice woman
   c. John was in Texas
   d. Mary was proud of John

Question:
Why does the sole DP argument of a passivized transitive construction have to be [nom], and cannot be [acc] in the LA?

Answer: [Case] on a DP is uninterpretable and must be deleted in syntax. If *the* and *book* bear the feature [acc] in the LA in (72), both [acc] and [*nom(Spec)*] on T will not be deleted, and the Interpretability Condition is violated.

1.6 Complementizer Phrases

Problem:
How can sentences like those in (75) be generated?

(75) a. I think that she likes John
   b. I think she likes John
   c. She wonders whether John likes her

Assumption:
LIs like *that* and *whether* in (75-ac) are C (complementizer) elements that select T. (75-b) has an empty C element that also selects T.

(76) *LA of (75-a)*:
   a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
   b. think: { [V], [1pers,–pl,+masc], [+fin,–past], [*C*,*–wh*] > [*D*] }
   c. John: { [N], [3pers,–pl,+masc], [acc] }
   d. she: { [D], [3pers,–pl,+fem], [nom] }

e. I: { [D], [1pers,–pl,+masc], [nom] }
f. Ø: { [D], [3pers,–pl,+masc], [acc], [*N*] }
g. Ø: { [T], [3pers,–pl,+fem], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
h. Ø: { [T], [1pers,–pl,+masc], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
i. that: { [C], [–wh], [+fin], [*T*,*+fin*] }

(77) *Derivation of (75-a)*:

a. Merge ([$_{\text{D}}$ Ø], [$_{\text{N}}$ John]) → [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]
b. Merge ([$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]], [$_{\text{V}}$ likes]) → [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]
c. Merge ([$_{\text{D}}$ she], [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]])
→ [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]
d. Merge ([$_{\text{T}}$ Ø], [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]])
→ [$_{\text{TP}}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]
e. Move ([$_{\text{DP}}$ she], [$_{\text{TP}}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ [$_{\text{DP}}$ she] [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]
f. Merge ([$_{\text{C}}$ that],
[$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]) →
[$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]
g. Merge ([$_{\text{V}}$ think], [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]])
→ [$_{\text{VP}}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]]
h. Merge ([$_{\text{D}}$ I], [$_{\text{VP}}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]])
→ [$_{\text{VP}}$ [$_{\text{DP}}$ I] [$_{\text{V}'}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]]]
i. Merge ([$_{\text{T}}$ Ø], [$_{\text{VP}}$ [$_{\text{DP}}$ I] [$_{\text{V}'}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]]])
→ [$_{\text{TP}}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ [$_{\text{DP}}$ I] [$_{\text{V}'}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{T}'}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ t$_1$ [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]]]]]]]]]]
j. Move ([$_{\text{DP}}$ I], [$_{\text{TP}}$ [$_{\text{T}}$ Ø] [$_{\text{VP}}$ [$_{\text{DP}}$ I] [$_{\text{V}'}$ [$_{\text{V}}$ think] [$_{\text{CP}}$ [$_{\text{C}}$ that] [$_{\text{TP}}$

[$_{DP_1}$ she] [$_{T'}$ [$_T$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_V$ likes] [$_{DP}$ [$_D$ Ø] [$_{NP}$ John]]]]]]]]]
)
→ [$_{TP}$ [$_{DP_2}$ I] [$_{T'}$ [$_T$ Ø] [$_{VP}$ t_2 [$_{V'}$ [$_V$ think] [$_{CP}$ [$_C$ that] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_T$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_V$ likes] [$_{DP}$ [$_D$ Ø] [$_{NP}$ John]]]]]]]]]

(78) *LA of (75-b)*:

a. likes: { [V], [3pers,–pl,+fem], [+fin,–past], [*D*] > [*D*], [*acc*] }
b. think: { [V], [1pers,–pl,+masc], [+fin,–past], [*C*,*–wh*] > [*D*] }
c. John: { [N], [3pers,–pl,+masc], [acc] }
d. she: { [D], [3pers,–pl,+fem], [nom] }
e. I: { [D], [1pers,–pl,+masc], [nom] }
f. Ø: { [D], [3pers,–pl,+masc], [acc], [*N*] }
g. Ø: { [T], [3pers,–pl,+fem], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
h. Ø: { [T], [1pers,–pl,+masc], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
i. Ø: { [C], [–wh], [+fin], [*T*,*+fin*] }

Observation:
The derivation of (75-b) is identical to that of (75-a).

Note:
Alternatively, one might assume that verbs like *think* can optionally select T instead of C. Under this assumption, postulation of an empty C element is not necessary. Whether such elements are needed or not must be determined on the basis of independent evidence (e.g., are there constraints that refer to C, CP, or SpecC in examples like (75-b)?).

(79) *LA of (75-c)*:

a. likes: { [V], [3pers,–pl,+masc], [+fin,–past], [*D*] > [*D*], [*acc*] }
b. wonders: { [V], [3pers,–pl,+fem], [+fin,–past], [*C*,*+wh*] > [*D*] }
c. John: { [N], [3pers,–pl,+masc], [nom] }
d. she: { [D], [3pers,–pl,+fem], [nom] }
e. her: { [D], [3pers,–pl,+fem], [acc] }
f. Ø: { [D], [3pers,–pl,+masc], [nom], [*N*] }
g. Ø: { [T], [3pers,–pl,+masc], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }

h. Ø: { [T], [3pers,–pl,+fem], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
i. whether: { [C], [–wh], [+fin], [*T*,*+fin*] }

(80) *Derivation of (75-c)*:

a. Merge ([$_{D}$ Ø], [$_{N}$ John]) → [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]
b. Merge ([$_{V}$ likes], [$_{D}$ her]) → [$_{VP}$ [$_{V}$ likes] [$_{DP}$ her]]
c. Merge ([$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]], [$_{VP}$ [$_{V}$ likes] [$_{DP}$ her]])
→ [$_{VP}$ [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]
d. Merge ([$_{T}$ Ø], [$_{VP}$ [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]])
→ [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]
e. Move ([$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]],
[$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]])
→ [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]
f. Merge ([$_{C}$ whether],
[$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]])
→ [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]
g. Merge ([$_{V}$ wonders],
[$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]])
→ [$_{VP}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]]
h. Merge ([$_{D}$ she],
[$_{VP}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]])
→ [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]]]
i. Merge ([$_{T}$ Ø],
[$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]]])
→ [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ likes] [$_{DP}$ her]]]]]]]]]
j. Move ([$_{DP}$ she],
[$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ wonders] [$_{CP}$ [$_{C}$ whether] [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t_1 [$_{V'}$

$[_V$ likes] $[_{DP}$ her]]]]]]]])
→ $[_{TP}$ $[_{DP_2}$ she] $[_{T'}$ $[_T$ Ø] $[_{VP}$ t_2 $[_{V'}$ $[_V$ wonders] $[_{CP}$ $[_C$ whether] $[_{TP}$ $[_{DP_1}$ $[_D$ Ø] $[_{NP}$ John]] $[_{T'}$ $[_T$ Ø] $[_{VP}$ t_1 $[_{V'}$ $[_V$ likes] $[_{DP}$ her]]]]]]]]]

Note:
Whereas verbs like *think* can only embed [–wh] CPs, and verbs like *wonder* can only embed [+wh] CPs, a verb like *know* can select either [–wh] CPs or [+wh] CPs.

(81) *Variable selectional features of 'know'*:
a. Mary knows that John likes her
b. Mary knows whether John likes her

Note:
Nouns can also embed CPs, i.e., N can have a [*C*] feature.

(82) a. John heard a rumour that you had read this book
b. Mary expressed the feeling that the meeting should be held

Problem:
The present system generates categories of variable complexity: NPs, DPs, VPs, TPs, CPs (= clauses). Recall that the main task of syntax is to determine what is a grammatical (well-formed) sentence in a given language, and what is not. But what qualifies as a sentence, and what does not? E.g., an embedded CP like *that she likes John* is not a well-formed sentence; neither is an NP like *rumour that you had read this book*, or indeed a VP like *arrived John*. Essentially, the question is: When does a derivation stop?

Solution:
There is a root C element in every LA that selects finite T, and that cannot be embedded any further. A derivation stops, yielding a sentence (that can then be grammatical or ungrammatical), when root C is merged.

(83) *Root Constraint*:
Every LA has a C marked [root]. No LI can have a feature [*root*].

(84) *Extended LA of (43-b)* (compare (55)):
a. taken: { [V], [–fin,+part], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. the: { [D], [3pers,–pl,–fem,–masc], [acc], [*N*] }
d. book: { [N], [3pers,–pl,–fem,–masc], [acc] }

e. has: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*–fin*,*+part*] > [*nom*] }
f. Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*].

(85) *Derivation of (43-b)* (compare (43-b)):
...
e. Move ([$_{DP}$ she],
[$_{TP}$ [$_{T}$ has] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ taken] [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]]]]
)
→ [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ has] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ taken] [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]]]]]
f. Merge ([$_{C}$ Ø],
[$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ has] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ taken] [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]]]]]) →
[$_{CP}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ has] [$_{VP}$ t_1 [$_{V'}$ [$_{V}$ taken] [$_{DP}$ [$_{D}$ the] [$_{NP}$ book]]]]]]]

1.7 Adjective Phrases

1.7.1 Adjectives as Predicates

(86) a. John is nice
b. She was proud of him

(87) *LA of (86-a)*:
a. John: { [N], [3pers,–pl,+masc], [nom] }
b. Ø: { [D], [3pers,–pl,+masc], [nom], [*N*] }
c. is: { [T], [3pers,–pl,+masc], [+fin,–past], [*A*] > [*nom*] }
d. nice: { [A], [*D*] }
e. Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*].

(88) *Derivation of (86-a)*:
a. Merge ([$_{D}$ Ø], [$_{N}$ John]) → [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]
b. Merge ([$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]], [$_{A}$ nice]) → [$_{AP}$ [$_{A}$ nice] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]
c. Merge ([$_{T}$ is], [$_{AP}$ [$_{A}$ nice] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]])
→ [$_{TP}$ [$_{T}$ is] [$_{AP}$ [$_{A}$ nice] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]]
d. Move ([$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]], [$_{TP}$ [$_{T}$ is] [$_{AP}$ [$_{A}$ nice] [$_{DP}$ [$_{D}$ Ø] [$_{NP}$ John]]]])
→ [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ is] [$_{AP}$ [$_{A}$ nice] t_1]]]
e. Merge ([$_{C}$ Ø], [$_{TP}$ [$_{DP_1}$ [$_{D}$ Ø] [$_{NP}$ John]] [$_{T'}$ [$_{T}$ is] [$_{AP}$ [$_{A}$ nice]

t$_1$]]])
→ [$_{\text{CP}}$ [$_{\text{C}}$ Ø], [$_{\text{TP}}$ [$_{\text{DP}_1}$ [$_{\text{D}}$ Ø] [$_{\text{NP}}$ John]] [$_{\text{T}'}$ [$_{\text{T}}$ is] [$_{\text{AP}}$ [$_{\text{A}}$ nice] t$_1$]]]]

(89) *LA of (86-b)*:
a. she: { [D], [3pers,–pl,+fem], [nom] }
b. him: { [D], [3pers,–pl,+masc], [acc] }
c. was: { [T], [3pers,–pl,+fem], [+fin,+past], [*A*] > [*nom*] }
d. proud: { [A], [*P:of*] > [*D*] }
e. of: { [P], [*D*], [*acc*] }
f. Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*].

(90) *Derivation of (86-b)*:
a. Merge ([$_{\text{P}}$ of], [$_{\text{D}}$ him]) → [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]
b. Merge ([$_{\text{A}}$ proud], [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]) → [$_{\text{AP}}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]
c. Merge ([$_{\text{D}}$ she], [$_{\text{AP}}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]])
→ [$_{\text{AP}}$ [$_{\text{DP}}$ she] [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]
d. Merge ([$_{\text{T}}$ was], [$_{\text{AP}}$ [$_{\text{DP}}$ she] [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]])
→ [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{AP}}$ [$_{\text{DP}}$ she] [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]]
e. Move ([$_{\text{DP}}$ she], [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{AP}}$ [$_{\text{DP}}$ she] [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]])
→ [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{AP}}$ t$_1$ [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]]]
f. Merge ([$_{\text{C}}$ Ø],
[$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{AP}}$ t$_1$ [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]]]) →
[$_{\text{CP}}$ [$_{\text{C}}$ Ø], [$_{\text{TP}}$ [$_{\text{DP}_1}$ she] [$_{\text{TP}}$ [$_{\text{T}}$ was] [$_{\text{AP}}$ t$_1$ [$_{\text{A}'}$ [$_{\text{A}}$ proud] [$_{\text{PP}}$ [$_{\text{P}}$ of] [$_{\text{DP}}$ him]]]]]]]

Note 1:
Similar analyses can be given for N predicates. As it stands, a noun like *woman* must have two lexical entries, corresponding to its use as a predicate or argument.

(91) *Two uses of N*:
a. The woman likes John
b. Mary is a woman

(92) *LIs*:

a. woman: { [N], [3pers,–pl,+fem] }
b. woman: { [N], [3pers,–pl,+fem], [nom], [*D*] }

1.7.2 Adjectives as Modifiers

Note:
As it stands, sentences like (93) cannot yet be accounted for.

(93) a. He met a nice woman
b. She probably likes him

Problem:
The problem is that there does not seem to be a selectional relation between the adjective *nice* and the noun *man*, or between the adverb *probably* and the verb *likes*. Suppose there were such a relation. First, it seems clear that the selectional feature would have to be on *nice* and *probably*, respectively (compare, e.g., the Linearization Constraint on Merge). (See, e.g., Longobardi (2001) on (93-a), Alexiadou (1997) and Cinque (1999) for cases like (93-b)). We would expect that D and T, respectively, would be the head of the projection after Merge. Consequently, D would have to select either A or N depending on whether *nice* is present or not in (93-a); and T would have to select either A or V, depending on whether or not *probably* is present in (93-b). This does not seem plausible.

Assumption:
The system has to be revised so as to permit modification. In particular, the notion of head and the Economy Constraint on Merge must be revised.

(94) *Head*:
a. If γ has been created by Merge of α and β, α is the *head* of γ (and γ is a *projection* of α) if (i) or (ii) holds:
(i) α selects β.
(ii) α is semantically modified by β.
b. A head and its projection share morpho-syntactic features.

Note:
In principle, there might be a situation in which a projection γ has two heads: α selects β, and α semantically modifies β. In practice, this situation does not arise.

(95) *Economy Constraint on Merge* (revised):
Merge can only apply if (a) or (b) holds:

a. Merge deletes the selectional feature of a LI.
b. Merge results in semantic modification.

Note:
In addition to the structural notions of complement and specifier, a third structural notion of (syntactic) modifier can now be introduced. It is worth noting that modifiers are not specifiers – the structural position may be the same, but selectional features are not involved in Merge.

(96) *Complements, specifiers, and modifiers* (see (13)):
a. β is a *complement* of α iff β has been merged with a LI α that selects it.
b. β is a *specifier* of α iff β has been merged with a non-LI α that selects it.
c. β is a *modifier* of α iff β has been merged with α, where α is the head and β neither selects nor is selected by α.

Consequence:
The linear precedence statements must extended to modifiers. For instance, A modifiers typically precede N, P modifiers follow N in English. In addition, preceding A modifiers can be merged with N only if they share ϕ-features and Case features with N. (As before, such agreement becomes much more obvious in more richly inflecting languages like German or Russian.)

(97) *LA of (93-a)*:
a. met: { [V], [3pers,–pl,+masc], [+fin,+past], [*D*] > [*D*], [*acc*] }
b. he: { [D], [3pers,–pl,+masc], [nom] }
c. a: { [D], [3pers,–pl,+fem], [acc], [*N*] }
d. woman: { [N], [3pers,–pl,+fem], [acc] }
e. nice: { [A], [3pers,–pl,+fem], [acc] }
f. Ø: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*+fin*,*+past*] > [*nom*] }
g. Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*].

(98) *Derivation of (93-a)*:
a. Merge ([$_{\text{A}}$ nice], [$_{\text{N}}$ woman]) → [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]
b. Merge ([$_{\text{D}}$ a], [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]) → [$_{\text{DP}}$ [$_{\text{D}}$ a] [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]]
c. Merge ([$_{\text{V}}$ met], [$_{\text{DP}}$ [$_{\text{D}}$ a] [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]])
→ [$_{\text{VP}}$ [$_{\text{V}}$ met] [$_{\text{DP}}$ [$_{\text{D}}$ a] [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]]]
d. Merge ([$_{\text{D}}$ he], [$_{\text{VP}}$ [$_{\text{V}}$ met] [$_{\text{DP}}$ [$_{\text{D}}$ a] [$_{\text{NP}}$ [$_{\text{AP}}$ nice] [$_{\text{N}}$ woman]]]]

)
→ $[_{VP}$ $[_{DP}$ he] $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]

e. Merge ($[_{T}$ Ø], $[_{VP}$ $[_{DP}$ he] $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]])
→ $[_{TP}$ $[_{T}$ Ø] $[_{VP}$ $[_{DP}$ he] $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]]

f. Move ($[_{DP_1}$ he],
$[_{TP}$ $[_{T}$ Ø] $[_{VP}$ $[_{DP}$ he] $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]]) →
$[_{TP}$ $[_{DP_1}$ he] $[_{T'}$ $[_{T}$ Ø] $[_{VP}$ t_1 $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]]]

g. Merge ($[_{C}$ Ø], $[_{TP}$ $[_{DP_1}$ he] $[_{T'}$ $[_{T}$ Ø] $[_{VP}$ t_1 $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]]]) →
$[_{CP}$ $[_{C}$ Ø] $[_{TP}$ $[_{DP_1}$ he] $[_{T'}$ $[_{T}$ Ø] $[_{VP}$ t_1 $[_{V'}$ $[_{V}$ met] $[_{DP}$ $[_{D}$ a] $[_{NP}$ $[_{AP}$ nice] $[_{N}$ woman]]]]]]]]

Note:
The reasoning is similar for (93-b).

1.8 Prepositional Phrases

Observation 1:
P can act as a predicate.

(99) *P as a predicate*:
John was in Texas

Observation 2:
P can act as a modifier of N or V.

(100) *P as a modifier*:
a. The woman in Texas likes John
b. She met him in Paris

1.9 Outlook

Infinitives:
The structure of infinitival clauses has not been addressed.

General question:

Do we actually need LAs? It seems that the analyses work just as well if the derivation has access to the whole lexicon/morphology component throughout (see work by Collins, Frampton & Gutmann). As it stands, most initially conceivable LAs will lead to derivational crashes. Chomsky suggests that we need LAs because they reduce complexity. If so, can or should there be meta-constraints on LAs that minimize derivational crashes? (A further argument for LAs arises if transderivational or translocal constraints are adopted. See below.)

1.10 Exercises

Exercise 1:
Give LAs and derivations of the following three sentences ((101-c) = (93-b), (101-d) = (99)):

(101) a. Few riders have won the race
b. I suspect that a lengthy discussion will ensue
c. She probably likes him
d. John was in Texas

Exercise 2:
The present system cannot yet correctly generate sentences with ditransitive verbs like (102-a), and sentences with question words like (102-b). Show this (based on appropriate LAs and derivations). Which (ill-formed) sentences are generated instead?

(102) a. He has given a book to Mary
b. What can I say?

Chapter 2

Classic Constraints

2.1 Introduction

Observation:
In addition to the lexicon and structure-building operations, a derivational approach to syntax along the lines sketched in [1] crucially relies on constraints. There are various constraint types. Most importantly, constraints can be *local* or *non-local*.

(1) *Constraint types*:
  a. A local *derivational* constraint (Con^d) applies to syntactic (Merge or Move) operations.
  b. A local *representational* constraint (Con^r) ("filter") applies to an output representation.
  c. A *global* constraint (Con^g) applies to a whole derivation; it correlates non-adjacent steps in the derivation.
  d. A *translocal* constraint (Con^{tl}) applies to sets of output representations; it picks out an optimal output representation among competing output representations.
  e. A *transderivational* constraint (Con^{td}) applies to sets of derivations; it picks out an optimal derivation among competing derivations.

(2) *Complexity of constraint types*:
derivational constraints, representational constraints $<$ global constraints $<$ translocal constraints $<$ transderivational constraints

Strategy:
If constraint C_1 and constraint C_2 can account for a given phenomenon in the same way and C_1 is less complex than C_2 then, other things being

equal, choose C_1.

Note:
This strategy does not imply that transderivational, translocal, or global constraints should be abandoned.

A meta-constraint on constraints:
Constraints should be as general as possible.

Observation 1:
Most of the constraints used so far are derivational constraints (see, e.g., the *Economy Constraint on Merge*, or the *Linearization Constraint on Merge*).

(3) *Economy Constraint on Merge^d*:
Merge can only apply if it deletes the highest-ranked selectional feature of a LI.

(4) *Linearization Constraint on Merge^d*:
The output of Merge in language L_i must conform to the linear precedence statements of L_i.

Observation 2:
In contrast, the *Interpretability Condition* is a representational constraint: It prohibits uninterpretable features in the final output representation, but not during the derivation (where such features are in fact essential).

(5) *Interpretability Condition^r*:
Features on LIs that are uninterpretable at level R_i must be removed at level R_{i-1}.

However, this presupposes that we know in the syntax which features are semantically interpretable, and which ones are not. This assumption is not innocuous; but if there is good reason to abandon it, then we end up with the result that the Interpretability Condition is a much more complex constraint, viz., a global constraint that takes into account aspects of the derivation that are post-syntactic (i.e., semantic).

Observation 3:
The constraint *Timing of Feature Deletion* (aka "Earliness") is not a local derivational constraint, a local representational constraint, or a global constraint. To find out whether this constraint is respected by a derivation or not, one has to compare it with other derivations: A derivation respects (6)

if, among a class of competing derivations that need to be defined appropriately (e.g., in terms of the same LA), the deletion of a given feature occurs at the earliest step. Suppose, e.g., that derivations D_1, D_2, D_3 and D_4 compete, and a feature F is deleted in step 3 of D_1, step 4 of D_2, step 6 of D_3, and step 9 of D_4. Then, only D_1 respects Timing of Feature Deletion.

(6) *The Timing of Feature Deletion*^{td}:
Uninterpretable features are deleted as soon as possible.

Question:
What kind of constraint is the Inclusiveness Condition?

(7) *Inclusiveness Condition*:
Material that is not part of the lexical array (other LIs, additional features) is inaccessible throughout a derivation.

Answer:
The Inclusiveness Condition should best be viewed not as a syntactic constraint at all; rather, it is a meta-constraint grammars, i.e., a hypothesis about the nature of syntactic operations. (That said, the Inclusiveness Condition *could* in principle be viewed as a local derivational constraint that restricts every Merge or Move operation.)

2.2 Movement to SpecC

2.2.1 Wh-Movement

2.2.1.1 Embedded Clauses

Problem:
So far, a sentence like (8) cannot be generated:

(8) *Embedded wh-questions in English*:
I wonder what she read

Solution:
A [+wh] C item in English requires movement of a *wh*-phrase in *wh*-questions. It has a selectional feature [*Q*] that must be deleted under identity with a *wh*-phrase bearing a [Q]-feature in SpecC. (The [Q]-feature of *wh*-phrases is often called [wh]-feature. To avoid ambiguity – cf. the relation between V and C on the one hand, the relation between C and D on the other –, the two features are distinguished here.)

(9) *LA of (8)*:

a. read: { [V], [3pers,–pl,+fem], [+fin,+past], [*D*] > [*D*], [*acc*] }
b. she: { [D], [3pers,–pl,+fem], [nom] }
c. what: { [D], [3pers,–pl,–fem,–masc], [acc], [Q] }
d. Ø: { [T], [3pers,–pl,+fem], [+fin,+past], [*V*,*+fin*,*+past*] > [*nom*] }
e. Ø: { [C], [+wh], [+fin], [*T*,*+fin*] > [*Q*].
f. I: { [D], [1pers,–pl,+masc], [nom] }
g. wonder: { [V], [1pers,–pl,+masc], [+fin,–past], [*C*,*+wh*] > [*D*] }
h. Ø: { [T], [1pers,–pl,+masc], [+fin,–past], [*V*,*+fin*,*–past*] > [*nom*] }
i. Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*].

(10) *Derivation of (8)*:

a. Merge ([$_{V}$ read], [$_{D}$ what]) → [$_{VP}$ [$_{V}$ read] [$_{DP}$ what]]
b. Merge ([$_{D}$ she], [$_{VP}$ [$_{V}$ read] [$_{DP}$ what]]) → [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]
c. Merge ([$_{T}$ Ø], [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]])
→ [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]]
d. Move ([$_{DP}$ she], [$_{TP}$ [$_{T}$ Ø] [$_{VP}$ [$_{DP}$ she] [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]])
→ [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]]]
e. Merge ([$_{C}$ Ø], [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]]])
→ [$_{CP}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]]]]
f. Move ([$_{DP}$ what],
[$_{CP}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] [$_{DP}$ what]]]]]])
→ [$_{CP}$ [$_{DP_2}$ what] [$_{C'}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] t$_2$]]]]]]
g. Merge ([$_{V}$ wonder],
[$_{CP}$ [$_{DP_2}$ what] [$_{C'}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{V'}$ t$_1$ [$_{VP}$ [$_{V}$ read] t$_2$]]]]]])
→ [$_{VP}$ [$_{V}$ wonder] [$_{CP}$ [$_{DP_2}$ what] [$_{C'}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] t$_2$]]]]]]]
... → [$_{CP}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_3}$ I] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_3$ [$_{V'}$ [$_{V}$ wonder] [$_{CP}$ [$_{DP_2}$ what] [$_{C'}$ [$_{C}$ Ø] [$_{TP}$ [$_{DP_1}$ she] [$_{T'}$ [$_{T}$ Ø] [$_{VP}$ t$_1$ [$_{V'}$ [$_{V}$ read] t$_2$]]]]]]]]]]]

(11) *More complex wh-phrases*:
Which book will she buy ?

(12) *Partial LA of (11)*:
a. which: { [D], [3pers,–pl,–fem,–masc], [acc], [Q], [*N*] }
b. book: { [N], [3pers,–pl,–fem,–masc], [acc] }

2.2.1.2 Root Clauses

Problem:
Root clauses pose an additional problem. There are two Move operations to the C domain: The *wh*-phrase moves as before; in addition, a finite auxiliary or modal verb is fronted. (Movement of a finite main verb is impossible in this context; cf. **What said she?*. In this case, a dummy auxiliary *do* must be inserted: *What did she say?*; this is called "do-support".)

(13) What has she said ?

Assumption:
There are two types of movement:
(i) XP movement = movement of a maximal projection to a specifier position.
(ii) X (head) movement = movement of a minimal projection (= a LI) to a LI position, via *adjunction* to the LI.

(14) *Intended structure of (13)*:
$[_{CP}$ $[_{DP_2}$ what] $[_{C}$ $[_{T_3}$ has] $[_{C}$ Ø]] $[_{TP}$ $[_{DP_1}$ she] $[_{T'}$ t_3 $[_{VP}$ t_1 $[_{V'}$ $[_{V}$ said] t_2]]]]]

Note:
So far, it has been (more or less tacitly) assumed that only maximal projections can be moved. Given that head movement also exists, further assumptions must be made.

(15) *XP vs. X movement*:
a. A feature [*F*] can only trigger XP movement.
b. A feature [*F-LI*] can only trigger movement of a LI.

(16) *Structure Preservation Principle*:
a. XP movement ends up in a specifier position.
b. LI movement ends up in an adjunction position of another LI.

Assumption:
A root C bearing [*Q*] has a [*T-LI*] feature that also triggers movement of the auxiliary.

(17) *LA of (13)*:
   a. said: { [V], [–fin,+part], [*D*] > [*D*], [*acc*] }
   b. she: { [D], [3pers,–pl,+fem], [nom] }
   c. has: { [T], [3pers,–pl,+masc], [+fin,+past], [*V*,*–fin*,*+part*] }
   d. what: { [D], [3pers,–pl,–fem,–masc], [acc], [Q] }
   e. Ø: { [C], [root], [+wh], [+fin], [*T*,*+fin*] > [*T-LI*] > [*Q*].

Note:
So far, it is only predicted that an empty [+wh] C element requires movement of a bare T. It does not yet follow that T is phonologically empty when a local *wh*-subject is moved, and is realized by an appropriate form of *do* otherwise. (And an attempt to account for this will not be made here.)

Question:
Can it be ensured that C in (17-e) never triggers two Merge operations with separate T LIs, rather than Merge with TP and Move of T?

Answer:
This follows from the assumption that there cannot be more than one T element per clause. Incidentally, similar questions arise with movement to SpecT and SpecC.

2.2.2 Topicalization

Observation:
Topicalization is similar to *wh*-movement, but it is movement to a [–wh] SpecC position. It is not accompanied by head movement in English.

(18) *Topicalization in English*:
   a. [$_{\text{CP}}$ John$_1$ C [$_{\text{TP}}$ she does not really like t$_1$]]
   b. I think that [$_{\text{CP}}$ John$_1$ C [$_{\text{TP}}$ she does not really like t$_1$]]

Note:
Topicalization systematically goes hand in hand with movement of the finite verb to C in German, Dutch, and the Scandinavian languages:

(19) *Topicalization in German*:

a. [$_{CP}$ Den Fritz$_1$ mag$_2$ [$_{TP}$ sie sehr t_1 t_2]]
ART Fritz$_{acc}$ likes she$_{nom}$ much
b. Ich glaube [$_{CP}$ den Fritz$_1$ mag$_2$ [$_{TP}$ sie sehr t_1 t_2]]
I think ART Fritz$_{acc}$ likes she$_{nom}$ much

Note:
German has linear precedence statements that are different from those of English. In particular, V (and perhaps T) heads follow their specifiers *and* their complements.

(20) *Verb-final VPs in German*:
a. ... dass Fritz Maria mag
that Fritz$_{nom}$ Maria$_{acc}$ likes
b. Den Fritz$_1$ hat$_2$ sie sehr t_1 gemocht t_2
ART Fritz$_{acc}$ has she$_{nom}$ much liked

Analysis:
Topicalization is triggered by a [*top*] feature on C and a corresponding [top] feature on some other XP. C is always marked [*V-LI*,*+fin*] in German if it bears the feature [root] (the *verb-second* effect). (Assumption: Auxiliaries and modals are also [V] in German.)

2.2.3 Relativization

Observation:
Relativization is similar to *wh*-movement and topicalization; it moves a relative pronoun (or relative phrase) to SpecC. The relative clause itself is a modifier of an N; it follows N.

(21) *Relativization*:
a. I know [$_{DP}$ a man [$_{CP}$ who$_1$ C [$_{TP}$ t_1' T [$_{VP}$ t_1 likes cars]]]]
b. She likes [$_{DP}$ the book [$_{CP}$ which$_2$ John gave her t_2]]

Analysis:
(i) A relative pronoun (D) has the feature [rel].
(ii) The head of a relative clause (C) has the feature [*rel*].

2.3 The A-over-A Principle

(22) *A-over-A Principle*[d] (Chomsky (1964)):
In a structure ... [$_A$ ... [$_A$ ...] ...] ..., an operation can only affect the higher, more inclusive category A.

Note:
The A-over-A Principle is a local derivational constraint. To find out whether a given derivation respects it or not, each (Move) operation must be checked, by taking into account the phrase marker constructed so far.

(23) *A first consequence of the A-over-A Principle*:
  a. [$_{\text{DP}_1}$ My letter to [$_{\text{DP}_2}$ a friend in Italy]] got lost
  b. *[$_{\text{DP}_2}$ Who] did [$_{\text{DP}_1}$ my letter to t_2] get lost ?
  c. [$_{\text{DP}_1}$ Which letter to [$_{\text{DP}_2}$ a friend in Italy]] got lost?
  d. *John is the friend [$_{\text{DP}_2}$ who] C [$_{\text{DP}_1}$ my letter to t_2] got lost
  e. This is the letter [$_{\text{DP}_1}$ which] t_1 got lost

(24) *Another consequence of the A-over-A Principle*:
  a. John heard [$_{\text{DP}_1}$ a rumour that you had read [$_{\text{DP}_2}$ this book]]
  b. *[$_{\text{DP}_2}$ What] did John hear [$_{\text{DP}_1}$ a rumour that you had read t_2]] ?
  c. [$_{\text{DP}_1}$ Which rumour that you had read [$_{\text{DP}_2}$ this book]] did John hear ?
  d. *This is a book [$_{\text{DP}_2}$ which] John heard [$_{\text{DP}_1}$ a rumour that you had read t_2]
  e. This a rumour [$_{\text{DP}_1}$ which] John heard t_1

Note:
The A-over-A Principle can be reformulated as a representational constraint on outputs.

(25) *A-over-A Principle^r* (representational version):
  *... A_2 ... [$_{A_1}$ ... t_2 ...] ...] ...

Note:
Crucially, this formulation relies on the existence of traces, and this is in fact one of the two main reasons why one would want to postulate traces in the first place (the other main reason being that traces are relevant for semantic interpretation).

Motivating traces:
Traces are needed by representational constraints.

Problem:
The A-over-A Principle is too strong and too weak. The first problem is potentially severe; the second problem makes the A-over-A Principle look less plausible.

(26) *Well-formed DP-over-DP examples ruled out by the A-over-A Principle*:

a. [$_{DP_2}$ Who would you approve of [$_{DP_1}$ my seeing t_2]] ?
b. [$_{DP_2}$ Which author] did you read [$_{DP_1}$ a book about t_2] ?

(27) *Well-formed CP-over-CP examples ruled out by the A-over-A Principle*:

a. John wouldn't say [$_{CP_1}$ that Mary thinks [$_{CP_2}$ that Bill is nice]]
b. [$_{CP_2}$ That Bill is nice] John wouldn't say [$_{CP_1}$ that Mary thinks t_2]
c. Fritz hat behauptet [$_{CP_1}$ Maria würde denken [$_{CP_2}$ dass er nett ist]]
Fritz$_{nom}$ has claimed Maria$_{nom}$ would think that he nice is
d. [$_{CP_2}$ Dass er nett ist] hat Fritz behauptet [$_{CP_1}$ würde Maria denken t_2]

(28) *Well-formed VP-over-VP (-over VP) examples ruled out by the A-over-A Principle*:

a. Fritz hat [$_{VP_1}$ [$_{VP_2}$ zu arbeiten] versucht]
Fritz$_{nom}$ has to work tried
b. [$_{VP_2}$ Zu arbeiten] hat Fritz [$_{VP_1}$ t_2 versucht]
to work has Fritz$_{nom}$ tried
c. [$_{VP_1}$ [$_{VP_2}$ Zu arbeiten] versucht] hat Fritz t_1
to work tried has Fritz$_{nom}$
d. Ich [$_{V_3}$ denke] nicht [$_{VP_0}$ t_3 [$_{CP}$ dass er [$_{VP_1}$ [$_{VP_2}$ zu arbeiten] versucht] hat]]
I think not that he to work tried has
e. ?[$_{VP_1}$ [$_{VP_2}$ Zu arbeiten] versucht] denke ich nicht [$_{VP_0}$ t_4 [$_{CP}$ dass er t_1 hat]]
to work tried think I not that he has
f. ?[$_{VP_2}$ Zu arbeiten] denke ich nicht [$_{VP_0}$ t_4 [$_{CP}$ dass er [$_{VP_1}$ t_2 versucht] hat]]
to work think I now that he tried has

(29) *An ill-formed example not ruled out by the A-over-A Principle – AP movement from DP*:

a. You have [$_{DP_1}$ a [$_{AP_2}$ very intelligent] sister]
b. [$_{DP_1}$ [$_{AP_2}$ How intelligent] a t_2 sister] do you have ?
c. *[$_{AP_2}$ How intelligent] do you have [$_{DP}$ a t_2 sister] ?

(30) *Another ill-formed example not ruled out by the A-over-A Principle – DP movement from PP ('preposition stranding')*:

a. Sie spielt [$_{PP_1}$ mit [$_{DP_2}$ dem grünen Auto]]
she plays with the green car
b. [$_{PP_1}$ Mit [$_{DP_2}$ welchem Auto]] spielt sie t_1 ?
with which car plays she
c. *[$_{DP_2}$ Welchem Auto] spielt sie [$_{PP_1}$ mit t_2] ?
which car plays she with
d. [$_{PP_1}$ Mit [$_{DP_2}$ dem grünen Auto]] spielt sie t_1
with the green car plays she
e. *[$_{DP_2}$ Diesem Auto] spielt sie [$_{PP_1}$ mit t_2]
this car plays she with

Outlook: the future:
The A-over-A Principle is formulated in terms of categorial features. The selectional features triggering movement that have been adopted so far ([*D*]/[*nom*], [*Q*], [*top*], [*rel*]) are not (necessarily) categorial, though. What would happen if the A-over-A Principle were revised as an F-over-F Principle?

(31) *F-over-F Principle*[d]:
In a structure $\alpha_{[*F*]}$... [$_{\beta_{[F]}}$ ... [$_{\gamma_{[F]}}$ ...] ...] ..., movement to [$_{*F*}$] can only affect the category bearing the [F] feature that is closer to [*F*].

Note:
This is in fact (a subcase of) a constraint that is widely adopted in most recent versions of the minimalist program (see below).

Back to the sixties:
In reaction to Chomsky's A-over-A Principle, Ross (1967) developed a theory of *islands*, i.e., categories that are opaque for movement.

2.4 The Complex NP Constraint

(32) *Complex NP Constraint*[d] (Ross (1967)):
No element contained in a CP dominated by a DP may be moved out of that DP.

Note on terminology:
It was a standard assumption until the late eighties that NP dominates DP, not DP NP, as assumed here (and in most current work). Hence, the original Complex NP Constraint is a constraint on movement from NP, not from DP. The constraint is still known under its original name, which is therefore also adopted here, even though "Complex DP Constraint" might be more appropriate. The Complex NP Constraint accounts for some of the data that motivated the A-over-A Principle.

(33) *A consequence of the Complex NP Constraint, relative clauses*:
a. *[$_{DP_1}$ Which book] did John meet [$_{DP_2}$ a child [$_{CP}$ who read t_1]] ?
b. *[$_{DP_1}$ Who] does Mary know [$_{DP_2}$ a girl [$_{CP}$ who is jealous of t_1]] ?

(34) *A consequence of the Complex NP Constraint, argument clauses* (see (24-b)):
a. ??[$_{DP_1}$ Which book] did John hear [$_{DP_2}$ a rumour [$_{CP}$ that you had read t_1]] ?
b. *[$_{PP_1}$ How] did John hear [$_{DP_2}$ a rumour [$_{CP}$ that you had fixed the car t_1]] ?
c. ?*The hat [$_{DP_1}$ which] I believed [$_{DP_2}$ the claim [$_{CP}$ that Otto was wearing t_1]] is red

Note:
Movement from argument clauses (selected categories) in complex DPs typically yields much better results than movement from relative clauses (non-selected, modifier categories). However, this does not hold for movement of modifiers themselves, which is completely impossible throughout (see (34-a) vs. (34-b)).

(35) *Complex NP Constraint^r* (representational version):
*... α_1 ... [$_{DP}$ ... [$_{CP}$ ... t_1 ...]] ...

2.5 The Sentential Subject Constraint

(36) *Sentential Subject Constraint^d* (Ross (1967)):
No element dominated by a CP may be moved out of that CP if that CP is a subject.

(37) *A consequence of the Sentential Subject Constraint*:
a. [$_{DP_1}$ Who] did the reporters expect [$_{CP}$ that the principal would fire t_1] ?

b. *[$_{DP_1}$ Who] was [$_{CP}$ that the principal would fire t_1] expected by the reporters ?

c. *[$_{DP_1}$ Who] did [$_{CP}$ that Mary was going out with t_1] bother you ?

(38) *Sentential Subject Constraint^r* (representational version):
*... α_1 ... [$_{CP}$ ... t_1 ...] ... if CP is a subject.

Note:
Given the terminology adopted so far, "subject" means "element in SpecT". However, movement to SpecT is triggered by [*nom*]. Does that mean that CPs actually bear abstract [nom] Case, so that the can move to SpecT if they are external arguments? If one does not want to make that assumption, the following options are available:
(i) The notion of subject is replaced by the notion of external argument in the formulation of the Sentential Subject Constraint; CPs are never in SpecT (they may be in VP or undergo topicalization).
(ii) CPs are in fact embedded by empty DPs that have abstract Case (compare Kiparsky & Kiparsky (1970)).

2.6 Subject Condition

Note:
The Sentential Subject Constraint can be generalized: DP subjects are also islands, even if they do not qualify as complex in the sense of the Complex NP Constraint.

(39) *Subject Condition^d* (Chomsky (1973), Huang (1982), Chomsky (1986), Freidin (1992)):
No element may be moved out of a subject.

(40) *Subject Condition* (see (23)):

a. *[$_{DP_2}$ Who(m)] has [$_{DP_1}$ a comment about t_2] annoyed you ?

b. *[$_{PP_3}$ About whom] has [$_{DP_1}$ a comment t_3] annoyed you ?

(41) *Subject Condition^r* (representational version):
*... α_1 ... [$_\beta$ ... t_1 ...] ... if β is a subject.

2.7 The Coordinate Structure Constraint

(42) *Coordinate Structure Constraint^d* (Ross (1967)):
In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(43) *A consequence of the Coordinate Structure Constraint – movement from a conjunct*:

a. John is $[_{AP}$ proud of $[_{DP_1}$ his father]] and $[_{AP}$ tired of $[_{DP_2}$ his mother]]

b. *$[_{DP_1}$ Who] is John $[_{AP}$ proud of t_1] and $[_{AP}$ tired of $[_{DP_2}$ his mother]] ?

c. *$[_{DP_2}$ Who] is John $[_{AP}$ proud of $[_{DP_1}$ his father]] and $[_{AP}$ tired of t_2] ?

Note:
It is not quite clear what the phrase structure of coordination looks like. An assumption that is sometimes made is that *and* is the head of a "coordination phrase", and this would get the word order facts right; but it also raises several problems. E.g.: What about coordinations with three conjuncts: α, β, and γ? If *and* is the head, how can the categorial features (like [A] in (43)) be visible for the selecting head (*is* bearing [*A*] in (43)?)

(44) *A second consequence of the Coordinate Structure Constraint – movement of a conjunct*:

a. John likes $[_{DP_1}$ Mary] and $[_{DP_2}$ Bill]

b. *$[_{DP_1}$ Who] does John like t_1 and $[_{DP_2}$ Bill] ?

c. *$[_{DP_2}$ Who] does John like $[_{DP_1}$ Mary] and t_2 ?

(45) *Coordinate Structure Constraint^r* (representational version):
*... α_1 ... $[_\beta$ ... t_1 ...] ..., where β is a coordinate structure.

Note:
This presupposes that a coordinate structure is a constituent. Indeed, it can be moved:

(46) *Coordinate structures are constituents*:

a. $[_{DP_1}$ Mary] and $[_{DP_2}$ Bill] are t in the garden

b. $[_{DP_1}$ Mary] and $[_{DP_2}$ Bill], John does not really like t

Note:
There is an interesting exception to the Coordinate Structure Constraint: If movement simultaneously affects both conjuncts, the Coordinate Structure Constraint does not hold. This is known as *Across-the-board rule application.* (See Ross (1967), Williams (1978), Gazdar (1981)).

(47) *Across-the-Board rule application*:

a. I wonder $[_{CP}$ $[_{DP_1}$ which books] John hates t_1 and Mary likes t_1

b. I know a man [$_{CP}$ [$_{DP_1}$ who] John [$_{VP}$ saw t_1] and [$_{VP}$ liked t_1]]
c. The doctor [$_{CP}$ [$_{DP_3}$ who] [$_{TP_1}$ John worked for t_3] and [$_{TP_2}$ Mary relied on t_3]] died

Problem:
It remains unclear how Across-the-board movement (two sources, one moved item) can be accounted for in the incremental approach adopted here.

2.8 The Upward Boundedness Constraint

Assumption:
Rightward movement (*extraposition*, *heavy NP (DP) shift*) exists. It is typically optional. (Some – optional – [*F*] features can only be deleted by movement to a right-peripheral specifier.)

(48) *Rightward movement*:
a. [$_{DP}$ The claim t_1] was refuted [$_{CP_1}$ that all languages are context-free]
b. John [$_{VP}$ returned t_1 [$_{PP}$ to the library]] [$_{DP_1}$ all the books [$_{CP}$ which he had borrowed]]
c. [$_{DP}$ A review t_1] came out yesterday [$_{PP_1}$ of this article]]

(49) *Upward Boundedness Constraint*^{d} (Ross (1967)):
No element that is moved rightward may be moved out of the next higher CP.

(50) *Upward Boundedness Constraint*^{r} (representational version):
*... [$_{CP}$ ... t_1 ...] ... α_1 ...

Note:
This constraint is also known as the *Right Roof Constraint* (see, e.g., Perlmutter & Soames (1979)).

(51) *A consequence of the Upward Boundedness Constraint*:
a. [$_{CP_0}$ It is catastrophic [$_{CP_1}$ that [$_{DP_2}$ a review [$_{PP_3}$ of this article]] came out yesterday]]
b. [$_{CP_0}$ [$_{CP_1}$ That [$_{DP_2}$ a review [$_{PP_3}$ of this article]] came out yesterday] is catastrophic]
c. [$_{CP_0}$ [$_{CP_1}$ That [$_{DP_2}$ a review t_3] came out yesterday [$_{PP_3}$ of this article]] is catastrophic]
d. *[$_{CP_0}$ [$_{CP_1}$ That [$_{DP_2}$ a review t_3] came out yesterday] is catastrophic] [$_{PP_3}$ of this article]

(52) *Another consequence of the Upward Boundedness Constraint*:

a. [$_{CP_0}$ Fritz denkt [$_{CP_1}$ dass Antje [$_{DP_2}$ den Versuch [$_{CP_3}$ mit
Fritz thinks that Antje the attempt with
fünf Bällen zu jonglieren]] aufgegeben hat] [$_{CP_4}$ weil er sie
five balls to juggle given up has because he her
nicht mehr sieht]]
not anymore sees

b. [$_{CP_0}$ Fritz denkt [$_{CP_1}$ dass Antje [$_{DP_2}$ den Versuch t_3]
Fritz thinks that Antje the attempt
aufgegeben hat] [$_{CP_3}$ mit fünf Bällen zu jonglieren] [$_{CP_4}$
given up has with five balls to juggle
weil er sie nicht mehr sieht]]
because he her not anymore sees

c. *[$_{CP_0}$ Fritz denkt [$_{CP_1}$ dass Antje [$_{DP_2}$ den Versuch t_3]
Fritz thinks that Antje the attempt
aufgegeben hat] [$_{CP_4}$ weil er sie nicht mehr sieht] [$_{CP_3}$
given up has because he her not anymore sees
mit fünf Bällen zu jonglieren]]
with five balls to juggle

2.9 The Left Branch Condition

(53) *Left Branch Condition^d* (Ross (1967)):
The leftmost item of an NP cannot be moved out of that NP.

(54) *Left Branch Condition^r* (representational version):
*... α_1 ... [$_{NP}$ t_1 ... N ...] ...

Note:
Like the Complex NP Constraint, the original Left Branch Condition presupposes a structure of nominal XPs that differs from the one adopted here, viz., (55-b) (where NP dominates DP) rather than (55-a) (where DP dominates NP), as assumed here and in most current literature.

(55) *DP vs. NP*:

a. [$_{DP}$ D [$_{NP}$ ... N ...] ...]
b. [$_{NP}$ [$_{DP}$ D] ... N ...]

(56) *A consequence of the Left Branch Condition under (55-b)*:

a. *[$_{DP_1}$ Which] did you buy [$_{NP}$ t_1 books] ?
b. *[$_{DP_1}$ Whose] did you meet [$_{NP}$ t_1 sister] ?

Note:
It seems that the Left Branch Condition is needed to rule out (56-a) only if structure (55-b) is adopted. If we assume structure (55-a), the prohibition against movement of *which* will not be needed because (a) the [*Q*] feature that triggers *wh*-movement does not permit head (LI) movement of D, and (b) if the whole DP moves, it has to carry the NP along. (Similar conclusions apply in the case of (56-b) if *whose* does not (fully) occupy SpecD – e.g., if *whose* is analyzed as *who* in SpecD plus *'s* in D.)

However:
The Left Branch Condition rules out sentences like (57-b) under either (55-b) or (55-a) if we understand "leftmost item" as "leftmost phonologically visible item".

(57) *A further consequence of the Left Branch Condition*:

a. Hans hat [$_{\text{DP}}$ D [$_{\text{NP}}$ [$_{\text{AP}_1}$ neue] Bücher]] gekauft
   Hans has new books bought

b. *[$_{\text{AP}_1}$ Neue] hat Hans [$_{\text{DP}}$ D [$_{\text{NP}}$ t_1 Bücher]] gekauft
   new has Hans books bought

Note:
Ross noted that there are Left Branch Condition violations in a number of languages; see, e.g., (58). Given (55-b), one can then simply assume that the Left Branch Condition does not hold in these languages; but it is a priori unclear how to reconcile the very existence of data such as those in (58) with the structure in (55-a).

(58) *Left Branch Condition violations in Russian*:

a. [$_{\text{NP}_1}$ [$_{\text{DP}_2}$ Čju] [$_{\text{N}}$ knigu]] ty čitaeš' t_1 ?
   whose book you read

b. [$_{\text{DP}_2}$ Čju] ty čitaeš' [$_{\text{NP}_1}$ t_2 [$_{\text{N}}$ knigu]] ?
   whose you read book

Note:
The Left Branch Condition can be generalized. This accounts for more data, but it also raises more problems.

(59) *Generalized Left Branch Condition*[d] (Ross (1967), Gazdar (1981)):
The leftmost item of an XP cannot be moved out of that XP.

(60) *Generalized Left Branch Condition effects, APs*:

a. [$_{\text{AP}_1}$ [$_{\text{XP}_2}$ How] sane] is John t_1 ?

b. *$[_{XP_2}$ How] is John $[_{AP_1}$ t_2 sane] ?
c. $[_{AP_1}$ $[_{XP_2}$ Ganz schön] neugierig] ist Maria t_1
d. *$[_{XP_2}$ Ganz schön] ist Maria $[_{AP_1}$ t_2 neugierig]

(61) *Generalized Left Branch Condition effects, TP ('that-trace effect')*:
a. $[_{DP_1}$ What] do you think $[_{CP}$ that John bought t_1] ?
b. $[_{DP_1}$ What] do you think $[_{CP}$ Ø John bought t_1] ?
c. *$[_{DP_1}$ Who] do you think $[_{CP}$ that $[_{TP}$ t_1 arrived]] ?
d. $[_{DP_1}$ Who] do you think $[_{CP}$ Ø $[_{TP}$ t_1 arrived]] ?

Note:
The *that*-trace configuration in (61-c) can be excluded; but unfortunately, the Generalized Left Branch Condition also excludes (61-d), which is well formed. Gazdar's solution: Movement from an embedded CP is only apparent here; the external argument *who* of *arrived* is in fact merged in the matrix VP domain. Furthermore, the analysis is incompatible with the idea that external arguments are merged in SpecV and move to SpecT.

(62) *Generalized Left Branch Condition effects, VP (problem)*:
a. $[_{CP}$ $[_{C}$ Ø] $[_{TP}$ $[_{DP_1}$ John] $[_{T}$ Ø] $[_{VP}$ t_1 likes Mary]]]
b. $[_{CP}$ $[_{DP_1}$ Who] $[_{C}$ Ø] $[_{TP}$ t_1' $[_{T}$ Ø] $[_{VP}$ t_1 likes Mary]]] ?

2.10 The Wh-Island Condition

(63) *Wh-Island Condition*^{d} (Chomsky (1973)):
Movement must not cross a CP with a *wh*-element in SpecC or C.

(64) *Wh-Island Condition*^{r} (representational version):
*... α_1 ... $[_{CP}$ β_2 ... t_1 ...] ..., where β is a *wh*-element in SpecC or C.

(65) *A consequence of the Wh-Island Condition*:
a. How_1 do you think $[_{CP}$ that Mary solved the problem t_1] ?
b. *How_1 do you wonder $[_{CP}$ whether Mary solved the problem t_1] ?
c. $[_{DP_1}$ Which book] do you think $[_{CP}$ that John read t_1] ?
d.?*$[_{DP_1}$ Which book] do you wonder $[_{CP}$ $[_{PP_2}$ to whom] John gave t_1 t_2] ?

Note:
Wh-Island effects are typically not that strong if the *wh*-clause is an infinitive and the moved item is a complement DP.

(66) *Weak Wh-Island Condition effects*:
??$[_{DP_1}$ Which book] don't you know $[_{CP}$ whether to read t_1] ?

Note:
A similar effect arises with topicalization to SpecC. Accordingly, a *Topic Island Condition* has been suggested, and further generalization seems possible.

(67) *Topic Island effects*:

a. [$_{\text{DP}_1}$ This book] Mary thinks that Bill gave t_1 [$_{\text{PP}_2}$ to John]

b. *[$_{\text{DP}_1}$ This book] Mary thinks that [$_{\text{PP}_2}$ to John] Bill gave t_1 t_2

c.(?)[$_{\text{DP}_1}$ Wen] denkst du [$_{\text{CP}}$ dass Maria t_1 mag] ?
whom think you that Maria likes

d. *[$_{\text{DP}_1}$ Wen] denkst du [$_{\text{CP}}$ Maria$_2$ mag$_3$ t_2 t_1 t_3] ?
whom think you Maria likes

2.11 The Superiority Condition

(68) *Superiority Condition*^{d} (Chomsky (1973)):
In a structure $\alpha_{[*F*]}$... [... $\beta_{[F]}$ ... [... $\gamma_{[F]}$ ...] ...] ..., movement to [*F*] can only affect the category bearing the [F] feature that is closer to [*F*].

Note:
The only difference to the (revised) A-over-A Principle (i.e., the F-over-F Principle) is that β c-commands γ in the Superiority Condition, wheres β dominates γ in the F-over-F Principle.

(69) *Superiority Condition*^{r} (representational version):
*... $\gamma_{[F]}$... [... $\beta_{[F]}$ ... [... t_γ ...] ...] ... if the head of which γ is the specifier bears a [*F*] feature in the LA.

(70) *A consequence of the Superiority Condition*:

a. Who$_1$ t_1 saw what$_2$?

b. *What$_2$ did who$_1$ see t_2 ?

c. I wonder [$_{\text{CP}}$ who$_1$ t_1 bought what$_2$]

d. *I wonder [$_{\text{CP}}$ what$_2$ who$_1$ bought t_2]

2.12 The Clause Non-final Incomplete Constituent Constraint

(71) *Clause Non-final Incomplete Constituent Constraint*^{d} (Kuno (1973)):
It is not possible to move any element of a category α (α = DP or CP) in a clause non-final position out of α if what is left over in α constitutes an incomplete α.

(72) *Incompleteness*:
A DP/CP α is incomplete if an obligatory element is missing.
(An obligatory element may, as a first approximation, be an element that is obligatorily selected.)

Origin:
Kuno suggests the Clause Non-final Incomplete Constituent Constraint as a more general version of the Sentential Subject Constraint, which it is therefore supposed to replace.

(73) *A consequence of the Clause Non-final Incomplete Constituent Constraint, object DPs*:

a. [$_{DP_1}$ Which man] did you buy [$_{DP}$ a picture of t_1] ? (see (26-b))
b. [$_{PP_2}$ Of which man] did John give [$_{DP}$ a picture t_2] to Bill ?
c. ?*[$_{DP_1}$ Which man] did John give [$_{DP}$ a picture of t_1] to Bill ?

Note:
In (73-a), the DP is clause-final; in (73-b), the DP counts as complete (recall that arguments of N are optional). Only in (73-c) are both requirements violated: The DP from which movement takes place is in a non-final position, and if movement occurs, it counts as incomplete (*of* has an obligatory [*D*] feature).

(74) *A consequence of the Clause Non-final Incomplete Constituent Constraint, subject DPs*:

a. [$_{DP_1}$ Which cars] did the explosion damage [$_{DP}$ the hoods of t_1] ?
b. [$_{PP_2}$ Of which cars] were [$_{DP}$ the hoods t_2] damaged by the explosion ?
c. *[$_{DP_1}$ Which cars] were [$_{DP}$ the hoods of t_1] damaged by the explosion ?

Note:
(74-b) is expected to be ungrammatical under the Subject Condition. However, it has been suggested that these kinds of PPs may in fact be merged outside the subject DP (see Cinque (1990)), in which case the Subject Condition would be compatible with (74-b) (and the Clause Non-final Incomplete Constituent Constraint would be vacuously fulfilled here).

(75) *A consequence of the Clause Non-final Incomplete Constituent Constraint, CPs* (see (37)):

a. [$_{DP_1}$ Who] did the reporters expect [$_{CP}$ that the principal would fire t_1] ?

b. [$_{\text{DP}_1}$ Who] was it expected by the reporters [$_{\text{CP}}$ that the principal would fire t_1] ?
c. *[$_{\text{DP}_1}$ Who] was [$_{\text{CP}}$ that the principal would fire t_1] expected by the reporters ?

Note:
The Clause Non-final Incomplete Constituent Constraint can be reformulated as a representational constraint on outputs.

(76) *Clause Non-final Incomplete Constituent Constraint*[r] (representational version):
*... α_1 ... [$_\beta$ ... t_1 ...] ... if (a)–(c) hold:
a. β = DP or CP.
b. β is in a clause non-final position.
c. β is incomplete.

(77) *An apparent problem*:
[$_{\text{DP}_1}$ Who] does John think [$_{\text{CP}_2}$ Mary has persuaded t_1 [$_{\text{CP}_3}$ that Bill is a spy]] ?

Note:
(77) does not violate the Clause Non-final Incomplete Constituent Constraint because the only CP from which movement takes place is CP_2; and CP_2 is incomplete after the movement operation, but it is in a clause-final position.

(78) *A real problem?*
a. [$_{\text{DP}_1}$ Which man] did you buy [$_{\text{DP}}$ a picture of t_1] from Mary ?
b. [$_{\text{DP}_1}$ Which tree] did you see [$_{\text{DP}}$ the leaves of t_1] in the yard ?

Note:
Kuno assumes that (78-ab) are well-formed, and he takes this to follow from the Clause Non-final Incomplete Constituent Constraint. The idea is that what is problematic about the starred data is "the fact that the incomplete ... phrases are followed by nonoptional elements [...] In [(78-ab)], ... incomplete ... phrases appear either clause-finally or, if not, are followed only by optional elements in the sentences." But does this follow from the constraint?

2.13 The Post-Sentential Subject Extraction Constraint

(79) *Post-Sentential Subject Extraction Constraint*[d] (Zaenen & Pinkham (1976)):
It is impossible to move a DP across a sentential subject.

Note:
Unlike the Clause Non-final Incomplete Constituent Constraint, this constraint is supposed to complement (rather than replace) the Sentential Subject Constraint.

(80) *A consequence for wh-movement*:
a. [$_{\text{DP}_1}$ Who] do you think [$_{\text{CP}_1}$ that [$_{\text{DP}_2}$ Bill's resignation] would surprise t_1] ?
b. *[$_{\text{DP}_1}$ Who] do you think [$_{\text{CP}_1}$ that [$_{\text{CP}_2}$ for Bill to resign] would surprise t_1] ?

(81) *A consequence for topicalization*:
a. [$_{\text{DP}_1}$ John] [$_{\text{DP}_2}$ Bill's resignation] would not surprise t_1
b. *[$_{\text{DP}_1}$ John] [$_{\text{CP}_2}$ for Bill to resign] would not surprise t_1

(82) *Post-Sentential Subject Extraction Constraint*^{r} (representational version):
*...α_1 ... [... β ... [... t_1 ...] ...] if β is a sentential subject.

A generalization?
(i) Sentential Subject Constraint:
All sentential subjects are islands.
(ii) Post-Sentential Subject Extraction Constraint:
The domain to the right of a sentential subject is an island.
$\rightarrow$
(iii) Most general constraint:
All sentences with sentential subjects are islands.

Problem:
Sentential subjects themselves can be moved.

(83) *Movement of sentential subjects*:
That John would be late, Mary didn't think was very likely.

2.14 Conclusion

The problem with most of the constraints discussed so far is the lack of generality; these constraints often look construction-specific. Should syntactic constraints be permitted to mention specific categorial features, or specific selectional features? Ideally, the answer is no. Still, some of the constraints are not subject to this critique. Most notably, this holds for the the A-over-A principle (in particular, its F-over-F revision) and for the Superiority

Condition. It therefore does not come as a surprise that the combination of these two conditions is widely considered valid nowadays. (The combined constraint is known as the *Minimal Link Condition*; more on this constraint is to come later.)

(84) *F-over-F Principle*^{d}:
In a structure $\alpha_{[*F*]}$... [$_{\beta_{[F]}}$ ... [$_{\gamma_{[F]}}$ ...] ...] ..., movement to [*F*] can only affect the category bearing the [F] feature that is closer to [*F*].

(85) *Superiority Condition*^{d} (Chomsky (1973)):
In a structure $\alpha_{[*F*]}$... [... $\beta_{[F]}$ ... [... $\gamma_{[F]}$ ...] ...] ..., movement to [*F*] can only affect the category bearing the [F] feature that is closer to [*F*].

Note:
The constraints discussed here (in [2]) are all local (derivational or representational). Is it possible to reformulate constraints like, e.g, the F-over-F Principle and the Superiority Condition as, e.g., transderivational constraints? Indeed, there is a straightforward reformulation, even though it is not fully equivalent.

(86) *Shortest Paths Condition*^{td} (Chomsky (1993)):
Minimize the length of movement paths.
(Given the set of derivations RS that are based on the same LA, choose the derivation in RS in which movement paths have minimal length.)

(87) *Movment path* (informal):
A movement path is the set of nodes that are crossed by movement operation. A movement path α is shorter than a movement path β if α has fewer nodes than β.

2.15 Exercises

Exercise 3:
Consider the following examples. They are all ungrammatical because they violate some constraint. Which example violates which constraint(s)?

(88) a. *What$_1$ did Bill buy potatoes and t$_1$?
b. *How$_1$ do you believe the stories [$_{CP}$ that John fixed your car t$_1$] ?
c. *The proof that the claim t$_1$ was made by the Greeks was given in 1492 [$_{CP_1}$ that the world was round]
d. *[$_{DP_1}$ Which rock star] were admirers of t$_1$ arrested ?

Exercise 4:
(89) looks like a violation of the Wh-Island Condition. Do the derivational and representational versions of the constraint in (63) and (64) on page 41 make identical predictions? Is there another constraint that also excludes (89)?

(89) *[$_{DP_1}$ Who] do you wonder [$_{CP}$ [$_{DP_2}$ which picture of t$_1$] John likes t$_2$] ?

Exercise 5:
The following grammatical sentences from French, German, and English are all potentially problematic for the system of constraints developed so far because they all appear to violate some constraint. Which constraints are violated by these examples, and why are they violated?

(90) a. Combien as-tu lu de livres ?
how many have you read of books
b. Was hat gelesen zu haben den Fritz geärgert ?
what$_{acc}$ has read to have the Fritz$_{acc}$ annoyed
c. Whose books did which students read ?
d. Was hat sie wem zu lesen empfohlen ?
what$_{acc}$ has she$_{nom}$ whom$_{dat}$ to read recommended
e.(?)This is a man to whom liberty we could never grant

Exercise 6:
Consider the following two sentences. Both are completely ungrammatical. As we have seen, (91-a) can be excluded by the Complex NP Constraint. What about (91-b), where DP$_2$ has been topicalized? Discuss the derivational and representational versions of the Complex NP Constraint. Is there another constraint that (91-b) violates?

(91) a. *I wonder $[_{CP}$ $[_{DP_1}$ which book] John met $[_{DP_2}$ a child who read t_1]

b. *$[_{DP_2}$ A child who read t_1], I wonder $[_{CP}$ $[_{DP_1}$ which book] John met t_2]

Exercise 7:

All *wh*-phrases must move to a clause-initial position in multiple *wh*-questions in Bulgarian (see Rudin (1988), Richards (1997), Bošković (2002)). Let us assume that all instances of such multiple *wh*-movement target $\text{SpecC}_{[+wh]}$. (This would seem to imply that $\text{C}_{[+wh]}$ can have more than one [*Q*] feature in Bulgarian.) Interestingly, the order of [Q]-marked DP arguments in $\text{SpecC}_{[\ wh]}$ positions must be identical to the base order of the DP arguments within VP; see (1-a) vs. (1-b). In simple *wh*-questions, the VP-internal order can be reversed by *wh*-movement; see (1-c) (where V has undergone LI-movement to C, which is irrelevant in the present context).

(1) *Multiple and simple wh-movement in Bulgarian*:

a. $[_{\text{CP}}$ Koj$_1$ $[_{\text{C}'}$ kogo$_2$ $[_{\text{C}'}$ $\text{C}_{[+wh]}$ $[_{\text{TP}}$ t'_1 T $[_{\text{VP}}$ t_1 običa t_2]]]]] ?
   who$_{nom}$ whom$_{acc}$ loves

b. *$[_{\text{CP}}$ Kogo$_2$ $[_{\text{C}'}$ koj$_1$ $[_{\text{C}'}$ $\text{C}_{[+wh]}$ $[_{\text{TP}}$ t'_1 T $[_{\text{VP}}$ t_1 običa t_2]]]]] ?
   whom$_{acc}$ who$_{nom}$ loves

c. $[_{\text{CP}}$ Kakvo$_2$ $[_{\text{C}'}$ $[_{\text{C}_{[+wh]}}$ pravi$_3$] $[_{\text{TP}}$ Ivan$_1$ T $[_{\text{VP}}$ t_1 t_3 t_2] ?
   what does Ivan

Questions:

(i) The phenomenon in (1-ab) is reminiscent of Superiority Condition effects in English. Does it follow from the Superiority Condition?

(ii) One might account for the difference between (1-a) and (1-b) by the constraint in (2). What kind of constraint is this (Con^d, Con^r, Con^g, Con^{td}, or Con^{tl})?

(iii) Would (2) also account for Superiority Condition effects in English?

(iv) Would (2) be compatible with (1-c)?

(v) Try to reformulate (2) without mentioning the levels "D-structure" and "S-structure", by exclusively referring to syntactic categories.

(vi) Why is (2) not really a "good" constraint?

(2) *[Q]-Isomorphism*:

If $\alpha_{[Q]}$ c-commands $\beta_{[Q]}$ at D-structure, $\alpha_{[Q]}$ also c-commands $\beta_{[Q]}$ at S-structure.

Exercise 8:

Languages like Italian exhibit so-called *pro-drop* constructions: A subject pronoun that is interpreted as a topic cannot be overtly realized. Suppose that there is a non-overt pronominal empty category *pro* in these contexts. Furthermore, a subject pronoun that is not a topic must be overtly realized, and cannot be *pro*. Account for this generalization by invoking two constraints; one of them should be transderivational/translocal.

(3) *Pro-drop in Italian*:

a. [$_{\text{TP}}$ pro$_{[top]}$ Ha cantato]
b. *[$_{\text{TP}}$ Lui$_{[top]}$ ha cantato]

Chapter 3

Constraints in the Principles-and-Parameters Approach

3.1 Introduction

Note:
The Principles-and-Parameters (aka *Government and Binding*) approach was first developed in Chomsky (1980; 1981); it was then refined throughout the eighties. A guiding idea was that the constraints (or "principles") should be as general as possible, and that they may contain open parameters which are fixed differently in different languages. Many of the constraints rely on the notions of government and binding.

A different approach to phrase structure:
The approach to phrase structure and derivations adopted in the Principles-and-Parameters approach is slightly different from the one presupposed so far:

(1) *Three levels of the syntactic component of a grammar in the Principles-and-Parameters approach*:
   a. *D-structure*:
      All pure Merge operations have applied; no Move operation has applied.
   b. *S-structure*:
      All overt Move operations have applied.
   c. *Logical Form*:
      All covert Move operations have applied.

Note:
Of these three *levels of representation*, S-structure is motivated indepen-

dently – it represents the accessible output form of a sentence. D-structure and Logical Form (LF) are theoretical constructs that are mainly motivated by the fact that constraints can refer to them (potentially exclusively so).

Note on (1-a):
D-structure is not to be confused with the lexical array: The latter contains to hierarchically organized structures, the former does.

Note on (1-b):
S-structure is the level of representation at which we have so far assumed representational constraints to hold. S-structure movement is phonologically visible.
(There are exceptions: Move may apply to certain empty categories – not to traces, but to an empty pronominal subject PRO in *control* infinitives (cf. *John tries* PRO_1 *to be elected* t_1, or to so-called *null operators* OP in, e.g., certain relative clauses (cf. *the man* OP_1 *I was talking to* t_1).

Note on (1-c):
Logical Form is an abstract level of representation that is supposed to act as the interface to the semantic representation. (Sometimes, it is in fact regarded as the semantic representation itself; see Heim & Kratzer (1998). In that case, LF has sometimes been referred to as *Transparent Logical Form.*) Movement of items in the LF-component is necessarily phonologically invisible.

(2) *Levels of grammar in the Principles-and-Parameters approach*:
Lexicon (plus Morphology) → D-structure → S-structure → Phonological Form, Logical Form

(3) *Constraints in the Principles-and-Parameters approach*:
   a. Local derivational constraints
      (Some constraints are of this type.)
   b. Local representational constraints
      (Most constraints are of this type.)
   c. Global constraints
      (A few constraints are of this type.)
   d. Transderivational/translocal constraints
      (The role of these constraints is minimal, but they exist nevertheless.)

3.2 Successive-Cyclic Movement

A question:
Movement operations like *wh*-movement appear to be unbounded in principle (as long as no constraints on Move are violated). Does such *long-distance movement* operate in one step, or does it result from the combination of smaller steps, i.e., operate *successive-cyclically*?

Assumption:
Long-distance movement applies successive-cyclically. Each intermediate SpecC position of a $C_{[-wh]}$ is targeted by Move on the way to the ultimate SpecC position of the $C_{[+wh]}$ node. Movement from a position created by Move (rather than by pure Merge) leaves an *intermediate* trace (t'_1, t''_2, etc.)

(4) *Successive-cyclic movement*:
  a. How_1 do you think [$_{CP}$ t'_1 [$_C$ (that)] Mary solved the problem t_1] ?
  b. [$_{DP_1}$ Which book] do you think [$_{CP}$ t'_1 [$_C$ (that)] John read t_1] ?
  c. [$_{DP_1}$ What] does John think [$_{CP}$ t''_1 [$_C$ (that)] Mary said [$_{CP}$ t'_1 [$_C$ (that)] Bill likes t_1]] ?

Problem:
If Move is a special case of Merge, and (except for modification operations) Merge is possible only if it deletes a selectional [*F*] feature (a [*Q*] feature in the case at hand), then how does movement to intermediate SpecC positions in (4) come about?

Three solutions:
(i) *Feature-based approach*:
There is in fact a selectional feature on the intermediate C nodes.
(ii) *Violability*:
The Economy Constraint on Merge (Move) is violable in favour of certain other constraints that force (certain) Move operations to be successive-cylcic.
(iii) *Form Chain*:
There is no selectional feature on the intermediate C nodes. Move operates in one step after all, targetting $C_{[+wh],[*Q*]}$ directly. But Move is a more complex operation (called 'Form Chain'): It inserts traces in all intermediate SpecC positions after re-merging an XP in its target position.

Note:

The Form Chain approach is problematic for conceptual reasons (it violates the Strict Cycle Condition that will be discussed below). The approach in terms of violability has a number of non-trivial further consequences since it presupposes that constraints can be violable in principle (which we have not assumed so far). Hence, for the time being, the feature-based approach will be adopted. This approach is also arguably the predominant one in recent work based on minimalist assumptions.

Assumption (Chomsky (2000; 2001)):
C can be assigned a [*F*] ([*Q*], [*top*], [*rel*]) feature during the derivation (that triggers movement to SpecC) only if this has an effect on output.

Note:
This raises a question with respect to the Inclusiveness Condition. Furthermore, the problem arises of how one can determine at a given stage in the derivation whether assigning a feature like [*Q*] will eventually be justified; this seems to require what is known as *look-ahead*. Alternatively, one might assume that in order to decide at a given stage of the derivation whether [*Q*] is to be assigned to a given C, one can look into the lexical array: Simplifying a bit, [*Q*] should better be assigned to $C_{[-wh]}$ if there is a $C_{[+wh],[*Q*]}$ left in the lexical array that needs a *wh*-phrase to delete its selectional [*Q*] feature (and no other *wh*-item is left in the lexical array).

An alternative:
Suppose that $C_{[-wh]}$ can optionally bear a feature like [*Q*] in the lexical array, without qualification. Then, the problem might arise to exclude ill-formed sentences like (5-b), depending on the features of the root C (sentences of this type are possible as long-distance questions in certain languages, though, like Iraqi Arabic and Ancash Quechua). Depending on whether root C is [+wh], [*Q*] or [–wh], (5-b) can be straightforwardly excluded ([*Q*] on C must be deleted under identity with a [Q] feature on a *wh*-phrase, which it is not in (5-b)), or requires additional assumptions (a [–wh] does not require a [*Q*] feature; hence, the illformedness of (5-b) must be accounted for by invoking additional assumptions).

(5) *A potential problem with [*F*] features on [–wh] C nodes*:
   a. *Partial LA*:
      (ia) Ø: { [C], [root], [+wh], [+fin], [*T*,*+fin*] > [*Q*] }
      (ib) Ø: { [C], [root], [–wh], [+fin], [*T*,*+fin*] }

(ii) Ø: { [C], [–wh], [+fin], [*T*,*+fin*] > [*Q*] }
(iii) which: { [D], [3pers,–pl,–fem,–masc], [acc], [Q], [*N*] }
(iv) book: { [N], [3pers,–pl,–fem,–masc], [acc] }

b. *[$_{CP}$ [$_{C}$ Ø] Mary thinks [$_{CP}$ [$_{DP_1}$ which book] [$_{C}$ Ø] John read t_1]]

Conclusion:
For present purposes, we simply assume that Move can operate successive-cyclically, via intermediate SpecC positions, and that such movement can be reconciled with the idea that Move takes place only if it deletes a selectional feature.

Note:
Whereas it is difficult to find evidence for selectional features on intermediate C nodes, there is evidence for successive-cyclic movement via SpecC as such. Some languages show morphological reflexes of successive-cyclicity in SpecC on either the C node (e.g., Modern Irish) or the adjacent SpecT position (e.g., Ewe). Some languages exhibit verb-second phenomena (V/T-to-C movement) in a CP exactly in those circumstances where (a certain type of *wh*-) movement has taken place from that CP (e.g., Spanish). Some languages have obligatory CP extraposition (which is otherwise optional) when movement takes place from that CP (German). Some languages (like Iraqi Arabic, Hungarian, Ancash Quechua, and German) even permit a stranding of the *wh*-phrase in an intermediate SpecC position (*partial wh-movement*).

(6) *Partial wh-movement in German*:

a. ?Wen$_1$ denkt Maria [$_{CP}$ t'_1 [$_{C}$ dass] Fritz t_1 mag] ?
whom thinks Maria that Fritz likes

b. Was denkt Maria [$_{CP}$ wen$_1$ [$_{C}$ Ø Fritz t_1 mag] ?
what thinks Maria whom Fritz likes

Note:
Thus far, the only goal was to derive that successive-cyclic movement via SpecC is possible. It remains to be shown that such movement is also necessary in long-distance movement constructions.

3.3 Derivational Constraints: The Subjacency Condition

3.3.1 The Constraint

(7) *Subjacency Condition*[d] (Chomsky (1977)):

a. In a structure α ... [$_\beta$ ... [$_\gamma$ ... δ ...] ...] ..., movement of δ to α cannot apply if β and γ are bounding nodes.
b. DP and TP are bounding nodes.

Note:
The Subjacency Condition is much more general than most of the constraints discussed so far. True, it does mention categorial features ([D] and [T]), but it turns out to account for many effects that separate constraints were so far needed for.

A first consequence of the Subjacency Condition:
It now follows that successive-cyclic movement is the only way to leave a CP: Otherwise, two TPs will be crossed in the course of a single Move operation. Thus, the SpecC position acts as an *escape hatch*. In the final output representation, the *wh*-phrase is separated from its initial trace by two bounding nodes, but given the derivational formulation of the Subjacency Condition and the availability of successive-cyclic movement, this is unproblematic.

(8) *The Subjacency Condition and successive-cyclic movement*:
[$_{DP_1}$ Which book] do [$_{TP_2}$ you think [$_{CP}$ t_1' [$_C$ (that)] [$_{TP_4}$ John read t_1]]] ?

Further consequences of the Subjacency Condition:
The Subjacency Condition accounts for Complex NP Constraint effects, Wh-Island Condition effects, Left Branch Condition effects, some Coordinate Structure Constraint effects, and Sentential Subject Constraint/Subject Condition effects (under certain assumptions).

3.3.2 Complex NP Constraint Effects

(9) *A consequence of the Complex NP Constraint, argument clauses*:
a.??[$_{DP_1}$ Which book] did [$_{TP_3}$ John hear [$_{DP_2}$ a rumour [$_{CP}$ t_1' that you had read t_1]]] ?
b. *[$_{PP_1}$ How] did [$_{TP_3}$ John hear [$_{DP_2}$ a rumour [$_{CP}$ t_1' that you had fixed the car t_1]]] ?
c.?*The hat [$_{DP_1}$ which] [$_{TP_3}$ I believed [$_{DP_2}$ the claim [$_{CP}$ t_1' that Otto was wearing t_1]]] is red

Note:
Movement in (9-abc) crosses TP_3 and DP_2 in the second step; TP_3 and DP_2 are bounding nodes. Hence, illformedness results. In contrast, the first

movement step crosses only one bound node – the embedded TP –, and therefore respects the Subjacency Condition.

(10) *A consequence of the Complex NP Constraint, relative clauses*:
a. *[$_{DP_1}$ Which book] did [$_{TP_3}$ John meet [$_{DP_2}$ a child [$_{CP}$ who [$_{TP_4}$ read t_1]]]] ?
b. *[$_{DP_1}$ Who] does [$_{TP_3}$ Mary know [$_{DP_2}$ a girl [$_{CP}$ who [$_{TP_4}$ is jealous of t_1]]]] ?

Note:
Movement in (10-ab) crosses the two bounding nodes TP_3 and DP_2. In addition, this time the embedded bounding node TP_4 is also crossed. The reason is this: First, there is some other category in the SpecC position of the relative clause already, viz., the relative pronoun. Second, it seems to be a fact about many languages (English and German among them) that C can have only one selectional feature that triggers a Move operation targetting SpecC. In other words: C can only have one specifier. Under these assumptions, it follows that a relative pronoun in SpecC blocks the use of SpecC as an escape hatch for successive-cyclic movement form CP.

A side effect:
Movement from DP-internal relative clauses crosses one more bounding node than movement from DP-internal argument clauses. This is often taken to account for the fact that Complex NP Constraint violations are typically more severe with relative clauses than with argument clauses.

3.3.3 Wh-Island Condition Effects

(11) *A consequence of the Wh-Island Condition*:
a. *How$_1$ does [$_{TP_3}$ she know [$_{CP}$ [$_{DP_2}$ which car] [$_{TP_4}$ Mary fixed t_2 t_1]]] ?
b.?*[$_{DP_1}$ Which book] do [$_{TP_3}$ you wonder [$_{CP}$ [$_{PP_2}$ to whom] [$_{TP_4}$ John gave t_1 t_2]]] ?
c. ??Who$_1$ do [$_{TP_3}$ you wonder [$_{CP}$ why C [$_{TP_4}$ Mary loves t_1]]] ?

Note:
Wh-movement in (11-ab) crosses two bounding nodes (TP_3 and TP_4), in violation of the Subjacency Condition. As with the relative clause case of the Complex NP Constraint, the problem is that the escape hatch SpecC is blocked by something else.

A problem:
Why does *wh*-movement in (12) violate the Subjacency Condition? (Recall that it violates the Wh-Island Condition only because of the stipulation that *wh*-elements in SpecC *or C* block movement from a CP.)

(12) *Wh-Islands created by C elements*:
*How_1 do you wonder [$_{CP}$ whether Mary solved the problem t_1] ?

A solution:
It is clear that LIs like *whether* and *if* do not need to have [*Q*] features, unlike an empty (non-root) $C_{[+wh]}$, which must have a [*Q*] feature in English-type languages with *wh*-movement in questions. Suppose that this assumption is strengthened: *whether* and *if*, as a lexical property, cannot have a [*Q*] feature, in contrast to [–wh] complementizers (*that* and Ø) which can have [*Q*] features. Then, *wh*-movement in (12) will have to proceed in one step, without an intermediate landing site in the embedded SpecC position, and a violation of the Subjacency Condition is ensured.

3.3.4 Left Branch Condition Effects

(13) *A consequence of the Left Branch Condition*:
a. *[$_{DP_1}$ Whose] did [$_{TP_3}$ you meet [$_{DP_2}$ t_1 sister]] ?
b. *[$_{AP_1}$ Neue] hat [$_{TP_3}$ Hans [$_{DP_2}$ D [$_{NP}$ t_1 Bücher]] gekauft]
new has Hans books bought

Note:
Movement crosses TP_3 and DP_2 in (13-ab). Hence, a violation of the Subjacency Condition results.

3.3.5 Coordinate Structure Constraint Effects

(14) *Consequences of the Coordinate Structure Constraint*:
a. *[$_{DP_1}$ Who] does John like [$_{DP_3}$ t_1 and [$_{DP_2}$ Bill]] ?
b. *[$_{DP_1}$ Who] is John [$_{AP}$ proud of t_1] and [$_{AP}$ tired of [$_{DP_2}$ his mother]] ?

Note:
(14-b) does not follow from the Subjacency Condition. (14-a) does so only if we assume (instead of postulating a "coordination phrase") that the two DP conjuncts are dominated by a DP again.

3.3.6 Sentential Subject Constraint/Subject Condition Effects

(15) *A consequence of the Subject Condition*:

a. *[$_{DP_2}$ Who(m)] has [$_{TP_3}$ [$_{DP_1}$ a comment about t_2] annoyed you] ?
b. *[$_{PP_3}$ About whom] has [$_{TP_3}$ [$_{DP_1}$ a comment t_3] annoyed you] ?

Note:
Movement crosses two bounding nodes, TP_3 and DP_1. Hence, a Subjacency Condition violation arises in both (15-a) and (15-b).

Problem:
Like the A-over-A Principle, the Subjacency Condition fails to distinguish between argument DPs that are external arguments merged in specifier positions (subjects) and argument DPs that are merged in complement positions (objects). This may be a desirable result for left branches of DPs, but it is less desirable for material that is merged to the right of N. Hence, it seems that the constraint is too strong; it rules out examples like (16-ab). However, it also rules out (16-cd), which is a welcome result – (16-c) has another type of embedding predicate, (16-d) has replaced the indefinite determiner of (16-abc) with a more specficic, definite determiner (a so-called *Specificity* effect).

(16) *A problem for the Subjacency Condition*:

a. [$_{DP_1}$ Which author] did [$_{TP_3}$ you read [$_{DP_2}$ a book about t_1]] ?
b. [$_{DP_1}$ Who$_1$] did [$_{TP_3}$ you see [$_{DP_2}$ a picture of t_1]] ?
c. *[$_{DP_1}$ Which author] did [$_{TP_3}$ you destroy [$_{DP_2}$ a book about t_1]] ?
d.?*[$_{DP_1}$ Which author] did [$_{TP_3}$ you read [$_{DP_2}$ the book about t_1]] ?

Sketch of a possible solution:
Suppose that certain types of verbs have a *reanalysis* property that in effect can break up the DP structure of its internal argument. Technically, we can assume that, e.g., *read* can have an additional [*P*] feature that does not correspond to a slot in the argument structure, and that can only be deleted by (string-vacuous) rightward movement of PP_1 (*about which author*) in (16-a) to a right-peripheral specifier in VP. Such PP movement from DP_2 crosses only one bounding node since the landing site is still below TP. In the next step, DP_1 would move from the extraposed PP_1 to SpecC, again crossing only one bounding node. Needless to say, such an analysis raises many further problems (e.g., moved items typically block further extraction, see below).

Note:
The Subjacency Condition may also account for Sentential Subject Constraint effects if we make some further assumptions. A first assumption might be that subject clauses are always embedded by DPs with empty D heads, as in (17-a). Then, it follows that movement from the subject CP_4 will have to cross two bounding nodes (TP_2 and DP_3), even if it proceeds successive-cyclically, via the intermediate SpecC position of CP_4. Alternatively, one might assume that whereas there is no empty determiner embedding subject clauses, a LI C that is merged in a specifier position cannot bear the [*Q*] that is otherwise optional (and needed to trigger successive-cylic movement, by assumption). If so, Sentential Subject Constraint effects will essentially be derivable in the same way as Wh-Island Condition effects: In (17-b), *wh*-movement will have to cross two bounding nodes (TP_2 and TP_5) in one step.

(17) *A consequence of the Sentential Subject Constraint*:
   a. *[$_{DP_1}$ Who] did [$_{TP_2}$ [$_{DP_3}$ [$_D$ Ø] [$_{CP_4}$ t_1' that [$_{TP_5}$ Mary was going out with t_1]]] bother you] ?
   b. *[$_{DP_1}$ Who] did [$_{TP_2}$ [$_{CP_4}$ that [$_{TP_5}$ Mary was going out with t_1]] bother you] ?

3.3.7 Parametrization

Note:
The Subjacency Condition is the classic example of a parametrized constraint. The idea is that languages may differ with respect to what counts as a bounding node, and what does not. Here is Rizzi's (1982) famous proposal for Italian:

(18) *Parametrization of bounding nodes*:
   a. English: DP, TP
   b. Italian: DP, CP

Evidence:
Italian seems to freely violate the Wh-Island Condition (see (19-a)), but it respects the Complex NP Constraint (see (19-b)). Wh-Island Condition effects can be derived from the Subjacency Condition in English because they involve a crossing of two TP bounding nodes in one movement step; if CP replaces TP as a bounding node in Italian, Wh-Island Condition violations are expected to disappear. Complex NP Constraint effects can be derived from the Subjacency Condition in English because they involve a crossing of

a DP and a TP bounding node; and if CP replaces TP as the second bounding node in Italian, these kinds of effects are still predicted. Furthermore, movement steps that cross *two wh*-islands are again correctly predicted to be impossible, even in Italian; see (19-c).

(19) *The Wh-Island Condition and the Complex NP Constraint in Italian*:

a. Tuo fratello [$_{CP_3}$ [$_{PP_1}$ a cui] mi domando [$_{CP_4}$ [$_{DP_2}$
your brother to whom myself I ask
che storie] abbiano raccontato t_2 t_1] era molto
which stories they have told was very
preoccupato
worried

b. *Tuo fratello [$_{CP_3}$ [$_{PP_1}$ a cui] temo [$_{DP_4}$ la possibilitá [$_{CP_5}$
your brother to whom I fear the possibility
t'_1 che abbiano raccontato tutto t_1]]] ...
that they have told everything

c. *Francesca [$_{CP_3}$ [$_{DP_1}$ che] non immagino [$_{CP_4}$ [$_{DP_2}$ quanta
Francesca who not I imagine how many
gente] t_2 sappia [$_{CP_5}$ [$_{PP_6}$ dove] hanno mandato t_1 t_6]]] ...
people know where they have sent

A cautionary note:
It is not really clear whether a parametrization of bounding nodes is the correct approach. First, the Italian examples above involve relativization; however, Italian *wh*-movement constructions that closely resemble constructions that are typically chosen to illustrate Wh-Island Condition effects in English are also fairly ill formed; the pertinent examples can already be found in Rizzi's (1982) original work. Second, some of the English examples that involve a Wh-Island Condition violation are often not judged to be that ill formed after all; see Grimshaw (1986), Chomsky (1986).

(20) *Wh-movement respects the Wh-Island Condition in Italian*:

a. *[$_{DP_1}$ Chi] ti domandi [$_{CP_3}$ [$_{DP_2}$ chi] t_2 ha incontrato
who yourself you ask who has met
t_1] ?

b.??[$_{DP_1}$ Chi] non sai [$_{CP_3}$ [$_{DP_2}$ che cosa] t_1 ha fatto t_2] ?
who not you know what has done

(21) *Wh-movement may violate the Wh-Island Condition in English*:

a. ?[$_{DP_1}$ Which book] did the students forget [$_{CP_3}$ who$_2$ t_2 wrote t_1] ?

b. ?[$_{DP_1}$ Which car] did John tell you [$_{CP_3}$ how$_2$ to fix t_1 t_2] ?

Conclusion:
The concept of parametrization of bounding nodes is often considered dubious nowadays. More generally:

(22) *Assumptions about parametrization*:
Languages differ (a) the feature structures of their LIs (including functional categories), and (b) their linear precedence statements, but not in (c) the make-up of the fundamental operations Merge and Move, or (d) the constraints.

3.3.8 A Gap in the Argument So Far?

Note:
The discussion so far presupposes a certain derivational order in Wh-Island Condition constructions: First, a *wh*-phrases XP$_2$ moves to the embedded SpecC$_{[+wh]}$, then, another *wh*-phrase XP$_1$ moves across it to the higher SpecC$_{[+wh]}$. But what about the reverse application of movement operations?

(23) *An alternative derivation for Wh-Island Condition constructions*:

a. ...
b. [$_{CP_4}$ C$_{[+wh]}$ [$_{TP}$ John gave [$_{DP_1}$ which book] [$_{PP_2}$ to whom]]]
c. [$_{CP_4}$ [$_{DP_1}$ which book] C$_{[+wh]}$ [$_{TP}$ John gave t_1 [$_{PP_2}$ to whom]]]
d. ...
e. [$_{CP_5}$ C$_{[+wh]}$ [$_{TP}$ you wonder [$_{CP_4}$ [$_{DP_1}$ which book] C$_{[+wh]}$ [$_{TP}$ John gave t_1 [$_{PP_2}$ to whom]]]]]
f. [$_{CP_5}$ [$_{DP_1}$ which book] C$_{[+wh]}$ [$_{TP}$ you wonder [$_{CP_4}$ t_1' C$_{[+wh]}$ [$_{TP}$ John gave t_1 [$_{PP_2}$ to whom]]]]]
g. [$_{CP_5}$ [$_{DP_1}$ which book] C$_{[+wh]}$ [$_{TP}$ you wonder [$_{CP_4}$ [$_{PP_2}$ to whom] C$_{[+wh]}$ [$_{TP}$ John gave t_1 t_2]]]]

A first solution:
The standard solution to this problem is that the last movement operation in (23) is *counter-cyclic*: It violates the *Strict Cycle Condition*:

(24) *Strict Cycle Condition*[d] (Chomsky (1973)):
No operation can apply to a domain dominated by a cyclic node α in such a way as to affect solely a proper subdomain of α dominated by a node β which is also a cyclic node.

Note:
There is disagreement as to what counts as a cyclic node. The strongest hypothesis is that every XP is a cyclic node.

(25) *Cyclic node*:
Every XP is a cyclic node.

Conclusion:
The last operation in the derivation in (23) violates the Strict Cycle Condition: Here, *wh*-movement of PP_2 affects only the embedded CP_4, which is dominated by several other cyclic nodes (matrix VP, matrix TP, matrix CP_5).

A second solution:
Suppose that the Subjacency Condition is reformulated as a representational constraint:

(26) *Subjacency Condition*^{r} (Freidin (1978; 1992)):
a. $*\ldots\ \alpha_1\ \ldots\ [_\beta\ \ldots\ [_\gamma\ \ldots\ t_1\ \ldots\ ]\ \ldots\ ]\ \ldots$, where β and γ are bounding nodes.
b. DP and TP are bounding nodes.

Conclusion:
It is now immaterial how Wh-Island Condition constructions are derived: The Subjacency Condition successfully rules out the final output representation. All the remaining evidence in favour of the Subjacency Condition that was discussed so far can still be accounted for under the representational reformulation.

A third solution:
Suppose that we maintain the derivational formulation of the Subjacency Condition. Counter-cyclic derivations of Wh-Island Condition constructions may then still be excluded without invoking the Strict Cycle Condition, given the assumptions about movement adopted above. Here is why:
(i) Because of the Economy Constraint on Merge and the definition of Move in terms of Merge, there can be no movement without deletion of a selectional feature.
(ii) C can only have one selectional [*Q*] feature in English-type languages (otherwise, multiple *wh*-movement would be predicted to occur, as in Bulgarian).
(iii) Once DP_1 has moved to SpecC of CP_4 in (23), no other XP (incuding

PP_2) will be able to move to that position in the remainder of the derivation, because of the Economy Constraint on Merge.

Note:
Even though it may not be needed for an account of Wh-Island phenomena, the Strict Cycle Condition is a fundamental constraint in derivational approaches to syntax. It is needed in many other domains.

3.3.9 Is the Subjacency Condition a Derivational or a Representational Constraint?

Note:
Chomsky (1981) presents a theory-internal argument in favour of a derivational formulation of the Subjacency Condition. It is based on *wh*-movement from *exceptional Case-marking* (ECM) constructions. Assumptions:
(i) In ECM constructions, the selectional [*acc*] Case feature of a matrix V can exceptionally be deleted under identity with an [acc] Case feature on a DP that V is not merged with; rather, the DP providing the matching [acc] feature is the specifier of an infinitival TP complement of V.
(ii) The Case feature of DP in SpecT of an infinitive cannot be deleted under identity with a selectional Case feature within the infinitive.
(iii) DP must move to the embedded SpecT position even though infinitival T cannot possibly have a [*nom*] feature. (This might argue for [*D*] as the feature triggering movement to SpecT after all.)

(27) *ECM constructions*:
I believe $[_{TP}$ $[_{DP_1}$ him] to be t_1 in love with Mary]

Problem:
If bare TP embedding is the correct analysis for (27), (28) is wrongly predicted to incur a violation of the Subjacency Condition: A SpecC escape hatch is missing.

(28) *A violation of the Subjacency Condition*:
Who_2 do $[_{TP}$ you believe $[_{TP}$ him_1 to be in love with t_2]] ?

Chomsky's solution:
ECM constructions initially involve Merge of V and an infinitval CP. As a lexical property, ECM predicates can then delete the CP shell later in the derivation; and they must do so in order to ensure [*acc*]/[acc] feature deletion on V and DP_1. However, *wh*-movement must take place prior to CP deletion, in order to satisfy the Subjacency Condition. This, in turn,

implies that the Subjacency Condition must be a derivational constraint; a representational Subjacency Condition can only check the ultimate output representation, in which CP has been deleted, and the *wh*-phrase is separated from its trace t_2 by two TP bounding nodes.

(29) *A relevant part of the derivation*:

a. $[_{CP}$ C $[_{TP}$ him to be in love with who_1]]]
b. $[_{CP}$ who_1 C $[_{TP}$ him to be in love with t_1]]] ...
c. who_1 do $[_{TP}$ you believe $[_{CP}$ t'_1 C $[_{TP}$ him to be in love with t_1]]]
d. who_1 do $[_{TP}$ you believe $[_{TP}$ him to be in love with t_1]] ?

3.3.10 A Note on LF Movement and the Subjacency Condition

Note:
Recall that the Principles-and-Parameters approach envisages an abstract level of Logical Form that is created on the basis of S-structure via so-called LF movement. LF movement has been suggested for the following types of categories, among others:

(30) *Items that undergo LF movement in the Principles-and-Parameters approach*:

a. *Wh*-phrases in multiple questions that are *in situ* at S-structure undergo movement to a specifier position of $C_{[+wh]}$ in the LF component.
b. Quantified XPs undergo so-called *quantifier raising* (QR) to a TP- or VP-specifier in the LF component.

Note:
The basic motivation behind postulating these abstract movement operations is semantic. We will not be concerned with the question of what triggers the movement operation (selectional [*F*] features that are somehow inert at S-structure being an obvious candidate), and what the exact landing site is. Furthermore, we can leave open whether there is or is not good reason to assume a level of Logical Form that is derived by syntactic movement in the first place. However, it seems clear that if LF exists, the Subjacency Condition can not be assumed to hold at this level (if it is formulated representationally), or to hold for movement operations that connect S-structure to LF (if it is formulated derivationally). Here is why:

(31) *Wh-in situ does not obey the Subjacency Condition*:

a. Who_1 t_1 remembers $[_{CP}$ why_2 we bought $what_3$ t_2] ?

b. Who$_1$ t$_1$ likes [$_{DP}$ D books [$_{CP}$ that critisize who$_2$]] ?
c. Who$_1$ t$_1$ thinks [$_{CP}$ that [$_{DP}$ pictures of who$_2$] are on sale] ?

Observation:
The same goes for other island phenomena.

(32) *Wh-in situ does not obey the Coordinate Structure Constraint*:
Who$_1$ t$_1$ saw John and who$_2$?

Observation:
The same goes for *wh*-in situ in a language like Chinese, which does not have [*Q*] features on C nodes marked [+wh] (see Huang (1982)).

(33) *Wh-in situ does not obey the Subjacency Condition*:
a. ni zui xihuan [$_{DP}$ shei mai de shu] ?
you most like who buy Comp book
'*Who$_1$ do you like the books that t$_1$ bought?'
b. [$_{CP}$ wo mai shenme] zui hao ?
I buy what most good
'*What$_1$ is that I buy t$_1$ best?'

Conclusion:
The argument for a syntactic derivation of a level of Logical Form is strengthened if it can be shown that LF-construction obeys constraints that are otherwise well motivated in syntax; it is weakened if it does not obey any of the well-established syntactic constraints. There is no general agreement with respect to this question so far.

3.4 Derivational Constraints: The Condition on Extraction Domains

Observation:
Items which do not enter the derivaiton via selectional Merge (modifiers, so-called *adjuncts*) are always islands. This can be formulated in a preliminary way as the Adjunct Condition:

(34) *Adjunct Condition*[d]:
Movement must not take place from an XP that has been merged without a deletion of selectional features.

The Adjunct Condition straightforwardly excludes Complex NP Constraint constructions in which a relative clause is crossed by movement. Furthermore:

(35) *A consequence of the Adjunct Condition*:

a. [$_{\text{DP}_1}$ Who] did you get jealous [$_{\text{CP}}$ because I talked to t_1] ?
b. [$_{\text{PP}_1}$ To whom] did they leave [$_{\text{CP}}$ before speaking t_1] ?
c. [$_{\text{DP}_1}$ Who] did they leave [$_{\text{CP}}$ before speaking to t_1] ?

Question:
Can (35-abc) also be excluded by the Subjacency Condition? The answer is yes if we can ensure that the adjunct CPs do not have a SpecC position that is available for successive-cyclic movement; otherwise (i.e., if the adjunct CPs have an available SpecC position) it is no.

Observation:
The Subject Condition and the Adjunct Condition can be unified as the Condition on Extraction Domains (CED). The basic insight was arguably first formulated by Cattell (1976). The notion CED is due to Huang (1982). Kayne (1984) employs a similar concept. Chomsky (1986) is the most comprehensive and careful study in this area; it centers around the notion of *barrier*. Cinque (1990) has useful simplifications. The following definition freely draws on all the concepts developed in these approaches.

(36) *Condition on Extraction Domain*:

a. Movement must not cross a barrier.
b. An XP is a barrier iff it is not a complement.

Note:
Conceptually, this is a step in the right direction because we move from an intrinsic definition to a contextual definition of locality domains: Whether some XP is a bounding node or not is simply listed; whether some XP is a barrier or not can be determined by looking at the syntactic context in which it occurs.

Consequence:
A barriers-based approach to locality in terms of the Condition on Extraction Domains can account for Subject Condition and Adjunct Condition effects. It also derives the relative clause case of the Complex NP Constraint. If argument clauses selected by N are in fact not merged in complement position (as suggested by Stowell (1981), Kiss (1986), among others), Complex NP Constraint phenomena can be explained in toto. A further constraint that can be dispensed with in favour of the Condition on Extraction Domains is the Freezing Principle. The reason is that movement can never end in a complement position.

(37) *Freezing Principle*[d] (based on Ross (1967), Wexler & Culicover (1980)):
Movement cannot take place from a moved XP.

Note:
Given that subject DPs are DPs that have been moved to SpecT, their opacity follows from both the Subject Condition and the Freezing Principle.

(38) *Consequences of the Freezing Principle*:

a. *Who$_1$ do you think [$_{CP}$ t$'_1$ that [$_{DP_2}$ pictures of t$_1$] were painted t$_2$] ?
b. *Who$_1$ do you think [$_{CP}$ t$'_1$ that [$_{DP_2}$ pictures of t$_1$] John would like t$_2$] ?
c. *Who$_1$ do you think [$_{CP}$ [$_{PP_2}$ to t$_1$] he will talk t$_2$] ?
d. *Who$_1$ don't you know [$_{CP}$ [$_{DP_2}$ which picture of t$_1$] Mary bought t$_2$] ?
e. *[$_{PP_1}$ Über Fritz]$_1$ glaube ich [$_{CP}$ [$_{DP_2}$ ein Buch t$_1$] hat Maria
about Fritz believe I a book has Maria
t$_2$ geschrieben]
written

An important distinction:
From a more general point of view, we can distinguish between two types of (local derivational or local representational) locality constraints – rigid locality constraints and relativized locality constraints (island constraints all belong to the first group).

(39) *Two types of locality constraints*:

a. *Rigid Locality*:
   (i) Complex NP Constraint
   (ii) Sentential Subject Constraint
   (iii) Subject Condition
   (iv) Coordinate Structure Constraint
   (v) Upward Boundedness Constraint
   (vi) Left Branch Condition
   (vii) Wh-Island Condition
   (viii) Clause Non-final Incomplete Constituent Constraint
   (ix) Post-Sentential Subject Extraction Constraint
   (x) Subjacency Condition
   (xi) Adjunct Condition
   (xii) Condition on Extraction Domains
b. *Relativized Locality*:

(i) A-over-A Principle
(ii) F-over-F Principle
(iii) Superiority Condition
(iv) Minimal Link Condition (= F-over-F Principle & Superiority Condition)
(v) Relativized Minimality (Rizzi (1990), has not yet been discussed)

Generalization:
Both types of constraints are needed, but it is far from clear which phenomena should be accounted for by which constraint type. Currently, there are two fundamental constraints that are widely adopted: The Condition on the Extraction Domains on the one hand, and the Minimal Link Condition (i.e., the combined F-over-F Principle/Superiority Condition) on the other.

Note:
There is an interesting correlation: The XPs that best tolerate movement out of them are also the ones that can be moved most easily themselves (from certain types of islands), viz.: complements. Ideally, this should be reflected in the theory.

3.5 Representational Constraints: The Empty Category Principle

Assumption:
The Empty Category Principle is a representational constraint that holds at LF.

(40) Empty Category Principle (ECP)^r:
Every trace must be marked $[+\gamma]$.

(41) γ-*Marking* (derivational):
A trace is marked $[+\gamma]$ iff it is properly governed.

(42) *Proper Government* (simplified):
A trace is properly governed if it is antecedent-governed or lexically governed.

(43) *Lexical Government* (simplified):
α lexically governs β iff

a. α is a LI belonging to a lexical category.
b. α and β are dominated by the same XPs.

(44) *Antecedent-Government* (simplified):
α antecedent-governs β iff

a. α and β are co-indexed.
b. α c-commands β.
c. There is no barrier between α and β.
d. There is no *wh*-phrase or complementizer in the C domain that intervenes between α and β.

Note:
(i) An item in SpecC cannot antecedent-govern a subject trace in SpecT across a lexical complementizer.
(ii) An item in an outer SpecC position cannot antecedent-govern a subject trace in SpecT across an item in an inner SpecC position.

(45) *The ECP accounts for that-trace effects*:

a. *Who$_1$ do you think [$_{CP}$ t_1'([$+\gamma$]) that [$_{TP}$ t_1([$-\gamma$]) left]] ?
b. Who$_1$ do you think [$_{CP}$ t_1'([$+\gamma$]) Ø [$_{TP}$ t_1([$+\gamma$]) left]] ?

Analysis:
The trace t_1 in (45-b) is antecedent-governed from SpecC; the trace t_1 in (45-a) is not. Since it is not lexically governed either, it cannot be assigned [$+\gamma$], and the ECP will be violated at LF.

General assumption (Lasnik & Saito (1984; 1992), Chomsky (1986)):
Intermediate traces of arguments can be deleted on the way to LF (intermediate traces of adjuncts cannot be deleted on the way to LF).

(46) *An Anti-that-trace effect*:
Who$_1$ do you think [$_{CP}$ t_1''([$+\gamma$]) that Mary said [$_{CP}$ t_1'([$-\gamma$]) Ø [$_{TP}$ t_1([$+\gamma$]) left]]] ?

Note:
There is no ECP violation in (46) because the intermediate argument trace t_1'([$-\gamma$]) can be deleted on the way to LF; but there is an ECP violation in (45-a) because the initial t_1([$-\gamma$]) cannot be deleted on the way to LF. This presupposes that the ECP is a representational constraint applying to LF representations, not to S-structure representations or derivations. If the ECP held at S-structure or in the derivation, we would expect the [$-\gamma$]-marked intermediate trace t_1' in (46) to induce a fatal ECP effect that would render the sentence ungrammatical.

Note:
Another application of the ECP concerns data that have so far been accounted for by the Superiority Condition:

(47) *Superiority Condition effects*:
  a. Who$_1$ t$_1$ saw what$_2$?
  b. *What$_2$ did who$_1$ see t$_2$?
  c. I wonder [$_{CP}$ who$_1$ t$_1$ bought what$_2$]
  d. *I wonder [$_{CP}$ what$_2$ who$_1$ bought t$_2$]

Analysis:
If all *wh*-in situ XPs must move to a SpecC$_{[+wh]}$ position in the LF component, and if they must occupy an *outer* specifier of C if some other *wh*-phrase has already moved to a specifier of C in the syntax, a subject trace created by LF *wh*-movement will not be marked [$+\gamma$]. Not being deletable, it will therefore incur a violation of the ECP.

Problem (Hendrick & Rochemont (1982), Pesetsky (1982)):
An ECP approach does not cover all Superiority Condition effects.

(48) *Superiority Condition effects that are not reducible to the ECP*:
  a. Whom$_1$ did John persuade t$_1$ [$_{CP}$ to visit whom$_2$] ?
  b. *Whom$_2$ did John persuade whom$_1$ [$_{CP}$ to visit t$_2$] ?

3.6 Global Constraints: The Projection Principle

Note:
The Projection Principle (Chomsky (1981)) applies to pairs of levels of representation; hence, it qualifies as a global constraint.

(49) *Projection Principle*[g]:

  a. If A selects B as a lexical property, then A selects B in C at level L$_i$.
  b. If A selects B in C at level L$_i$, then A selects B in C at level L$_j$.

(50) *A consequence of the Projection Principle*:
  a. What$_1$ did John [$_{VP}$ see t$_1$]?
  b. *What$_1$ did John [$_{VP}$ see]?

Note:
To find out whether the Projection Principle is violated, it does not suffice

to simply look at a level of representation, or at a step in the derivation – to show that (50-b) is an impossible S-structure representation, we have to know that there is an object DP within VP at D-structure.

3.7 Transderivational/Translocal Constraints: Avoid Pronoun

Note:
Chomsky (1981) proposes a non-local, non-global Avoid Pronoun principle as a genuinely grammatical (i.e., non-pragmatic) constraint. The empirical evidence comes from English gerunds. A background assumption is that all entries in the argument structure (Θ-grid) of a predicate must be represented as arguments in the syntax. In those cases where no external argument DP is visible, there is a non-overt argument PRO.

(51) *PRO in English gerunds*:
   a. John_1 would much prefer [PRO_1 going to the movie]
   b. *John_1 would much prefer [$\text{PRO}_{2/arb}$ going to the movie]

(52) *Constraint on Control*^{r} (Manzini (1983)):
   If PRO is minimally dominated by a declarative clausal complement α, then it must be controlled by an antecedent within the minimal CP that dominates α.

(53) *Pronouns in English gerunds*:
   a. *John_1 would much prefer [his_1 going to the movie]
   b. John_1 would much prefer [his_2 going to the movie]
   c. John_1 would much prefer [his_1 book]

Observation:
It is unclear why (53-a) is ungrammatical. (Constraints of the Binding theory cannot be involved, see (53-c) and below). Proposal:

(54) *Avoid Pronoun*$^{td/tl}$ (Chomsky (1981)):
   Lexical pronouns are blocked by empty pronouns if possible.

Note:
To make the Avoid Pronoun account work, we cannot adopt the null hypothesis according to which derivations (or output representations) compete with each other (i.e., are in the same *reference set*) if they go back to the same LA; see (56). Otherwise, (51-a) could not block (53-a). Thus, an independent way must be found to determine the reference set, i.e., the set of competing derivations (or output representations). Furthermore, we must assume that

a derivation that violates some local constraint (like the Constraint on Control) cannot block another derivation; see (57). A more precise definition of the Avoid Pronoun constraint might look as follows.

(55) *Avoid Pronoun*$^{td/tl}$ (different formulation):
If two derivations D_1 and D_2 are in the same reference set and D_1 uses a lexical pronoun where D_2 uses an empty pronoun, then D_1 is to be preferred over D_2.

(56) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff:
a. D_1 and D_2 start with the same LA.
b. D_1 and D_2 do not violate local or global constraints.

(57) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff:
a. D_1 and D_2 have identical lexical categories in the LA.
b. D_1 and D_2 have the same semantic interpretation.
c. D_1 and D_2 do not violate local or global constraints.

Chapter 4

Constraints on Binding

4.1 Principles A, B, C of the Binding Theory

Terminology (Chomsky (1981)):
An item is called *anaphor* if it is a *reflexive pronoun* or a *reciprocal pronoun*. An item is referred to as a *pronoun* if it is a *personal pronoun*. An item is called an *R-expression* ("referential expression") if it has the categorial feature [D] and does not qualify as an anaphor or as a pronoun (in the technical sense). Typically, R-expressions are *names* or *definite DPs*; but they may also include other kinds of DPs.

Note:
These notions are used by the three fundamental constraints of binding theory, viz., Principle A, Principle B, and Principle C. The constraints are representational; for the time being, we can take them to apply to syntactic output (S-structure) representations.

(1) a. *Principle A^r*:
An anaphor is bound in its binding domain.
b. *Principle B^r*:
A pronoun is not bound in its binding domain.
c. *Principle C^r*:
An R-expression is not bound.

Note:
The notions of *binding* and *binding domain* need to be clarified. For present purposes, (2) and (3) will suffice.

(2) *Binding*:
α binds β iff (a), (b), and (c) hold:

a. α and β are co-indexed.
b. α occupies an A-position.
c. α c-commands β.

Question:
What is an A-position? For present purposes, we can assume that A-positions are specifiers of lexical categories (N, V, A, P), and of the functional catgories D and T. SpecC is not an A-position (neither are modifier positions or, irrelevantly, complement positions).

(3) *Binding domain*:
The binding domain of some category α is the minimal XP that dominates a category β ($\beta \neq \alpha$) such that (a) or (b) holds:
a. β is an external argument.
b. β is a finite T.

Note:
In general, principles A and B predict that anaphors and pronouns are in *complementary distribution*. By and large, this seems to be correct (but there are a number of principled exceptions that we will ignore here, in particular in DP-internal contexts).

(4) *Consequences of Principle A*:
a. [$_{CP}$ C [$_{TP}$ John$_1$ [$_{T'}$ Ø [$_{VP}$ t$_1$ likes himself$_1$]]]]
b. *[$_{CP}$ C [$_{TP}$ John$_1$ [$_{T'}$ Ø [$_{VP}$ t$_1$ thinks [$_{CP}$ that [$_{TP}$ Mary$_2$ [$_{T'}$ Ø [$_{VP}$ t$_2$ likes himself$_1$]]]]]]]]
c. *[$_{CP}$ Fritz$_1$ glaubt [$_{CP}$ dass sich$_1$ dumm ist]]
Fritz$_{nom}$ believes that self stupid is
d. [$_{CP}$ C John$_1$ believes [$_{TP}$ himself$_1$ to be [$_{AP}$ t$_1$ clever]]]
e. John$_1$ likes [$_{DP}$ Ø [$_{NP}$ stories [$_{PP}$ about himself$_1$]]]
f. *John$_1$ likes [$_{DP}$ Bill's$_2$ [$_{NP}$ stories [$_{PP}$ about himself$_1$]]]

(5) *Consequences of Principle B*:
a. *[$_{CP}$ C [$_{TP}$ John$_1$ [$_{T'}$ Ø [$_{VP}$ t$_1$ likes him$_1$]]]]
b. [$_{CP}$ C [$_{TP}$ John$_1$ [$_{T'}$ Ø [$_{VP}$ t$_1$ thinks [$_{CP}$ that [$_{TP}$ Mary$_2$ [$_{T'}$ Ø [$_{VP}$ t$_2$ likes him$_1$]]]]]]]]
c. [$_{CP}$ Fritz$_1$ glaubt [$_{CP}$ dass er$_1$ dumm ist]]
Fritz$_{nom}$ believes that he stupid is
d. *[$_{CP}$ C John$_1$ believes [$_{TP}$ him$_1$ to be [$_{AP}$ t$_1$ clever]]]
e. ?John$_1$ likes [$_{DP}$ Ø [$_{NP}$ stories [$_{PP}$ about him$_1$]]]
f. John$_1$ likes [$_{DP}$ Bill's$_2$ [$_{NP}$ stories [$_{PP}$ about him$_1$]]]

(6) *Consequences of Principle C*:
  a. *[$_{\text{CP}}$ C [$_{\text{TP}}$ He$_1$ [$_{\text{T}'}$ Ø [$_{\text{VP}}$ t$_1$ likes John$_1$]]]]
  b. *[$_{\text{CP}}$ C [$_{\text{TP}}$ He$_1$ [$_{\text{T}'}$ Ø [$_{\text{VP}}$ t$_1$ thinks [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary$_2$ [$_{\text{T}'}$ Ø [$_{\text{VP}}$ t$_2$ likes John$_1$]]]]]]]]
  c. *He$_1$ likes [$_{\text{DP}}$ Bill's$_2$ [$_{\text{NP}}$ stories [$_{\text{PP}}$ about John$_1$]]]

4.2 Levels of Representation

Question:
Is there evidence that, e.g., a representational principle A must apply at S-structure, but not at D-structure or LF? Yes, there is:

(7) *An argument against Principle A at D-structure: Movement to SpecT makes A-binding possible*:
  a. *D-structure representation*:
  [$_{\text{CP}}$ C [$_{\text{TP}}$ T [$_{\text{VP}}$ [$_{\text{V'}}$ seems [$_{\text{PP}}$ to himself$_1$]] [$_{\text{TP}}$ to be [$_{\text{AP}}$ John$_1$ clever]]]]]
  b. *S-structure representation*:
  [$_{\text{CP}}$ C [$_{\text{TP}}$ John$_1$ [$_{\text{T}'}$ T [$_{\text{VP}}$ [$_{\text{V'}}$ seems [$_{\text{PP}}$ to himself$_1$]] [$_{\text{TP}}$ t$'_1$ to be [$_{\text{AP}}$ t$_1$ clever]]]]]]

Note:
At D-structure, *himself* is not A-bound; it finds an A-binder only after movement of the DP *John* to the matrix subject position (where [*nom*]/[*D*] on the finite T is deleted).

An independent problem:
The structure of VP adopted here is not unproblematic: It does not matter whether PP is a modifier or optionally selected – it seems clear that TP cannot occupy a complement position. Hence, it should be predicted to be a barrier blocking movement of the DP *John* via the Condition on Extraction Domains. Essentially, this reflects the recurring problem with double object constructions in the present system. Ultimately, the solution will have to be that there is an additional empty verb-like functional head that *seems* raises to by LI-movement in (7). If so, PP can be viewed as a specifier (or modifier), and TP as the complement of *seems* prior to LI-movement to the higher head position.

A remark on notation:
So far, we have assumed that movement leaves a trace that is co-indexed

with the moved item. Now we assume that binding also involves co-indexing; but this time, two separate categories are involved that are not related via movement. If the indices for binding and the indices for movement are treated in the same way (and they usually are), ambiguities may arise in syntactic representations. To avoid such ambiguities, a letter (*a* or *b*) is added to indices where needed in what follows. Only those items are related by movement that have an identical letter accompanying the general index – but for the purposes of binding theory, an identical number is sufficient to ensure co-indexing.

(8) *Why Principle A cannot apply (only) at LF: Quantifier raising makes A-binding impossible*:

a. *S-structure representation*:
*[$_{CP}$ C [$_{TP}$ Each other$_1$ [$_{T'}$ T [$_{VP}$ t$_1$ like [$_{DP_1}$ all students]]]]]

b. *Logical Form representation*:
*[$_{CP}$ C [$_{TP}$ [$_{DP_{1/b}}$ all students] [$_{TP}$ each other$_{1/a}$ [$_{T'}$ T [$_{VP}$ t$_{1/a}$ like t$_{1/b}$]]]]]

Note:
This analysis presupposes that quantifier raising (QR) is an LF movement operation that moves quantified phrases like *all students* to an outer SpecT position at LF. Given that SpecT is an A-position, a reciprocal in a lower SpecT would be predicted to be A-bound within TP at LF. Hence, under these assumptions, the evidence in (8) might be taken to suggest that Principle A does not solely apply at LF. (If it applies at S-structure and LF, (8) does not raise a problem anymore: The derivation is ill formed because there is one level of representation where its output representation violates a constraint.)

4.3 Problems with Principle A: Legitimate Unbound Anaphors

Note:
Movement operations applying to an anaphor or an XP that contains an anaphor can create contexts in which the anaphor is not bound at S-structure. Hence, we would expect a violation of Principle A^r at S-structure, and therefore ungrammaticality. However, ungrammaticality does not arise in these S-structure configurations (see van Riemsdijk & Williams (1981), Barss (1984; 1986), Chomsky (1995)).

(9) *Topicalization of the anaphor*:

a. *D-structure representation*:
[$_{\text{CP}}$ C [$_{\text{TP}}$ does not really [$_{\text{DP}_{1/a}}$ John] like [$_{\text{DP}_{1/b}}$ himself]]]
b. *S-structure representation*:
[$_{\text{CP}}$ [$_{\text{DP}_{1/b}}$ Himself] C [$_{\text{TP}}$ [$_{\text{DP}_{1/a}}$ John] does not really $t_{1/a}$ like $t_{1/b}$]]

(10) *Topicalization or wh-movement of an XP containing the anaphor*:
a. (i) *D-structure representation*:
Mary wondered [$_{\text{CP}}$ C [$_{\text{TP}}$ T [$_{\text{VP}}$ [$_{\text{DP}_{2/a}}$ Bill] saw [$_{\text{DP}_1}$ which picture of himself$_{2/b}$]]]]
(ii) *S-structure representation*:
Mary wondered [$_{\text{CP}}$ [$_{\text{DP}_1}$ which picture of himself$_{2/b}$] C [$_{\text{TP}}$ [$_{\text{DP}_{2/a}}$ Bill] T [$_{\text{VP}}$ $t_{2/a}$ saw t_1]]]
b. (i) *D-structure representation*:
[$_{\text{CP}}$ C [$_{\text{TP}}$ does not really [$_{\text{DP}_{1/a}}$ John] like [$_{\text{DP}_2}$ [$_{\text{D}}$ Ø] books about [$_{\text{DP}_{1/b}}$ himself]]]
(ii) *S-structure representation*:
[$_{\text{CP}}$ [$_{\text{DP}_2}$ [$_{\text{D}}$ Ø] Books about [$_{\text{DP}_{1/b}}$ himself]] C [$_{\text{TP}}$ [$_{\text{DP}_{1/a}}$ John] does not really $t_{1/a}$ like t_2]]

Note:
A weaker version of this problem arises if we assume that *wh*-in situ phrases move to a SpecC$_{[+wh]}$ position at LF, as in (11): The wellformedness of (11) shows that Princile A cannot apply solely at LF, like (8) did.

(11) *LF wh-movement of a wh-phrase containing an anaphor*:
a. *S-structure representation*:
Mary wondered [$_{\text{CP}}$ [$_{\text{DP}_{1/a}}$ who] C [$_{\text{TP}}$ $t'_{1/a}$ T [$_{\text{VP}}$ $t_{1/a}$ saw [$_{\text{DP}_2}$ which picture of himself$_{1/b}$]]]]
b. *LF representation*:
Mary wondered [$_{\text{CP}}$ [$_{\text{DP}_2}$ which picture of himself$_{1/b}$] [$_{\text{C}'}$ [$_{\text{DP}_{1/a}}$ who] C [$_{\text{TP}}$ $t'_{1/a}$ T [$_{\text{VP}}$ $t_{1/a}$ saw t_2]]]]

Note:
Belletti & Rizzi (1986) observe the same kind of phenomenon in *psych verb* constructions. A basic assumption (for which they provide independent motivation) is that the arguments that act as subjects in these constructions are not the external argument of the psych verb; rather, they are "derived" subjects in the sense that they must move across a higher argument into the subject position, as in (12). (As with double object constructions, problems arise with respect to linear precedence statements unless we are willing to

adopt a more complex structure of VP. For now, we put those problems aside.)

(12) *Structure of psych verb constructions*:
$[_{TP}$ $[_{DP_1}$ This picture] T $[_{VP}$ $[_{V'}$ bothers t_1] $[_{DP_2}$ John]]]

(13) *Movement to SpecT of a DP containing an anaphor, English*:
a. *$[_{TP}$ $[_{DP_1}$ Each other's$_2$ parents] T $[_{VP}$ t_1 promised $[_{DP_2}$ the girls] to buy cars]]
b. $[_{TP}$ $[_{DP_1}$ This picture of himself$_2$] T $[_{VP}$ $[_{V'}$ bothers t_1] $[_{DP_2}$ John]]]

(14) *Movement to SpecT of a DP containing an anaphor, Italian*:
a. $[_{TP}$ $[_{DP_1}$ Questi pettegolezzi su di sé$_2$] T $[_{VP}$ $[_{V'}$
these gossips about himself
preoccupano t_1] Gianni$_2$ più di ogni altra cosa]]
worry Gianni more than anything else
b. *$[_{TP}$ $[_{DP_1}$ Questi pettegolezzi su di sé] T $[_{VP}$ t_1 $[_{V'}$
these gossips about himself
descrivono Gianni$_1$ meglio di ogni biografia ufficiale]]]
describe Gianni better than any official biography

Note:
Not only can an anaphor contained in a moved XP escape the structural binding domain of a subject antecedent without inducing ungrammaticality; it can also find a new binder in the matrix clause this way. The ambiguity of examples like (15), (17) thus provides a second argument against assuming that Principle A applies at D-structure (recall (7)), and an argument against assuming that Principle A applies at S-structure (compare (9), (10), (13), (14)).

(15) *Wh-movement to SpecC makes A-binding possible, first example*:
a. *D-structure representation*:
$[_{CP}$ $C_{[-wh]}$ $[_{TP}$ T $[_{VP}$ $[_{DP_1}$ John] wondered $[_{CP}$ $C_{[+wh]}$ $[_{TP}$ T $[_{VP}$ $[_{DP_2}$ Bill] saw $[_{DP_3}$ which picture of himself$_{1,2}$]]]]]]]
b. *S-structure representation*:
$[_{CP}$ $C_{[-wh]}$ $[_{TP}$ $[_{DP_1}$ John] T $[_{VP}$ t_1 wondered $[_{CP}$ $[_{DP_3}$ which picture of himself$_{1,2}$] $C_{[+wh]}$ $[_{TP}$ $[_{DP_2}$ Bill] T $[_{VP}$ t_2 saw t_3]]]]]]

(16) *Long-distance binding is impossible without movement*:
$[_{CP}$ $C_{[-wh]}$ $[_{TP}$ $[_{DP_1}$ John] T $[_{VP}$ t_1 wondered $[_{CP}$ whether $[_{TP}$ $[_{DP_2}$ Bill] T $[_{VP}$ t_2 saw $[_{DP_3}$ a picture of himself$_{*1,2}$]]]]]]]

(17) *Wh-movement to SpecC makes A-binding possible, second example*:

a. *D-structure representation*:
[$_{CP}$ C$_{[+wh]}$ [$_{TP}$ does [$_{VP}$ [$_{DP_1}$ John] think [$_{CP}$ [$_{C_{[-wh]}}$ that] [$_{TP}$ T [$_{VP}$ [$_{DP_2}$ Bill] liked [$_{DP_3}$ which picture of himself$_{1,2}$]]]]]]]

b. *S-structure representation*:
[$_{CP}$ [$_{DP_3}$ Which picture of himself$_{1,2}$] [$_{C_{[+wh]}}$ does] [$_{TP}$ [$_{DP_1}$ John] T [$_{VP}$ t$_1$ think [$_{CP}$ t$'_3$ [$_{C_{[-wh]}}$ that] [$_{TP}$ [$_{DP_2}$ Bill] T [$_{VP}$ t$_2$ liked t$_3$]]]]]] ?

(18) *Long-distance binding is impossible without movement*:
[$_{CP}$ C$_{[-wh]}$ [$_{TP}$ [$_{DP_1}$ John] T [$_{VP}$ t$_1$ thinks [$_{CP}$ [$_{C_{[-wh]}}$ that] [$_{TP}$ [$_{DP_2}$ Bill] T [$_{VP}$ t$_2$ liked [$_{DP_3}$ this picture of himself$_{*1,2}$]]]]]]]

Conclusion:
If a representational version of Principle A is to be maintained, it must be revised in such a way that the effects of movement (no loss of binding options after S-structure movement, new binding options may arise after S-structure movement) can be "imitated" by the constraint. Intuitively, the creation of "new" binding options can be taken to support the idea that the constraint applies at S-structure; and the persistence of "old" binding options requires a concept of *reconstruction*. Here is a possible solution (that essentially goes back to Barss (1984); also see Barss (1986) for an even more complicated version of the general idea in terms of so-called *Chain Accessebility Sequences*).

(19) *Principle A^r* (revised):
At S-structure, an anaphor is chain-bound in its binding domain.

(20) *Chain-Binding*:
α chain-binds β iff (a), (b), and (c) hold:

a. α and β are co-indexed.
b. α occupies an A-position.
c. (i) α c-commands β, or
   (ii) α c-commands a trace of γ, where $\gamma = \beta$ or γ dominates β.

(21) *Binding domain* (as before):
The binding domain of some category α is the minimal XP that dominates a category β ($\beta \neq \alpha$) such that (a) or (b) holds:

a. β is an external argument in SpecX.
b. β is a finite X.

Note:
There is a potential technical problem: Suppose that an anaphor has been topicalized in a root clause, as in (9-b), repeated here in (22). Here, the anaphor does not seem to have any binding domain: The only XP that dominates the anaphor is CP, which does not have a β in the sense of (21). How, then can the anaphor in (22) fulfill Principle A in (19)? One assumption could be that the definition of binding domain is modified in such a way that the root CP qualifies as a binding domain if otherwise no binding domain can be determined.

(22) *Topicalization of the anaphor*:
$[_{CP}$ $[_{DP_{1/b}}$ Himself] C $[_{TP}$ $[_{DP_{1/a}}$ John] does not really $t_{1/a}$ like $t_{1/b}$]]

Question:
Why is there no ambiguity in (23-ab) (see Barss (1986), Huang (1993))?

(23) a. $[_{AP_3}$ t_1 How proud of himself$_{1/*2}$] did John$_2$ say $[_{CP}$ t'_3 Bill$_1$ became t_3]?
b. $[_{VP_3}$ t_1 Criticize himself$_{1,*2}$] John$_2$ thinks $[_{CP}$ t'_3 Bill$_1$ will not t_3]

Answer:
Recall that the structure of VP and AP is based on the argument structure of V and A, respectively (all arguments of a predicate are merged within that predicate's maximal projection); and that only maximal projections (XP) can undergo *wh*-movement or topicalization to SpecC. Hence, the fronted XPs in (24-ab) have (unbound) traces in specifier position that continue to erect a binding domain for the anaphors after the movement operation.

General problem:
The new representational Principle A is not conceptually attractive because it simply states properties of binding that should independently result from the role of movement in syntax.

4.4 Problems with Principle C: Illegitimate Unbound R-Expressions

Observation:
Just as movement does not destroy anaphoric options, it does not create new options for R-expressions (or pronouns).

(24) *Topicalization of R-expressions*:

a. *[$_{CP}$ C [$_{TP}$ He$_{1/a}$ does not really t$_{1/a}$ like John$_{1/b}$]]
b. *[$_{CP}$ John$_{1/b}$ C [$_{TP}$ he$_{1/a}$ does not really t$_{1/a}$ like t$_{1/b}$]]

(25) *Wh-movement of an XP containing the R-expression*:
a. *He$_1$ was willing to discuss [$_{DP_2}$ the claim [$_{CP}$ that John$_1$ was asleep]]
b. *[$_{DP_2}$ Which claim [$_{CP}$ that John$_1$ was asleep]] was he$_1$ willing to discuss t$_2$?

Note:
Again, a problem arises for the assumption that Principle C applies at S-structure. And again, a reformulation of Principle C that relies on the notion of chain-binding will fix the problem. (Similar conclusions hold for Principle B.)

(26) *Principle C'* (revised):
An R-expression is not chain-bound.

Note:
However, there is an interesting exception to the generalization that movement does not change binding options for R-expressions. Examples like (27) seem well-formed for many speakers. This is known as an *anti-reconstruction* effect.

(27) *Anti-reconstruction with wh-movement of an XP containing the R-expression*:
*[$_{DP_2}$ Which claim [$_{CP}$ that John$_1$ made]] was he$_1$ willing to discuss t$_2$?

Note:
(25-b) and (27) form a minimal pair. The crucial difference is that CP is an argument of N in (25-b), and a modifier of N in (27).

4.5 A Derivational Reinterpretation of the Principles of Binding Theory

Note:
Throughout, the system in [1] ("Phrase Structure and Derivations") is adopted again, i.e., Move and Merge alternate throughout a derivation, and sentences grow until they reach the root C.

4.5.1 Principle C

Note:
Except for (27), Principle C can straightforwardly be reinterpreted as a derivational constraint that holds at every step of the derivation. No recourse to concepts like chain-binding is necessary: As soon as an R-expression is bound, the constraint will be violated, and ungrammaticality arises.

(28) *Principle C^d*:
An R-expression is not bound.

(29) *An illustration of Principle C^d effects*:
a. *[$_{\text{CP}}$ C [$_{\text{TP}}$ [$_{\text{DP}_1}$ He] [$_{\text{T}'}$ Ø [$_{\text{VP}}$ t_1 likes [$_{\text{DP}_1}$ John]]]]]
b. *Derivation*:
Merge ([$_{\text{V}}$ likes], [$_{\text{DP}_1}$ Ø John]) → [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}_1}$ Ø John]]
c. Merge ([$_{\text{DP}_1}$ he], [$_{\text{VP}}$ [$_{\text{V}}$ likes] [$_{\text{DP}_1}$ Ø John]]) →
*[$_{\text{VP}}$ [$_{\text{DP}_1}$ he] [$_{\text{V}'}$ [$_{\text{V}}$ likes] [$_{\text{DP}_1}$ Ø John]]]
(→ Violation of Principle C^d)

Question:
Is there anything that can be done to prevent a violation of Principle C^d in (27)?

Assumption (Freidin (1986), Lebeaux (1988), Chomsky (1995), Epstein et al. (1998)):
Modifiers can be merged counter-cyclically, arguments and other selected items cannot be merged counter-cyclically.

(30) *Strict Cycle Condition^d* (revised):
No operation *that involves the deletion of a selectional feature* can apply to a domain dominated by a cyclic node α in such a way as to affect solely a proper subdomain of α dominated by a node β which is also a cyclic node.

Note:
The constraint Timing of Feature Deletion in (29) on page 7 of [1] can actually be viewed as a version of the Strict Cycle Condition that has this effect.

4.5.2 Principle A

Assumption:
Let us assume that Principle A is a derivational constraint that restricts every syntactic operation.

(31) *Principle A^d*:
An anaphor is bound in its binding domain.

Note:
Since anaphors are usually first be merged with a predicate before its antecedent enters the phrase marker (except for cases like (7)), a straightforward derivational reinterpretation of Principle A makes problematic predictions: The anaphor may not have a binding domain yet at the point where it is introduced. But then, a presupposition failure would arise, and Principle A could not be fulfilled.

(32) *A wrong prediction*:
a. $[_{CP}$ C $[_{TP}$ John$_1$ $[_{T'}$ Ø $[_{VP}$ t$_1$ likes himself$_1$]]]]
b. *Derivation*:
Merge ($[_V$ likes], $[_{DP_1}$ himself]) → *$[_{VP}$ $[_V$ likes] $[_{DP_1}$ himself]]
(→ Violation of Principle A^d!)

Note:
One might want to fix this problem by revising Principle A^d as in (33):

(33) *Principle A^d* (revised):
If an anaphor α has a binding domain β, then α is bound in β.

Consequence:
This still does not help: It will be impossible for an anaphor to extend its binding domain (and find an antecedent in a higher clause) by movement; but this is needed for cases like (15), (17).

Conclusion:
The problem with Principle A^d in (31)/(33) is that it is assumed to hold at every derivational step. The *universal quantification* embodied in this assumption works well for constraints like Principle B and Principle C (and for locality constraints like those discussed in [2] and [3]), but not for a constraint like Principle A. Here, an *existential quantification* is needed, as in (34) (see Belletti & Rizzi (1986), Epstein et al. (1998)).

(34) *Principle A^g* (second revision):
An anaphor is bound in its binding domain at some point of the derivation.

Note:
This accounts for all the data discussed so far. However, it seems that (34) cannot simply be checked at any given step of the derivation. Rather, the whole derivation must be considered, and there must be at least one step where the anaphor is bound within its binding domain. Hence, (34) does in fact qualify as a global constraint; it is not local anymore.

(35) *A final interesting example*:
Mary wondered [$_{CP}$ [$_{DP_3}$ which claim [$_{CP}$ that pictures of herself disturbed Bill]] he made t_3]

Chapter 5

Transderivational Constraints in the Minimalist Program

5.1 Fewest Steps

5.1.1 V-to-T Movement in Chomsky (1991)

Background:
Chomsky (1991) is concerned with deriving the difference between French and English with respect to V-to-T movement (cf. Pollock (1989)): French has overt V-to-T movement of finite verbs; English does not have such movement (except for auxiliaries). This is shown in (1).

(1) *V-to-T in French vs. English*:
  a. Jean embrasse$_1$ souvent [$_{VP}$ t$_1$ Marie]
  b. *Jean souvent [$_{VP}$ embrasse$_1$ Marie]
  c. *John kisses$_1$ often [$_{VP}$ t$_1$ Mary]
  d. John often [$_{VP}$ kisses$_1$ Mary]

Questions:
Why do French verbs have to move to T? Why must English main verbs not move to T? Assumption: French has strong T nodes, English has weak T nodes.

(2) *Strength of T^d*:
  Strong T tolerates adjunction of all Vs; weak T tolerates adjunction only of "light" Vs (auxiliaries).

Note:
This excludes (1-c) in English: Overt V-to-T movement violates Strength of T. In contrast, overt V-to-T movement in (1-a) in French does not violate

this constraint. Still, something extra needs to be said about (1-b) in French, which vacuously fulfills Strength of T, just like (1-d) in English does. Thus, the question is: Why is overt V-to-T raising obligatory if it is possible? Chomsky's (1991) background assumption is that inflection is base-generated in T. If V does not raise to T, T must lower to V in overt syntax, so as to fulfill another local constraint, the *Stray Affix Filter*, which prohibits inflectional affixes that are not attached to a verbal host. The crucial idea now is that overt T lowering creates an unbound trace that must be undone by LF, via covert raising of the whole V-T complex to the position of the trace of T. The derivations underlying (1-a) and (1-b) in French are given in (3) and (4), respectively.

(3) *The V-to-T Derivation in French*:
   a. Jean T_2 souvent [$_{VP}$ embrasse$_1$ Marie] (overt raising)
   b. Jean [$_{T_2}$ embrasse$_1$-T] souvent [$_{VP}$ t_1 Marie]

(4) *The V-in situ Derivation in French*:
   a. Jean T_2 souvent [$_{VP}$ embrasse$_1$ Marie] (overt lowering)
   b. Jean t_2 souvent [$_{VP}$ [$_{V_1}$ embrasse$_1$-T_2] Marie] (covert raising)
   c. Jean [$_{V_1}$ embrasse$_1$-T_2] souvent [$_{VP}$ t_1 Marie]

Idea:
The second derivation has more movement steps than the first one, and it is therefore filtered out as uneconomical by the transderivational economy constraint *Fewest Steps* that can be formulated as follows:

(5) *Fewest Steps^{td}*:
   If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer operations than D_2, then D_1 is to be preferred over D_2.

(6) *Reference Set*:
   Two derivations D_1 and D_2 are in the same reference set iff:
   a. D_1 and D_2 start with the same LA.
   b. D_1 and D_2 do not violate local or global constraints.

Consequences:
(i) (6-a): (7-b) cannot accidentally block (7-a) even though it involves fewer syntactic operations.
(ii) (6-b): (7-c) cannot accidentally block (7-a) even though it involves fewer syntactic operations by leaving the *wh*-phrase and the auxiliary in situ – (7-c) violates the constraint that selectional features ([*Q*] on C) have to be deleted.

(7) *Motivation for reference sets*:

a. What$_1$ have$_2$ you t$_2$ seen t$_1$?
b. You have seen a car
c. *You have$_2$ seen what$_1$?

Historical comment:
This particular application of a transderivational constraint in the minimalist program is not generally accepted anymore. Still, it can serve as an illustration of certain recurring properties and problems of transderivational constraints, and of competition-based approaches in general.

5.1.2 Problems

Problem no. 1:
Chomsky (1991, 433) observes: This system "tends to eliminate the possibility of optionality in derivation. Choice points will be allowable only if the resulting derivations are all minimal in cost ... This may well be too strong a conclusion, raising a problem for the entire approach."

(8) *Optional topicalization*:

a. Mary gave a book to John$_1$
b. To John$_1$ Mary gave a book t$_1$

Solutions:
(i) Chomsky (1991): Certain movement operations might be "assigned to some other component of the language system, perhaps a 'stylistic' component of the mapping ... to PF;" movement operations of this type might then be exempt from the Fewest Steps constraint.
(ii) We might revise the definition of reference set appropriately, such that the two derivations in (8) do not compete anymore. For instance, we might add the requirement that competing derivations must have identical LF representations, as in (9). Assuming that sentences which differ only with respect to whether topicalization has applied must have different LFs, this would yield the desired result.

(9) *Reference Set* (revised):
Two derivations D$_1$ and D$_2$ are in the same reference set iff:

a. D$_1$ and D$_2$ start with the same LA and have the same LF representation.
b. D$_1$ and D$_2$ do not violate local or global constraints.

Problem no. 2:
Transderivational economy constraints increase complexity: To find out whether a given sentence is grammatical, it does not suffice to look at internal properties of the sentence (Does it violate a local constraint?); rather, the properties of other sentences have to be taken into account as well (Does the sentence have the most economical derivation in the reference set?). Chomsky (1991, 448): "Language design as such appears to be in many respects 'dysfunctional,' yielding properties that are not well adapted to the functions language is called upon to perform."

Problem no. 3:
Successive-cyclic movement: It is standardly assumed that long-distance *wh*-movement must be successive-cyclic; otherwise, a locality constraint will be violated:

(10) *Wh-Island Effects*:
*How$_1$ do you wonder [$_{CP}$ which car$_2$ to fix t$_2$ t$_1$] ?

(11) *Successive-cyclic movement and Fewest Steps*:
a. How$_1$ do you think [$_{CP}$ t''_1 that John said [$_{CP}$ t'_1 that Bill fixed the car t$_1$]] ?
b. *How$_1$ do you think [$_{CP}$ that John said [$_{CP}$ that Bill fixed the car t$_1$]] ?

A caveat:
(10) violates the Subjacency Condition; but given that this constraint is not widely adopted anymore, it is by no means clear how (11-b) can in fact be excluded. As noted before, the main locality constraints adopted in the minimalist program are (a) the Minimal Link Condition (= F-over-F Principle & Superiority Condition), and (b) the Condition on Extraction Domains. The Minimal Link Condition is in fact violated in (10), but not in (11-b); and neither is the Condition on Extraction Domains. How, then, can we ensure that locality constraints demand successive-cyclic movement?

Chomsky's (2000; 2001) proposal:
Successive-cyclic movement is forced by the Phase Impenetrability Condition.

(12) *Phase Impenetrability Condition*[d]:
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and SpecX are accessible to such operations.

Note:
The constraint relies on the existence of special derivational units: phases.

(13) *Phase*:
The propositional categories CP and VP are phases; other XPs (except perhaps for DP) are not.

Note:
It follows that movement must be successive-cyclic in an even stricter sense that assumed so far:
(i) Movement from CP must take place via SpecC.
(ii) Movement from VP must take place via SpecV.
However, (ii) will be disregarded in what follows. (I.e., for present purposes, we can assume that only CPs are phases.)

A consequence:
Successive-cyclic movement creates a potential problem for the Fewest Steps condition: Successive-cyclic movement as in (11-a) should always be blocked by one-step movement as in (11-b). We would then incorrectly expect all long-distance dependencies to be ill formed.

But:
There is no problem under the notion of reference set in (6) or (9): Only those derivations can compete that respect all local constraints of grammar, i.e., that are otherwise well formed. The derivation that generates the surface representation (11-b) violates the Phase Impenetrability Condition; hence, it cannot compete with the derivation that generates (11-a), and (11-a) is chosen by Fewest Steps because there is no competing derivation that would be more economical.

However:
Suppose that we were to dispense with clause (b) in the definition of reference set, or that clause (b) would be weakened in such a way that some derivations violating local constraints can compete after all. (As we will see below, there is some evidence for this latter option.) If so, the derivations generating (13-a) and (13-b) might compete, and the problem of accounting for successive-cyclic movement under Fewest Steps persists. How, then, can we permit successive-cyclic movement in (13)? We have seen Chomsky's solution already in [3]: Successive-cyclic movement is reinterpreted as a single (albeit complex) operation:

(14) *Form Chain*:
Move α to its target position and freely insert intermediate traces in appropriate positions.

Note:
The French case of T-lowering at S-structure followed by V-raising at LF in (4) would still have to involve two applications of Form Chain.

5.1.3 Wh-Topicalization in Epstein (1992)

Goal:
Epstein accounts for the ban on *wh*-topicalization in English in terms of the transderivational Fewest Steps constraint.

(15) *Optional topicalization: different LF representations*:
a. Who_1 t_1 said [$_{CP}$ that [$_{TP}$ Mary gave a book to $John_2$]] ?
b. Who_1 t_1 said [$_{CP}$ that to $John_2$ [$_{TP}$ Mary gave a book t_2]] ?

(16) *The ban on wh-topicalization*:
a. Who_1 t_1 said [$_{CP}$ that [$_{TP}$ Mary gave a book to $whom_2$]] ?
b. *Who_1 t_1 said [$_{CP}$ that to $whom_2$ [$_{TP}$ Mary gave a book t_2]] ?

Assumption:
All *wh*-phrases must be in $SpecC_{[+wh]}$ at LF.

(17) *LF representations of (16-a), (16-b)*:
*Who_1 to $whom_2$ t_1 said [$_{CP}$ that [$_{TP}$ Mary gave a book t_2]] ?

Analysis:
D_1 reaches this LF by applying one (covert) instance of *wh*-movement to the embedded object DP who_2. There is only one movement operation in this case, either because LF movement of arguments does not have to be successive cyclic, or because successive-cyclic covert movement can be analyzed as one instance of Form Chain. D_2, on the other hand, reaches the same LF by applying two instances of *wh*-movement to the embedded object DP who_2 (viz., one overtly and one covertly – given the intervening S-structure from which the PF derivation branches, these two movement operations cannot be reanalyzed as one instance of Form Chain). Hence, D_1 blocks D_2 via Fewest Steps.

Note:
The same analysis can be given for ill-formed *wh*-scrambling in German (see Müller & Sternefeld (1996)).

(18) *The ban on wh-scrambling in German*:

a. Warum$_1$ hat der Fritz was$_1$ gelesen ?
why has ART Fritz what read

b. *Warum$_1$ hat was$_1$ der Fritz t$_1$ gelesen ?
why has what ART Fritz read

(19) *The ban on wh-topicalization in German*:

a. Wer$_1$ sagte t$_1$ [$_{CP}$ dass Maria wem$_2$ ein Buch gegeben hat$_3$] ?
who said that Maria whom a book given has

b. *Wer$_1$ sagte t$_1$ [$_{CP}$ wem$_2$ hat$_3$ Maria t$_2$ ein Buch gegeben t$_3$] ?
who said whom has Maria a book given

Problem:
German topicalization always requires verb-second movement, which is incompatible with the presence of a complementizer in German. Hence, D_1 generating (19-b) and D_2 generating (19-a) do not share an identical lexical array. Solution:

(20) *Reference Set* (second revision):
Two derivations D_1 and D_2 are in the same reference set iff:

a. D_1 and D_2 have the same LF representation.

b. D_1 and D_2 do not violate local or global constraints.

Assumption:
A complementizer *dass* is deleted (or invisible) at LF.

Another problem:
Optional partial *wh*-movement to a SpecC$_{[-wh]}$ position in Ancash Quetchua (see Cole (1982)).

(21) *Partial wh-movement in Ancash Quechua*:

a. [$_{CP}$ Ima-ta-taq$_1$ (qam) kreinki [$_{CP}$ t$''_2$ Maria muna-nqa-n-ta [$_{CP}$
what$_{acc}$ you believe Maria want-nom-3-acc
t$'_1$ José t$_1$ ranti-na-n-ta]]] ?
José buy-nom-3-acc

b. [$_{CP}$ – (Qam) kreinki [$_{CP}$ ima-ta-ta$_1$ Maria muna-nqa-n-ta [$_{CP}$ t$'_1$ José t$_1$ ranti-na-n-ta]]] ?

c. [$_{CP}$ – (Qam) kreinki [$_{CP}$ – Maria muna-nqa-n-ta [$_{CP}$ ima-ta-ta$_1$ José t$_1$ ranti-na-n-ta]]] ?

d. [$_{CP}$ – (Qam) kreinki [$_{CP}$ – Maria muna-nqa-n-ta [$_{CP}$ –José ima-ta-ta$_1$ ranti-na-n-ta]]] ?

(22) *Reference Set* (third revision):
Two derivations D_1 and D_2 are in the same reference set iff:
a. D_1 and D_2 have the same S-structure representation.
b. D_1 and D_2 do not violate local or global constraints.

Result:
Partial *wh*-movement is permitted, but most of the original evidence in favour of Fewest Steps is lost: Thus, on this view, neither French V-in situ, nor English (or German) *wh*-topicalization can be ruled out by Fewest Steps anymore. As noted in Sternefeld (1997), this situation might be viewed as indicative of a general problem with transderivational constraints: A significant reduction of competition in reference sets may be empirically desirable so as to account for cases of optionality (as in partial *wh*-movement constructions); but as an unwanted side effect, it also threatens to undermine the notion of transderivational economy itself: Many ill-formed derivations that could be ruled out by transderivational constraints will now survive because the more economical derivation is not part of the same reference set anymore. Finding a suitable definition of reference set that is weak enough to permit optionality and strong enough to actually do some work is one of the fundamental concerns of all versions of the minimalist program that employ the notion of competition.

5.1.4 Freezing in Collins (1994)

Background:
Evidence for yet another definition of reference sets comes from Collins' (1994) account of freezing effects with A-movement in English. We have seen that subject DPs are islands in English, whereas complement DPs permit extraction (with certain types of verbs). An interesting case is that of subject DPs that originate in complement position, as in the case of passivization. As can be seen in (23-c), such derived subject DPs are also islands.

(23) *Derived subject DPs as islands*:
a. Who$_1$ did John take [$_{DP}$ a picture of t$_1$] ?
b. *Who$_1$ is [$_{DP}$ a picture of t$_1$] on sale ?
c. *Who$_1$ was [$_{DP_2}$ a picture of t$_1$] taken t$_2$ by John ?

(24) *D_1 violates the Condition on Extraction Domains*:
a. [$_{VP}$ taken [$_{DP_2}$ a picture of who$_1$] by John] ...
b. [$_{TP}$ [$_{DP_2}$ a picture of who$_1$] T [$_{VP}$ taken t$_2$ by John]]
c. *[$_{CP}$ who$_1$ was [$_{TP}$ [$_{DP_2}$ a picture of t$_1$] T [$_{VP}$ taken t$_2$ by John]]]

(25) *Condition on Extraction Domains*:
  a. Movement must not cross a barrier.
  b. An XP is a barrier iff it is not a complement.

(26) *D_2 violates the Strict Cycle Condition*:
  a. [$_{VP}$ taken [$_{DP_2}$ a picture of who$_1$] by John] ...
  b. [$_{CP}$ who$_1$ was [$_{TP}$ T [$_{VP}$ taken [$_{DP_2}$ a picture of t$_1$] by John]]]
  c. [$_{CP}$ who$_1$ was [$_{TP}$ [$_{DP_2}$ a picture of t$_1$] T [$_{VP}$ taken t$_2$ by John]]]

(27) *Strict Cycle Condition*^{d} (simplified):
No operation can apply to a domain dominated by a cyclic node α in such a way as to affect solely a proper subdomain of α dominated by a node β which is also a cyclic node.

(28) *D_3 exhibits chain interleaving*:
  a. [$_{VP}$ taken [$_{DP_2}$ a picture of who$_1$] by John]
  b. [$_{VP}$ who$_1$ [$_{V'}$ taken [$_{DP_2}$ a picture of t$_1$] by John]]
  c. [$_{TP}$ [$_{DP_2}$ a picture of t$_1$] T [$_{VP}$ who$_1$ [$_{V'}$ taken t$_2$ by John]]]
  d. [$_{CP}$ who$_1$ was [$_{TP}$ [$_{DP_2}$ a picture of t$_1$] T [$_{VP}$ t$'_1$ [$_{V'}$ taken t$_2$ by John]]]]

Problem:
How is chain interleaving in D_3 ruled out?

Solution:
D_3 is blocked by D_1 and D_2 via Fewest Steps: Other things being equal, D_3 needs three movement steps where D_1 and D_2 make do with two movement steps.

Note:
The argument presupposes that the first movement operation in D_3 can somehow respect the Economy Constraint on Merge (i.e., there must be a selectional feature that is deleted by the Move operation to SpecV).

Consequence:
We need yet another definition of reference sets. The three derivations D_1, D_2, and D_3 yield the same surface string, which is ill formed. Thus, the more economical derivations that D_3 is blocked by via Fewest Steps are not well-formed derivations, as in the applications of Fewest Steps discussed above, but rather ill-formed derivations that violate local constraints. This reasoning implies that reference sets can in fact not be defined as

assumed so far, by requiring that only those derivations can compete that satisfy all local constraints – in the case at hand, D_1 and D_2 violate local constraints. Still, we cannot simply drop this requirement in the definition of reference sets; otherwise, all instances of movement would invariably be blocked in favour of in-situ derivations by Fewest Steps, and syntactic derivations would be fairly trivial. It seems that what is needed in view of this conflicting evidence is a relativized notion of local constraint satisfaction.

The solution: convergence (Chomsky (1993; 1995)):
Only those derivations can compete with respect to transderivational constraints that converge. Essentially, whereas all violations of local constraints lead to ungrammaticality, only a subset of violations of local constraints also lead to non-convergence. Ungrammatical derivations that converge may then still be used to block other derivations as ungrammatical, as in the freezing construction discussed by Collins (1994).

(29) *Convergence*:
  a. Violations of constraints related to selectional features and their deletion lead to non-convergence.
  b. Violations of locality constraints and the Strict Cycle Condition lead to convergence.

(30) *Reference Set* (fourth revision):
Two derivations D_1 and D_2 are in the same reference set iff:
  a. D_1 and D_2 have the same LA/S-structure/LF.
  b. D_1 and D_2 converge.

5.2 Shortest Paths

(31) *Shortest Paths*^{td} (Chomsky (1993; 1995)):
If two derivations D_1 and D_2 are in the same reference set and the movement paths of D_1 are shorter than the movement paths of D_2, then D_1 is to be preferred over D_2.

5.2.1 Superiority Effects in Chomsky (1993) and Kitahara (1993)

(32) *Superiority effects revisited*:
  a. I wonder [$_{CP}$ who$_1$ C [$_{TP}$ t$_1$ bought what$_2$]]
  b. *I wonder [$_{CP}$ what$_2$ C [$_{TP}$ who$_1$ bought t$_2$]]
  c. Whom$_1$ did John persuade t$_1$ [$_{CP}$ to visit whom$_2$] ?
  d. *Whom$_2$ did John persuade whom$_1$ [$_{CP}$ t$'_2$ to visit t$_2$] ?

(33) *Superiority Condition*[d] (Chomsky (1973)):
In a structure $\alpha_{[*F*]}$... [... $\beta_{[F]}$ ... [... $\gamma_{[F]}$ ...] ...] ..., movement to [*F*] can only affect the category bearing the [F] feature that is closer to [*F*].

Assumption:
Superiority effects can be accounted for by the transderivational condition Shortest Paths. For instance, the movement path from t_1 to *whom*$_1$ in (32-c) is shorter than the movement path from t_2 to *whom*$_2$ in (32-d).

Problem:
At least some of the evidence for Fewest Steps (the ban on V-in situ in French, the ban on *wh*-topicalization and *wh*-scrambling in English and German) has relied on the assumption that covert movement counts in the same way that overt movement does. But assuming that LF movement also counts for the Shortest Paths condition leads straightforwardly into a dilemma: In the case at hand, the derivation that has the shorter overt *wh*-movement path invariably has the longer covert *wh*-movement path, and it seems that by LF, both derivations have *wh*-movement paths of equal length. Hence, ceteris paribus, both should be well formed.

Solutions:
(i) There is in fact no covert *wh*-movement of any kind (this is incompatible with Fewest Steps approach outline above).
(ii) Whereas Fewest Steps compares whole derivations, Shortest Paths compares only the overt parts of derivations.

Another problem:
LF-optionality: Sentences like (34-a) have two possible readings (see Baker (1970)) that correspond to two different LF representations, given LF movement of *wh*-in situ elements.

(34) *Baker sentences*: Who_1 t_1 wonders $[_{CP}$ $where_2$ we bought $what_3$ t_2] ?
a. who_1 $what_3$ t_1 wonders $[_{CP}$ $where_2$ we bought t_3 t_2]
Answer: John wonders where we bought the books, Mary wonders where we bought the records, etc.
b. who_1 t_1 wonders $[_{CP}$ $where_2$ $what_3$ we bought t_3 t_2]
Answer: John wonders where we bought what, Mary wonders where we bought what, etc.

Problem:
Given that all *wh*-in situ phrases must undergo movement to a $\text{SpecC}_{[+wh]}$ position at LF, D_2 (creating (34-b)) should block D_1 (creating (34-a)) because D_1's paths are longer.

Solutions:
(i) As before: Covert *wh*-movement either does not exist, or does not count with respect to the Shortest Paths condition.
(ii) Reference sets are defined in such a way that competing derivations must have identical LF representations.

5.2.2 Yo-Yo Movement in Collins (1994)

Note:
Yo-yo movement characterizes a combination of lowering and raising operations affecting a single item in the course of a derivation. Collins (1994) shows that the availability of yo-yo movement would make a wrong prediction for the West African language Ewe, and attempts to derive a ban on yo-yo movement from the Shortest Paths condition.

Background:
Ewe is among the languages that show reflexes of of successive-cyclic *wh*-movement in the C domain. The reflex of successive cyclicity concerns the morphological form of the 3.Pers.Sing. subject pronoun in the canonical subject position. The regular form of the pronoun is *é*; cf. (35-a). The regular pronoun *é* can be replaced by *wo* in cases of long-distance extraction (focus movement, in the case at hand); cf. (35-b).

(35) *Successive cyclicity in Ewe*:

a. Kofi gblǒ [$_{\text{CP}}$ be é/*wo ʃo Kǒsi]
   Kofi said that he hit Kǒsi

b. Kofi$_1$ ϵ me gblǒ [$_{\text{CP}}$ (t$'_1$) be é/wo ʃo t$_1$]
   Kofi Foc I said that he hit

Assumptions:
(i) *é* is replaced by *wo* iff the local SpecC position is filled.
(ii) Long-distance movement of argument DPs in Ewe may apply either successive-cyclically, via SpecC (in which case *wo* is obligatory), or in one step (in which case *wo* is impossible).

(36) *Movement originating in the matrix CP*:

a. Kofi$_1$ ϵ me gblŏ na t$_1$ [$_{CP}$ be é/*wo ʃo Kŏsi]
Kofi Foc I said to that he hit Kŏsi

b. *A derivation that involves lowering*:
(i) [$_{VP}$ said [$_{PP}$ to Kofi$_1$] [$_{CP}$ that [$_{TP}$ he hit Kŏsi]]]
(ii) [$_{VP}$ said [$_{PP}$ to t$_1$] [$_{CP}$ Kofi$_1$ that [$_{TP}$ he hit Kŏsi]]] ...
(iii) [$_{CP}$ Kofi$_1$ Foc [$_{TP}$ T [$_{VP}$ said [$_{PP}$ to t$_1$] [$_{CP}$ t$'_1$ that [$_{TP}$ he hit Kŏsi]]]]]

Note:
The illformedness of this derivation also follows from the Strict Cycle Condition. Still, Collins (1994) observes that D_1 in (36) is blocked by the derivation D_2 in (37) via Shortest Paths. D_2 proceeds without intermediate lowering. (One might think that Fewest Steps would also suffice to block D_1 in favour of D_2. However, assuming that the two movement operations in D_1 can be reanalyzed as a single instance of Form Chain, this is not the case.)

(37) *A derivation that does not involve lowering*:
a. [$_{VP}$ said [$_{PP}$ to Kofi$_1$] [$_{CP}$ that [$_{TP}$ he hit Kŏsi]]]
b. [$_{CP}$ Kofi$_1$ Foc [$_{TP}$ T [$_{VP}$ said [$_{PP}$ to t$_1$] [$_{CP}$ that [$_{TP}$ he hit Kŏsi]]]]]

A consequence for the definition of reference set:
"He" is a *wo* in D_1 and an *é* in D_2. Hence, this difference must not suffice to create different reference sets.

5.2.3 Tagalog Wh-Movement in Nakamura (1998)

Generalization:
Only the highest A-position of a given clause (the subject position SpecT) is accessible for *wh*-movement in the Austronesian language Tagalog. In constructions in which an agent DP occupies the highest A-position (the so-called *Agent Topic* (AT) constructions), this DP can be *wh*-moved; an DP bearing a different Theta-role that shows up in an object position cannot undergo *wh*-movement.

(38) *Wh-movment in AT constructions*:
a. [$_{CP}$ Sino$_1$ ang [$_{TP}$ t$'_1$ b-um-ili [$_{VP}$ t$_1$ t$_V$ ng damit$_2$]]] ?
who Ang bought$_{AT}$ dress$_{inh}$
'Who is the one that bought the dress?'

b. *[$_{CP}$ Ano$_2$ ang [$_{TP}$ *si Juan*$_1$ b-um-ili [$_{VP}$ t$_1$ t$_V$ t$_2$]]]?
what Ang Juan$_{abs}$ bought$_{AT}$
'What is the thing that Juan bought?'

Observation:
A different marking on the verb triggers the so-called *Theme Topic* (TT) construction. Here, the theme DP occupies the structural subject position SpecT; and indeed, only the theme DP can undergo *wh*-movement.

(39) *Wh-movment in TT constructions*:

a. *[$_{CP}$ Sino$_1$ ang [$_{TP}$ *ang damit*$_2$ b-in-ili [$_{VP}$ t$_1$ t$_V$ t$_2$]]] ?
who Ang dress$_{abs}$ bought$_{TT}$
'Who is the one that bought the dress?'

b. [$_{CP}$ Ano$_2$ ang [$_{TP}$ t$'_2$ b-in-ili [$_{VP}$ ni Juan t$_V$ t$_2$]]] ?
what Ang bought$_{TT}$ Juan$_{erg}$
'What is the thing that Juan bought?'

Idea:
The derivations generating (38-a) and (39-a) compete, as do the derivations generating (38-b) and (39-b). The derivations underlying (38-a) and (39-b) can then block their respective competitors as ungrammatical because of the Shortest Paths constraint. Consider *wh*-movement of the theme DP in (38-b) and (39-b). The movement path from the VP-internal object position to the SpecC target position in (38-b) is longer than the path from the subject position SpecT to SpecC in (39-b).

Consequences:
(i) Derivations can compete even though they do not have identical lexical material – the Agent Topic and the Theme Topic constructions clearly differ in lexical make-up: "identical lexical array" is replaced by "non-distinct lexical array" in the definition of reference sets. The latter is defined in such a way that two lexical arrays that only differ with respect to functional features do not count as distinct.
(ii) The derivation that generates (39-b) may minimize the *wh*-path in comparison with the derivation that generates (38-b), but it increases path lengths in the TP-domain. Nakamura (1998) replaces the notion of "movement paths" in the definition of the Shortest Paths condition with the more specific notion of "comparable chain links".

5.2.4 Freezing in Chomsky (1995)

(40) *Derived subject DP islands again*:
*Who$_1$ was [$_{DP_2}$ a picture of t$_1$] taken t$_2$ by John ?

Recall from the discussion of Collins (1994):
(i) D_1 violates the Condition on Extraction Domains.
(ii) D_2 violates the Strict Cycle Condition.
(iii) D_3 violates the Fewest Steps condition.

Proposal:
D_2 can be accounted for without invoking the Strict Cycle Condition, by invoking the Shortest Paths Condition.

(41) a. *D_1: Condition on Extraction Domains*:
   (i) [$_{VP}$ taken [$_{DP_2}$ a picture of who$_1$] by John] ...
   (ii) [$_{TP}$ [$_{DP_2}$ a picture of who$_1$] T [$_{VP}$ taken t_2 by John]]
   (iii) *[$_{CP}$ who$_1$ was [$_{TP}$ [$_{DP_2}$ a picture of t_1] T [$_{VP}$ taken t_2 by John]]]
  b. *D_2: Shortest Paths Condition*:
   (i) [$_{VP}$ taken [$_{DP_2}$ a picture of who$_1$] by John]] ...
   (ii) [$_{CP}$ who$_1$ was [$_{TP}$ T [$_{VP}$ taken [$_{DP_2}$ a picture of t_1] by John]]]
   (iii) [$_{CP}$ who$_1$ was [$_{TP}$ [$_{DP_2}$ a picture of t_1] T [$_{VP}$ taken t_2 by John]]

Analysis (Chomskys (1995)):
"Passive [i.e., DP raising] is the same in both [derivations]; *wh*-movement is 'longer' in the illicit one in an obvious sense, object being more remote from SpecC than subject in terms of number of XPs crossed. The distinction might be captured by a proper theory of economy of derivation." D_2 in (41-b), which does not violate a local constraint (the SCC, by assumption, being irrelevant), is blocked as ungrammatical via Shortest Paths by the more economical D_1 in (41-a), which does violate a local constraint (the CED) but converges.

5.3 Merge before Move

Note:
Given that Merge and Move alternate in the derivation, the situation can arise that it must be decided whether the next step is a Merge or a Move operation. The following transderivational condition settles the issue by preferring Merge to Move if both are possible as such; the specific formulation is based on Frampton & Gutman (1999).

(42) *Merge before Move^{td}*:
Suppose that two derivations D_1 and D_2 are in the same reference set

and respect all local constraints, and $D_1 = <\Sigma_0, ..., \Sigma_n, \Sigma_{n+1}, ...\Sigma_k>$ and $D_2 = <\Sigma_0, ..., \Sigma_n, \Sigma'_{n+1}, ...\Sigma'_{k'}>$. Then D_1 is to be preferred over D_2 if $\Sigma_n \rightarrow \Sigma_{n+1}$ is an instance of Merge and $\Sigma_n \rightarrow \Sigma'_{n+1}$ is an instance of Move.

(43) *Expletive constructions in English*:

a. There$_1$ seems [$_{TP}$ t$_1$ to be [$_{PP}$ someone$_2$ in the room]]

b. *There$_1$ seems [$_{TP}$ someone$_2$ to be [$_{PP}$ t$_2$ in the room]]

Assumption:
Non-finite T in a control infinitive has a [*D*] feature (as does the finite T in the matrix clause).

Note:
The lexical array is identical in (43-a) and (43-b). The transderivational constraint Merge before Move ensures that D_1 generating (43-a) is to be preferred over D_2 generating (43-b).

Question:
Why is (44) possible, which does not have expletive Merge at all?

(44) *Successive-cyclic movement to SpecT*:
Someone$_1$ seems [$_{TP}$ t$'_1$ to be t$_1$ in the room]

Answer:
Reference sets must be assumed to be defined via identical lexical arrays. The derivation that generates (44) has a lexical array that is different from the one underlying the derivations in (43).

Conceptual problem:
Why should it be that Merge is preferred over Move?

Solution:
Move is defined on the basis of Move, which makes it inherently more complex. Chomsky (2000): "Good design conditions would lead us to expect that simpler operations are preferred to more complex ones, so that Merge ... preempt[s] Move, which is a 'last resort,' chosen when nothing else is possible."

5.4 Conclusion

Note:
There are more transderivational (or translocal) constraints that have been suggested in the minimalist literature. Among them we find:
(i) *Procrastinate* (transderivational, Chomsky (1993; 1995));
(ii) *Economy of Representation* (translocal, Chomsky (1991));
(iii) *Preference Principle for Reconstruction* (transderivational, Chomsky (1993)).

A meta-theoretical consideration:
Grammars that employ transderivational (or translocal) constraints in addition to local (or local and global) constraints all have the format that Prince & Smolensky (1993) attribute to an optimality-theoretic grammar: A first generator part (called *Gen*) creates the *candidate set* (= reference set, in minimalist syntax); Gen has only local (or global) constraints. A second harmony-evaluation part (called *H-Eval*) determines the optimal candidate(s) (= derivation(s), in minimalist syntax) in the candidate set. More generally, all kinds of competition-based syntax have this structure, which is schematically shown in (45).

(45) *Structure of a competition-based syntax*:
  a. Gen creates the candidate set $\{C_1, C_2, ...\}$.
  b. H-Eval determines the optimal candidate(s) C_i (C_j, ...) in $\{C_1, C_2, ...\}$.

Question:
Does optimality equal grammaticality in the minimalist program?

Answer:
Not necessarily: Given the notion of convergence, an optimal candidate may be one that violates a local (or global) constraint, and is therefore ungrammatical.

Problem:
How does optimality evaluation proceed in the presence of more than one transderivational constraint in the grammar? In this case, conflicts may arise. As a simple, abstract example, suppose that there are two transderivational constraints (T_1 and T_2), and only three derivations (D_1, D_2, and D_3) in the reference set. Suppose further that T_1 prefers D_1 over D_2 and D_3; that T_2 prefers D_2 over D_1 and D_3; and that a derivation D_0 that

would be preferred by both T_1 and T_2 fails to converge, so that it cannot participate in the competition. In such a situation, various possibilities arise:
(i) *Tolerance*:
It suffices to be selected by one transderivational constraint to be optimal; consequently, both D_1 and D_2 would be classified as optimal.
(ii) *Ranking*:
The conflict among transderivational constraints is resolved by a ranking, such that that derivation is optimal that is preferred by the higher-ranked constraint in the case of conflict. If T_1 is ranked higher than T_2, this would imply that only D_1 is optimal.
(iii) *Breakdown*:
In the case of conflicting instructions made by transderivational constraints, no derivation can emerge as optimal.

Note:
Option (iii) is generally adopted in the literature (see Collins (1994), Sternefeld (1997), and Müller (2000)).

(46) *Grammaticality in minimalist syntax*:
A derivation D_i is grammatical iff (a) and (b) hold:
   a. D_i does not violate a local constraint.
   b. D_i is optimal.

(47) *Optimality in minimalist syntax*:
A derivation D_i is optimal iff there is no derivation D_j in the same reference set that is preferred over D_i by a transderivational or translocal constraint.

Overall conclusion:
Transderivational constraints can offer insightful analyses, but they are problematic from the point of view of complexity (competition in potentially infinite candidate/reference sets). Hence: There is a strong tendency in recent versions of the minimalist program to dispense with transderivational constraints – and hence, with the concept of competition – altogether; see in particular Collins (1997), Frampton & Gutman (1999), and also Chomsky (1995; 2000; 2001). However, the transderivational Merge before Move constraint is still widely adopted (see, e.g., Chomsky (2000; 2001)).

A historical note:
It is interesting to note that the fall of transderivational constraints (and with

it the fall of the concept of competition) in minimalist syntax goes hand in hand with the rise of optimality theory, and hence optimality-theoretic syntax, which inherently relies on transderivationality and competition.

Chapter 6

Optimality-Theoretic Syntax

6.1 Basic Concepts

Note:
By definition, an optimality-theoretic syntax takes the general form in (1), with the grammar divided into a Gen part that creates the competing candidates, and a H-Eval part that selects the optimal candidate(s).

(1) *Structure of a competition-based syntax*:
  a. Gen creates the candidate set $\{C_1, C_2, ...\}$.
  b. H-Eval determines the optimal candidate(s) C_i (C_j, ...) in $\{C_1, C_2, ...\}$.

Recall:
The notion of optimality in a minimalist syntax is a comparatively simple one: Optimality is determined by a small set of simple transderivational or translocal economy constraints, as in (2).

(2) *Optimality in minimalist syntax*:
A derivation D_i is optimal iff there is no derivation D_j in the same reference set that is preferred over D_i by a transderivational or translocal constraint.

But:
In optimality-theoretic syntax, there is only one translocal or transderivational constraint that determines optimality: Optimal (and grammatical) is a candidate that has the best constraint profile – or, more precisely, a candidate for which there is no competitor that has a better constraint profile; cf. (3).
(Whether this optimality principle is transderivational or translocal depends

on whether the competing candidates are assumed to be derivations or output representations.)

(3) *Optimality in optimality-theoretic syntax*:
A candidate C_i is optimal (= grammatical) iff there is no candidate C_j in the same candidate set that has a better constraint profile.

(4) *Constraint Profile*:
A candidate C_j has a better constraint profile than a candidate C_i iff there is a constraint Con_k such that (a) and (b) hold:
a. C_j satisfies Con_k better than C_i.
b. There is no constraint Con_l ranked higher than Con_k on which C_i and C_j differ.

Note:
C_j satisfies a constraint Con better than C_i iff C_j has fewer violations of Con. This implies the case that C_i violates Con once (or more often), and C_j does not violate Con at all.

Note:
This presupposes that in addition to the local (or global) constraints employed by the Gen component, which are inviolable and unranked, the H-Eval component relies on a system of local (or global) constraints that are violable and ranked (and, by assumption, universal) in order to determine the best constraint profile, hence, optimality. The ranking among the violable local constraints of the H-Eval component is indicated by the symbol $\gg$; the H-Eval constraints themselves are typically written with small capitals. Optimality-theoretic competitions are often illustrated by tables (so-called *tableaux*); optimality of a candidate is indicated by the *pointing finger*: ☞; violation of a local constraint is shown by a star * in the appropriate column of the table; if this violation is fatal for a candidate (i.e., responsible for its suboptimality), an exclamation mark ! is added. In the abstract H-Eval competition in table $T_{6.1}$, in which the candidate set consists of C_1–C_5, C_1 emerges as the optimal candidate: It avoids a violation of the high-ranked constraints A and B (unlike C_3–C_5), and it minimizes a violation of the low-ranked constraint C (unlike C_2). Hence, there is no competing candidate with a better constraint profile than C_1.
Constraint reranking = parametrization:
By reranking the constraints B and C in $T_{6.1}$, candidate C_3 would emerge as the optimal candidate. Reranking of constraints forms the basis of the

$T_{6.1}$: Determining optimality

Candidates	A	B	C
☞C_1			*
C_2			**!
C_3		*!	
C_4	*!		
C_5		*!	*

concept of parametrization in optimality-theoretic syntax.

Non-cumulativity:
A further characteristic feature of this approach is that it is essentially non-cumulative; i.e., no number of violations of a low-ranked constraint can outweigh a single violation of a higher-ranked constraint. Thus, suppose that there were an additional, lowest-ranked constraint D in $T_{6.1}$ that C_1 violates, say, five times, and that C_2–C_5 do not violate at all. This would not undermine C_1's optimality.

Candidates and violable constraints:
Something needs to be said about the nature of candidates and candidate sets. Optimality-theoretic syntax is strongly influenced by work in optimality-theoretic phonology. Since the latter is characterized by an orientation that is predominantly representational (cf. Prince & Smolensky (1993) and McCarthy & Prince (1995)), it does not come as a surprise that many approaches in optimality-theoretic syntax postulate that the competing candidates created by Gen are output representations. This holds, e.g., for what can arguably be viewed as the three most influential analyses in optimality-theoretic syntax so far (outside the LFG work by Bresnan, Choi, Sells, and others), viz., Grimshaw (1997), Pesetsky (1998), and Legendre, Smolensky & Wilson (1998). However, there is no inherent reason why the candidates that are subject to optimization should not be syntactic objects of a more complex type, like <D-structure,S-structure,LF> tuples as in the Principles-and-Parameters approach, or, indeed, complete derivations, as in the minimalist program. The choice of candidate type goes hand in hand with the choice of local (or global) constraint type that shows up in the H-Eval part as a violable and ranked: If candidates are representations, constraints will be representational, if candidates are derivations, constraints will be derivational (or global), and if candidates are <D-structure,S-structure,LF> tuples, constraints can take any of the

forms sketched in [2].

Candidate sets:
Similarly, candidate sets can be defined in various ways, which of course significantly influences the nature of the competition. Basically, all of the definitions of reference sets in minimalist syntax that have been proposed (see [5], and Sternefeld (1997) for an overview) are also potential definitions of candidate sets in optimality-theoretic syntax. A further influential definition of candidate sets comes from Grimshaw (1997). She postulates that two candidates (S-structure representations) compete iff they are realizations of the same predicate/argument structure and have non-distinct logical forms (or non-distinct interpretation).

Note:
By making optimality depend on an intricate system of violable and ranked constraints, H-Eval – and hence, the concept of competition – becomes even more important than in minimalist syntax and blocking syntax. As a matter of fact, much work in optimality-theoretic syntax tries to minimize the role of the Gen component, and maximize the role of the H-Eval component (but see Pesetsky (1997; 1998) for some cautionary remarks).

What is optimality theorety good for in syntax?
An optimality-theoretic approach gains immediate support in all those contexts where postulating a competition of syntactic objects is initially plausible. This includes, but is by no means confined to, contexts where notions of economy seem to play a role. A prototypical case is one in which the wellformedness of a sentence S_i that exhibits an otherwise peculiar property seems to depend on the unavailability of another sentence S_j that exhibits the property one would normally expect. Here, S_i is often referred to as a "repair" or "last resort" form; a typical instance is the English *do*-support construction. Accordingly, *do*-support was among the first phenomena to be tackled in optimality-theoretic syntax (see Speas (1995) and Grimshaw (1997)). Most of the constructions discussed in [5] can also be viewed as suggesting an underlying competition; and indeed, they can fruitfully be addressed in optimality-theoretic syntax.

6.2 Anaphors vs. Pronouns

(5) *Consequences of Principle A*:

a. $[_{CP}$ C $[_{TP}$ $John_1$ $[_{T'}$ Ø $[_{VP}$ t_1 likes $himself_1$]]]]
b. *$[_{CP}$ C $[_{TP}$ $John_1$ $[_{T'}$ Ø $[_{VP}$ t_1 thinks $[_{CP}$ that $[_{TP}$ $Mary_2$ $[_{T'}$ Ø $[_{VP}$ t_2 likes $himself_1$]]]]]]]]

(6) *Consequences of Principle B*:
a. *$[_{CP}$ C $[_{TP}$ $John_1$ $[_{T'}$ Ø $[_{VP}$ t_1 likes him_1]]]]
b. $[_{CP}$ C $[_{TP}$ $John_1$ $[_{T'}$ Ø $[_{VP}$ t_1 thinks $[_{CP}$ that $[_{TP}$ $Mary_2$ $[_{T'}$ Ø $[_{VP}$ t_2 likes him_1]]]]]]]]

Generalization:
By and large, pronouns are allowed to express binding relations in English in just those cases in which anaphors are not allowed to do so. This has been taken to indicate that there is a competition of the two strategies: We do not have to stipulate both Principle A and Principle B. Rather, one of the two constraints (usually Principle A) is adopted, and the other is derived as an elsewhere case by invoking syntactic competition (see Fanselow (1989; 1991), Burzio (1991; 1998), Reinhart (1991), and Richards (1997), among others). The following analysis essentially goes back to Wilson (2001).

(7) a. LOC-ANT^r ("Local Antecedent"):
If a binding domain contains an anaphor, then it must also contain the anaphor's antecedent.
b. REF-ECON^r ("Referential Economy"):
An argument must not have lexical ϕ-feature specification.

(8) *Ranking in English*:
LOC-ANT $\gg$ REF-ECON.

Note:
(i) LOC-ANT is a version of Principle A. Other things being equal, this constraint favours pronouns.
(ii) REF-ECON inherently prefers anaphors to pronouns, assuming that anaphors do not have a lexical ϕ-feature specification, whereas pronouns do.

$T_{6.2}$: Reflexivization

Candidates	LOC-ANT	REF-ECON
☞C_1: $John_1$ likes $himself_1$		
C_2: $John_1$ likes him_1		*!

$T_{6.3}$: *Pronominalization*

Candidates	LOC-ANT	REF-ECON
C_1: $John_1$ thinks that Mary likes $himself_1$	*!	
☞C_2: $John_1$ thinks that Mary likes him_1		*

6.3 Complementizer-Trace Effects

Note:
Recall from [3] that the Principles-and-Parameters approach accounts for complementizer-trace effects (= that-trace effects) on a purely local basis, without postulating a competition with the complementizer-less variant from which only the latter would emerge as optimal.

(9) *The ECP account of that-trace effects*:

a. *Who_1 do you think [$_{CP}$ t'_1([$+\gamma$]) that [$_{TP}$ t_1([$-\gamma$]) will leave]] ?

b. Who_1 do you think [$_{CP}$ t'_1([$+\gamma$]) Ø [$_{TP}$ t_1([$+\gamma$]) will leave]] ?

Background assumptions:
This view is abandoned in Déprez (1991), which is the basis of the optimality-theoretic account advanced in Grimshaw (1997). As background, Grimshaw assumes that the size of clauses is variable. Clauses are extended projections of V; they are minimally VPs, but they can be TPs, CPs, or functional projections of an even bigger size, depending on the outcome of optimization. Bridge verbs in English permit both CP-embedding (with a complementizer – a declarative CP without a complementizer will typically fatally violate a high-ranked constraint that precludes empty head positions) and TP- or VP-embedding (without a complementizer). In the latter case, TP must be chosen if an auxiliary or *do* is present (i.e., if the need arises to accomodate an additional lexical head); VP otherwise. A co-occurrence of CP- and TP- (VP-) embedding is possible if the two candidates have an identical constraint profile. This implies that the presence of *that* does not violate any constraint, an assumption that Grimshaw (1997) makes even though it is not completely unproblematic (as she acknowledges). The main constraints that are needed in the account of complementizer-trace effects are listed in (10).

(10) a. OP-SPEC^r ("Operator in Specifier"):
Wh-operators must occupy a specifier position from which they c-command all elements of the extended V projection over which they take scope.

b. T-LEX-GOV^r ("Lexical Government of Traces"):
A trace is lexically governed.

c. STAY^r ("Economy of Movement"):
Trace is not allowed.

(11) *Ranking in English*:
OP-SPEC ≫ T-LEX-GOV ≫ STAY.

Note:
STAY is a local version of the translocal economy constraint Fewest Steps; OP-SPEC is a version of the *Wh*-Criterion that has often been postulated in the Principles-and-Parameters approach (see, e.g., Lasnik & Saito (1992)). (However, we have seen that the effects of such a constraint can be derived from postulating [*Q*] features on $C_{[+wh]}$ items.)

Assumption:
Candidates with and without *that* compete.

$T_{6.4}$: *Subject wh-movement*

Candidates	OP-SPEC	T-LEX-GOV	STAY
C_1: ... who$_1$ you think [$_{CP}$ that [$_{TP}$ t$_1$ will leave]]			*
☞C_2: ... who$_1$ you think [$_{TP}$ t$_1$ will leave]		*!	*
C_3: ... you think [$_{CP}$ that [$_{TP}$ who$_1$ will leave]]	*!		
C_4: ... you think [$_{TP}$ who$_1$ will leave]	*!		

Analysis:
C_3 and C_4 fatally violate OP-SPEC. Both C_1 and C_2 violate STAY, but C_2 violates T-LEX-GOV in addition: t_1 in C_2 is not lexically governed (*that* being unable to do so), whereas t_1 in C_1 is lexically governed (by the matrix V). In contrast, an embedded V governs object traces throughout, irrespective of the presence or absence of a complementizer *that*; hence, T-LEX-GOV is satisfied equally well by C_1 and C_2 in table $T_{6.5}$. Given that C_1 and C_2 do not differ with respect to any other constraint either, optionality of a complementizer is correctly predicted in cases of object extraction, due to an identical constraint profile.
Observation:
Thus far, there is no evidence for treating T-LEX-GOV as a violable constraint in the H-Eval part of the grammar (rather than as an inviolable constraint in the Gen part). Such evidence can be gained by considering adjunct extraction. In this case, T-LEX-GOV is violated by both candidates involving *wh*-movment (C_1, C_2). However, given that there is no competing candidate that can avoid a violation of T-LEX-GOV without violating a

$T_{6.5}$: Object wh-movement

Candidates	OP-SPEC	T-LEX-GOV	STAY
☞C_1: ... who_1 you think [$_{CP}$ that [$_{TP}$ she will invite t_1]]			*
☞C_2: ... who_1 you think [$_{TP}$ she will invite t_1]			*
C_3: ... you think [$_{CP}$ that [$_{TP}$ she will invite who_1]]	*!		
C_4: ... you think [$_{TP}$ she will invite who_1]	*!		

higher-ranked constraint (e.g., a candidate that employs a resumptive pronoun; see below), C_1 and C_2 can emerge as optimal despite this violation.

$T_{6.6}$: Adjunct wh-movement

Candidates	OP-SPEC	T-LEX-GOV	STAY
☞C_1: ... why_1 you think [$_{CP}$ that [$_{TP}$ she has left t_1]]		*	*
☞C_2: ... why_1 you think [$_{TP}$ she has left t_1]		*	*
C_3: ... you think [$_{CP}$ that [$_{TP}$ she has left why_1]]	*!		
C_4: ... you think [$_{TP}$ she has left why_1]	*!		

6.4 Subjacency and Resumptive Pronouns

(12) *Resumptive pronouns in Complex NP Constraint contexts*:

a. the man [$_{CP}$ $who(m)_1$ I saw t_1]

b. *the man [$_{CP}$ $who(m)_1$ I don't believe [$_{DP}$ the claim [$_{CP}$ t'_1 that anyone saw t_1]]]

c. *the man [$_{CP}$ $who(m)_1$ I saw him_1]

d. ?the man [$_{CP}$ $who(m)_1$ I don't believe [$_{DP}$ the claim [$_{CP}$ t'_1 that anyone saw him_1]]]

Note:
Resumptive pronouns often seem to be possible only as last resort strategies, in cases where traces are blocked. Competition-free approaches to syntax have no obvious means to relate one construction to the other; but the case is different in optimality-theoretic syntax. An optimality-theoretic account of resumptive pronoun strategies is developed in Legendre, Wilson & Smolensky (1998) (on the basis of evidence from Chinese) and Pesetsky (1998) (on the basis of English data comparable to those in (12), as well as evidence from Hebrew, Russian, and Polish). (A similar last resort analysis is given in Hornstein (2000) within the minimalist program.) The details of the analyses

differ a lot, but the gist of the explanation is identical; it centers around two constraints like those in (13).

(13) a. CNPC^r ("Complex NP Constraint"):
*... α_1 ... [$_{DP}$ ... [$_{CP}$ ... t_1 ...]] ...
b. RES^r ("Resumptive Pronoun Constraint"):
Resumptive pronouns are prohibited.

(14) *Ranking in English*:
CNPC $\gg$ RES.

Note:
The CNPC prohibits traces in certain (non-local) environments; RES disfavours resumptive pronouns (i.e., pronouns that have c-commanding co-indexed antecedent in an A-bar position) in general. As with the *wh*-movement construction discussed in the last section, it must be ensured that overt movement of the relative operator takes place in examples like those in (12). Suppose that this is independently taken care of – e.g., by a high-ranked (or inviolable Gen-) constraint demanding deletion of [*rel*].

$T_{6.7}$: Trace vs. resumptive pronoun in transparent contexts

Candidates	CNPC	RES
☞C_1: the man [$_{CP}$ who(m)$_1$ I saw t_1]		
C_2: the man [$_{CP}$ who(m)$_1$ I saw him$_1$]		*!

Analysis:
Both candidates respect CNPC. Consequently, the RES violation incurred by the resumptive pronoun in C_2 becomes fatal, and C_1 is optimal. However, in the competition illustrated in table $T_{6.8}$, C_1 violates the CNPC. In this case, C_2's RES violation is tolerable, and the resumptive pronoun strategy emerges as optimal.

$T_{6.8}$: Trace vs. resumptive pronoun in CNPC contexts

Candidates	CNPC	RES
C_1: the man [$_{CP}$ who(m)$_1$ [$_{TP}$ I don't believe [$_{DP}$ the claim [$_{CP}$ t'_1 that anyone saw t_1]]]]	*!	
☞C_2: the man [$_{CP}$ who(m)$_1$ [$_{TP}$ I don't believe [$_{DP}$ the claim [$_{CP}$ that anyone saw him$_1$]]]]		*

6.5 Avoid Pronoun

Recall:
As noted in [3], PRO may block overt pronouns in English gerunds. PRO and a lexical pronoun can both occur in principle; however, PRO must be used instead of a lexical pronoun if it can fulfill the Constraint on Control. This was derived by adopting a transderivational/translocal Avoid Pronoun constraint in Chomsky (1981).

(15) *PRO in English gerunds*:
  a. $John_1$ would much prefer [PRO_1 going to the movie]
  b. *$John_1$ would much prefer [$PRO_{2/arb}$ going to the movie]

(16) *Pronouns in English gerunds*:
  a. *$John_1$ would much prefer [his_1 going to the movie]
  b. $John_1$ would much prefer [his_2 going to the movie]

(17) CONTROL^r ("Constraint on Control", Manzini (1983)):
If PRO is minimally dominated by a declarative clausal complement α, then it must be controlled by an antecedent within the minimal CP that dominates α.

(18) *Avoid Pronoun*$^{td/tl}$:
If two derivations D_1 and D_2 are in the same reference set and D_1 uses a lexical pronoun where D_2 uses an empty pronoun, then D_1 is to be preferred over D_2.

Note:
Chomsky's (1981) approach can straightforwardly be translated into optimality theory:
(i) The Constraint on Control in (17) can directly be viewed as an optimality-theoretic constraint.
(ii) The Avoid Pronoun constraint in (18) can be simplified by turning this transderivational/translocal constraint into a local (though violable) one; cf. (19).

(19) *PRON^r ("Avoid Pronoun"):
Pronouns are prohibited.

(20) *Ranking in English*:
CONTROL $\gg$ *PRON.

Candidate sets:
Suppose that candidate sets are defined in such a way that candidates with PRO and candidates with a lexical pronoun can compete, but, crucially, that sentences with different indexings (hence, different logical forms) do not compete. Then, the facts fall into place. The blocking of a lexical pronoun by PRO in cases where CONTROL can be satisfied is illustrated in table $T_{6.9}$.

$T_{6.9}$: PRO vs. pronoun under co-indexing

Candidates	CONTROL	*PRON
C_1: $John_1$ would much prefer [his_1 going to the movie]		*!
☞C_2: $John_1$ would much prefer [PRO_1 going to the movie]		

But:
Table $T_{6.10}$ illustrates the case where PRO is not co-indexed with the matrix antecedent, thereby violating CONTROL. Here, the *PRON violation incurred by all pronouns is non-fatal, and the pronoun strategy is optimal.

$T_{6.10}$: PRO vs. pronoun under contra-indexing

Candidates	CONTROL	*PRON
☞C_1: $John_1$ would much prefer [his_2 going to the movie]		*
C_2: $John_1$ would much prefer [PRO_2 going to the movie]	*!	

6.6 Open Issues

Note:
Optimality-theoretic syntax inherits the *complexity* problem from minimalist syntax. Candidate sets are typically large (as in Pesetsky (1998)), often infinite (as in Grimshaw (1997)). In addition, there are several open issues that are specific to the optimality-theoretic approach. The focus is on two of these in what follows: inputs and absolute ungrammaticality.

6.6.1 Inputs and Faithfulness

The input and its functions:
An important optimality-theoretic concept that has played no role so far is the notion of *input*. In optimality theory (cf. Prince & Smolensky (1993)), Gen does not create competing candidates (the *outputs*) freely; rather, it does so on the basis of a given input. In phonology, inputs are underlying representations stored in the lexicon; here, inputs qualify as roughly the same

types of objects as outputs. In syntax, it is much less clear what the input might look like (see Archangeli & Langendoen (1996)). The null hypothesis – that the input is a completely articulated potential sentence of the same type as the output candidates – is not unproblematic because it would seem to imply the assumption that all possible sentences are "stored," which cannot possibly be true. To find out what the input in syntax is, it is instructive to consider its theory-internal functions. By and large, there are two. First, the input is standardly taken to define the competition, i.e., candidate sets. Second, the input serves as a basis for *faithfulness* constraints that demand input/output identity and thereby minimize deviations from the input in the optimal output. Let us consider the second function first. Faithfulness constraints can be viewed as global constraints because they cannot be checked by considering solely an output representation or a derivational step.

Faithfulness constraints:
Faithfulness constraints play an important role in phonology. Constraints of the PARSE (or MAX) family prohibit deletion of input material in the output; constraints of the FILL (or DEP) family prohibit insertion of output material that is not part of the input; constraints of the IDENT family prohibit modifying input material. Faithfulness constraints are also adopted in much recent work in optimality-theoretic syntax. The following two constraints are taken from Legendre, Smolensky & Wilson (1998) and Baković & Keer (2001), respectively.

(21) *Syntactic faithfulness constraints*:
   a. PARSE[SCOPE]^g:
      Scope assignment in the input must be realized by chain formation in the output.
   b. FAITH[COMP]^g:
      The output value of [±COMP] is the same as the input value.

Note:
(21-a) implies that the input is a more complex object than just a lexical array or a predicate/argument structure; it must be a highly structured representation that exhibits the relative scope of operators. (21-b) presupposes an abstract feature [±COMP] that for present purposes we can assume to be located on a V that selects a proposition. Let us consider candidates that violate these constraints. Suppose that (22-a) is the input for output candidate (22-b), and (22-c) is the input for output candidate (22-d). Legendre, Smolensky & Wilson (1998) assume that (22-b) violates PARSE[SCOPE]

because matrix scope for *how*$_1$ in the input (22-a) (indicated by [+wh]$_1$) is reduced to embedded scope in the output (again indicated by [+wh]$_1$). Similarly, Baković & Keer (2001) assume that (22-d) violates FAITH[COMP] because a [–COMP] specification in the input contrasts with a [+COMP] specification (hence, a complementizer) in the output.

(22) *Violations of syntactic faithfulness constraints*:

a. [+wh]$_1$ ... wonder$_{[+wh]}$ [[+wh]$_2$ ... what$_2$ ... how$_1$ ...] (input)
b. You wonder$_{[+wh]}$ [$_{CP}$ [+wh]$_1$ [+wh]$_2$ how$_1$ John did what$_2$] (output)
c. ... V$_{[-Comp]}$ [...] (input)
d. I think [$_{CP}$ that [$_{PP}$ on him]$_1$ no coat looks good t$_1$] (output)

A reinterpretation of faithfulness constraints:
At this point, we need not go into the actual analyses in which these constraints play a role (as it happens, both faithfulness violations turn out to be non-fatal, i.e., (22-bd) are optimal). The crucial question is: Is it really necessary to refer to the concept of input here, or is it possible to read the respective violations off the output forms, without any reference to inputs? At least for the cases at hand, the answer is straightforward: By enriching output representations in ways that have independently been proposed, a reference to inputs becomes unnecessary. (22-ab) is a case where the intended matrix scope is not reached by chain formation in the candidate. Employing abstract scope markers (Σ) in S-structure representations (cf., e.g., Williams (1986)), we can equivalently encode this input information in the output, as in (23-a). As for the case in (22-cd), the only assumption that we have to make (which is completely standard) is that selectional properties of lexical heads are accessible in syntax; cf. (23-b).

(23) *Violations of syntactic markedness constraints*:

a. Σ$_1$ you wonder$_{[+wh]}$ [$_{CP}$ [+wh]$_1$ [+wh]$_2$ how$_1$ John did what$_2$] (output)
b. I think$_{[-Comp]}$ [$_{CP}$ that [$_{PP}$ on him]$_1$ no coat looks good t$_1$] (output)

PARSE[SCOPE] and FAITH[COMP] can now be modified in obvious ways, without reference to inputs, as regular local (so-called *markedness*) constraints.

(24) *Faithfulness constraints as markedness constraints*:

a. PARSE[SCOPE]^r (revised):
Scope markers must be reached by chain formation.

b. FAITH[COMP]^r (revised):
Lexical [±COMP] selection requirements must be respected.

Note:
If this result can be generalized, and all syntactic faithfulness constraints can be reanalyzed in this way, we can conclude that these constraints do not support the concept of input anymore. Why should it be that the notion of input is relevant for phonological faithfulness constraints, but not for their syntactic counterparts? The answer, we believe, follows from what appears to be a fundamental difference between syntax and phonology: Syntax is an *information-preserving system* with richly structured output candidates, whereas phonology is a system that is standardly taken to lose information in the course of a derivation (e.g., from input to output), so that reference to an underlying input is necessary in constraints.

Candidate Sets:
With this in mind, consider the other input function noted above, that of defining candidate sets. Since syntactic output candidates are richly structured, all the relevant information that they must share in order to compete can be read off them, independently of what notion of candidate set is adopted; again, this is in sharp distinction to phonology. Hence, reference to inputs seems unnecessary for this purpose in syntax.

Conclusion:
From all this, it seems that one can conclude that it may eventually be possible to dispense with the notion of input in syntax; but further research is needed in this domain (see Heck et al. (2002)).

6.6.2 Absolute Ungrammaticality

Observation:
Another important open question in optimality-theoretic syntax is how to account for the phenomenon of ***absolute ungrammaticality*** or ***ineffability***, i.e., cases where there does not seem to be a candidate in a candidate set that is grammatical. As an example, consider the following ungrammatical example involving *wh*-extraction across an adjunct island in German:

(25) *A consequence of the Adjunct Condition*:

*Was$_1$ ist Fritz eingeschlafen [$_{CP}$ nachdem er t$_1$ gelesen hat] ?
what is Fritz fallen asleep after he read has

(26) *Adjunct Condition*^{d}:
Movement must not take place from an XP that has been merged without a deletion of selectional features.

(27) ADJ-CON^r ("Adjunct Condition", representational version):
A trace must not be separated by a non-selected XP from its antecedent.

Approach no. 1: Gen:
Pesetsky (1997; 1998) emphasizes that certain sentences may be ungrammatical not because they are classified as suboptimal in the H-Eval part of the grammar, but because they cannot be generated by Gen in the first place. Thus, a constraint like (27) might be part of Gen.

Approach no. 2: semantic illformedness:
It is suggested in Grimshaw (1994) and Müller (1997) that certain optimal candidates may have properties that make them inaccessible for other domains of the language faculty like, e.g., semantic interpretation. ADJ-CON might be part of H-Eval, but ranked higher than OP-SPEC. On this view, (28) could block (27) as suboptimal; but this optimal candidate would be uninterpretable (indicated by #) and, hence, unusable.

(28) *Semantic illformedness*:

#Fritz ist eingeschlafen [$_{CP}$ nachdem er was$_1$ gelesen hat] ?
Fritz is fallen asleep after he what read has

Note:
These two approaches have in common that they allow the possibility that absolute ungrammaticality is not located in the H-Eval component of grammar; but in a component that precedes (Gen) or follows (interpretation) optimization. If, however, H-Eval is to be held responsible for the ungrammaticality of (25), there must be a competing candidate with a better constraint profile that blocks it.

Approach no. 3: last resort (fails):
A priori, this might be a candidate that employs a resumptive pronoun strategy which is only legitimate in this context, as a last resort. If this were so, the ineffability problem would be spurious in the case at hand. However, (29) shows that the resumptive pronoun strategy is not an option in German (a constraint like RES must outrank ADJ-CON and other locality constraints in German):

(29) *Last resort*:

*Was₁ ist Fritz eingeschlafen [$_{CP}$ nachdem er es$_1$ gelesen hat] ?
what is Fritz fallen asleep after he it read has

Approach no. 4: null parse:
What, then, could the optimal candidate blocking (25) look like? Following Prince & Smolensky (1993), Ackema & Neeleman (1998) propose that the empty candidate Ø (the "null parse") is part of every candidate set. This candidate violates the constraint in (30), which is typically ranked high.

(30) *Ø^r ("Avoid Null Parse"):
Ø is prohibited.

Note:
Constraints that are ranked higher than *Ø in effect become inviolable (given that there is no other constraint except *Ø that Ø can violate). In this sense, *Ø introduces a dividing line into rankings. Thus, if both ADJ-CON and the constraint that triggers *wh*-movement (e.g., OP-SPEC) outrank *Ø, adjunct islands become inviolable. This is shown in table $T_{6.11}$.

$T_{6.11}$: Adjunct islands and the null parse

Candidates	ADJ-CON	OP-SPEC	*Ø
C_1: was$_1$ ... [$_{CP}$ nachdem er t_1 V]	*!		
C_2: – ... [$_{CP}$ nachdem er was$_1$ V]		*!	
☞C_3: Ø			*

Approach no. 5: neutralization:
A final possibility to be discussed here is the *neutralization* approach to absolute ungrammaticality in syntax. Such an approach has been adopted by Legendre, Smolensky & Wilson (1998), Schmid (1998), Baković & Keer (2001), and Wilson (2001), among others. For the present case, a neutralization analysis might posit that the optimal candidate blocking (25) is (31).

(31) *Neutralization*:

Fritz ist eingeschlafen [$_{CP}$ nachdem er was$_1$ gelesen hat]
Fritz is fallen asleep after he something read has

Note:
The crucial difference to (28) is that *was*$_1$ is turned into an indefinite pronoun, and the matrix $C_{[+wh]}$ is turned into a $C_{[-wh]}$. Thus, there is a feature change from [+wh] in (25) to [–wh] in (31), and the sentence is interpreted as

declarative, rather than a as question. If (31) is to block (25) as suboptimal, this presupposes that candidates that differ in their *wh*-feature specification can compete. But then, the problem arises that we would also wrongly expect one of the sentences in (32) to block the other.

(32) *A problem?*

a. Was_1 hat er t_1 gelesen ?
   what has he read

b. Er hat was_1 gelesen
   he has something read

The solution:
The neutralization approach solves this problem as follows. The [±wh]-specification is unambiguously specified in the input; an input with a [+wh] specification on some item and a minimally different input with a [–wh] specification count as different, and define different candidate sets. The important assumption is that there is a faithfulness constraint that demands preservation of the [±wh] feature specification in the output:

(33) FAITH[WH]^g:
The output value of [±wh] is the same as the input value.

Analysis:
Suppose now that ADJ-CON and OP-SPEC are ranked higher than FAITH[WH]. Then, (31) will have a better constraint profile than (25) both in the competition that has a [–wh] specification in the input, and in the competition that has a [+wh] specification in the input. Thus, there is a "neutralization" of different input specifications in the output. This is shown in tables $T_{6.12}$ and $T_{6.13}$.

$T_{6.12}$: Adjunct islands and neutralization; [–wh] in the input

Candidates	ADJ-CON	OP-SPEC	FAITH[WH]
C_1: $was_{1[+w]}$ ... [$_{CP}$ nachdem er t_1 V]	*!		*
C_2: – ... [$_{CP}$ nachdem er $was_{1[+w]}$ V]		*!	*
☞C_3: – ... [$_{CP}$ nachdem er $was_{1[-w]}$ V]			

Note:
In transparent contexts, where movement may occur without a violation of a high-ranked locality constraint like ADJ-CON (cf. (32)), FAITH[WH] violations become fatal, and the candidate that maintains the [±wh]

$T_{6.13}$: Adjunct islands and neutralization; [+wh] in the input

Candidates	ADJ-CON	OP-SPEC	FAITH[WH]
C_1: $was_{1[+w]}$ ... [$_{CP}$ nachdem er t_1 V]	*!		
C_2: – ... [$_{CP}$ nachdem er $was_{1[+w]}$ V]		*!	
☞C_3: – ... [$_{CP}$ nachdem er $was_{1[-w]}$ V]			*

specification of the input emerges as optimal.

Conclusion:
Of the four working approaches to absolute ungrammaticality discussed here (Gen, semantic illformedness, null parse, neutralization), the neutralization approach is arguably the most elegant one. Still, it is not without problems. One conspicuous peculiarity is that neutralization creates massive derivational ambiguity. A well-formed sentence like (31) can have different "histories," being an optimal candidate in two candidate sets with different inputs. This vacuous ambiguity may be considered problematic from the point of view of language acquisition and parsing; and it can only be avoided by additional meta-optimization procedures that compare the competitions in $T_{6.12}$ and $T_{6.13}$; cf. the notion of *input optimization* in Prince & Smolensky (1993) (called *lexicon optimization* in phonology).

6.6.3 Conclusion

Note:
This does not exhaust the list of open issues that are currently under debate in optimality-theoretic syntax. Here are some others:
(i) Optionality
(ii) Degrees of Grammaticality
(iii) Cumulativity
(iv) Parametrization
(v) Serial vs. parallel optimization

Outlook:
Do optimality-theoretic and minimalist approaches differ? Yes, they do, but not as much as one might think at first sight, and in comparing these two approaches, it is important to identify issues that are orthogonal (e.g., derivational vs. representational constraints/candidates, and the issue of transderivational/translocal constraints).

Chapter 7

Formal Grammars

Lit.:
Partee et al. (1993)

(1) *Grammatik*:
Eine Grammatik ist ein Quadrupel $<V_T, V_N, S, R>$, wobei gilt:
a. V_T = Vokabular (Alphabet) der terminalen Symbole;
b. V_N = Vokabular (Alphabet) der nicht-terminalen Symbole (V_T und V_N sind disjunkt);
c. S = Startsymbol;
d. R = endliche Menge von Regeln der Form $\psi \rightarrow \omega$, wobei ψ und ω Ketten sind.
Interpretation der Regeln: Wenn ψ irgendwo als Teilkette auftritt, kann es durch ω ersetzt werden und so eine neue Kette erzeugen.

(2) *Sprache* (Chomsky (1957, 13)):
Eine Sprache ist eine (potentiell infinite) Menge von Ketten von terminalen Symbolen (= Sätzen), die durch eine Grammatik erzeugt werden.

(3) *Beispielgrammatik*:
a. $V_T = \{a,b\}$
b. $V_N = \{S,A,B\}$
c. $S \in V_N$
d. $$R = \left\{ \begin{array}{ccc} S & \rightarrow & ABS \\ S & \rightarrow & e \\ AB & \rightarrow & BA \\ BA & \rightarrow & AB \\ A & \rightarrow & a \\ B & \rightarrow & b \end{array} \right\}$$

Notationskonvention:
Kleibuchstaben: terminales Alphabet; Großbuchstaben: nicht-terminales Alphabet.

$$R = \left\{ \begin{array}{ccc} S & \rightarrow & ABS \\ S & \rightarrow & e \\ AB & \rightarrow & BA \\ BA & \rightarrow & AB \\ A & \rightarrow & a \\ B & \rightarrow & b \end{array} \right\}$$

(4) *Erzeugung von 'abba'*:

a. S $\Longrightarrow$ ABS
b. ABS $\Longrightarrow$ ABABS
c. ABABS $\Longrightarrow$ ABAB
d. ABAB $\Longrightarrow$ ABBA
e. ABBA $\Longrightarrow$ ABbA
f. ABbA $\Longrightarrow$ aBbA
g. aBbA $\Longrightarrow$ abbA
h. abbA $\Longrightarrow$ abba

Bemerkung:
Diese Grammatik erzeugt die Sprache L_0.

(5) L_0: $\{x \in \{a,b\}^* \mid$ *x enthält die gleiche Anzahl von a's und b's*$\}$

Kleene-Stern:
A^* bezeichnet die Menge aller Ketten, die über dem Alphabet A gebildet werden können (der Abschluss oder 'Kleene-Stern' auf einer Menge von Ketten).

Bäume:
Syntaktische Bäume erfüllen eine Reihe von Wohlfgeformtheitsbedingungen.

(6) *Bedingung der singulären Wurzel*:
In einem wohlgeformten Strukturbaum gibt es genau einen Knoten, der jeden anderen Knoten dominiert.

(7) *Exklusivitätsbedingung*:
In einem wohlgeformten Strukturbaum gilt für alle Knoten x und y: x und y stehen in einer Präzedenzbeziehung P (d.h., entweder $<x,y> \in P$, oder $<y,x> \in P$) gdw. x und y nicht in einer Dominanzrelation D stehen (d.h., es gilt weder $<x,y> \in D$, noch $<y,x> \in D$).

(8) *Bedingung der Nicht-Verwirrung*:
In einem wohlgeformten Strukturbaum gilt für alle Knoten x und y: Wenn x y vorangeht, dann gehen alle Knoten, die von x dominiert werden, allen Knoten, die von y dominiert werden, voran.

Von Grammatiken zu Bäumen:
Die durch grammatische Regeln erzeugten Ketten korrespondieren Strukturbäumen.

(9) Eine Grammatik (mit Regeln der Art A $\rightarrow \psi$) erzeugt einen Baum gdw. wenn gilt:
   a. Die Wurzel ist mit dem Startsymbol der Grammatik gelabelt.
   b. Die Kette, die durch die gemäß der Präzedenzrelation geordneten Blätter des Baumes gebildet wird, ist eine Kette von terminalen Symbolen der Grammatik.
   c. Für jeden Teilbaum der Form $[_{\mathrm{A}}\ \alpha_1\ ...\ \alpha_n\ ]$ im Baum (wobei A α_1 ... α_n unmittelbar dominiert) gibt es eine Regel A $\rightarrow \alpha_1$ ... α_n in der Grammatik.

(10) *Beispielgrammatik*:
   a. $\mathrm{V}_T = \{\mathrm{a,b}\}$
   b. $\mathrm{V}_N = \{\mathrm{S,A,B}\}$
   c. $\mathrm{S} \in \mathrm{V}_N$
   d. $\mathrm{R} = \left\{ \begin{array}{ccc} \mathrm{S} & \rightarrow & \mathrm{AB} \\ \mathrm{A} & \rightarrow & \mathrm{aAb} \\ \mathrm{A} & \rightarrow & \mathrm{e} \\ \mathrm{B} & \rightarrow & \mathrm{Bb} \\ \mathrm{B} & \rightarrow & \mathrm{b} \end{array} \right\}$

(11) *Erzeugte Bäume*:
   a. $[_{\mathrm{S}}\ [_{\mathrm{A}}\ \mathrm{a}\ [_{\mathrm{A}}\ \mathrm{e}\ ]\ \mathrm{b}\ ]\ [_{\mathrm{B}}\ [_{\mathrm{B}}\ \mathrm{b}\ ]\ \mathrm{b}\ ]\ ]$
   b. $[_{\mathrm{S}}\ [_{\mathrm{A}}\ \mathrm{e}\ ]\ [_{\mathrm{B}}\ \mathrm{b}\ ]]$
   c. $[_{\mathrm{S}}\ [_{\mathrm{A}}\ \mathrm{e}\ ]\ [_{\mathrm{B}}\ [_{\mathrm{B}}\ \mathrm{b}\ ]\ \mathrm{b}\ ]]$
   d. $[_{\mathrm{S}}\ [_{\mathrm{A}}\ \mathrm{a}\ [_{\mathrm{A}}\ \mathrm{e}\ ]\ \mathrm{b}\ ]\ [_{\mathrm{B}}\ \mathrm{b}\ ]]$

Die Chomsky-Hierarchie

Beobachtung:
Gemäß der Art der zugelassenen Regeln unterscheiden sich Grammatiken bezüglich ihrer *generativen Kapazität* (Mächtigkeit).

(12) *Beschränkungen für Regeln*:

a. *Typ-0-Grammatiken*:
–

b. *Typ-1-Grammatiken*:
Jede Regel hat die Form $\alpha A \beta \rightarrow \alpha \psi \beta$, wobei $\psi \neq e$.

c. *Typ-2-Grammatiken*:
Jede Regel hat die Form $A \rightarrow \psi$.

d. *Typ-3-Grammatiken*:
Jede Regel hat die Form $A \rightarrow xB$ oder $A \rightarrow x$.

Es gilt:
(i) α, β, ψ sind beliebige Ketten (u.U. leere) über der Vereinigung der terminalen und nicht-terminalen Alphabete.
(ii) A, B sind nicht-terminale Symbole.
(iii) x ist eine Kette von terminalen Symbolen.

Bemerkung:
- *Typ-0-Grammatiken* $\leftrightarrow$ *unbeschränkte Ersetzungssysteme*
- *Typ-1-Grammatiken* $\leftrightarrow$ *kontextsensitive Grammatiken*
- *Typ-2-Grammatiken* $\leftrightarrow$ *kontextfreie Grammatiken*
- *Typ-3-Grammatiken* $\leftrightarrow$ *reguläre Grammatiken* (finite state grammars)

(13) *Festlegung*:
Eine Sprache ist vom Typ *n* gdw. wenn sie generiert wird von einer Grammatik vom Typ *n*.

Konsequenz:
Eine Sprache kann von mehr als einem Typ sein. L_0 z.B. kann durch eine Typ-0-Grammatik erzeugt werden (nämlich die in (3)); aber auch durch eine (kontextfreie) Typ-2-Grammatik.

(5) L_0: $\{x \in \{a,b\}^* \mid x$ *enthält die gleiche Anzahl von a's und b's*$\}$

(14) *Typ-2-Grammatik*:

a. $V_T = \{a,b\}$
b. $V_N = \{S,A,B\}$
c. $S \in V_N$
d. $R = \left\{ \begin{array}{ccc} S & \rightarrow & e \\ S & \rightarrow & aB \\ S & \rightarrow & bA \\ B & \rightarrow & b \\ B & \rightarrow & bS \\ A & \rightarrow & a \\ A & \rightarrow & aS \\ A & \rightarrow & bAA \\ B & \rightarrow & aBB \end{array} \right\}$

(3) *Typ-0-Grammatik*:

a. $V_T = \{a,b\}$
b. $V_N = \{S,A,B\}$
c. $S \in V_N$
d. $R = \left\{ \begin{array}{ccc} S & \rightarrow & ABS \\ S & \rightarrow & e \\ AB & \rightarrow & BA \\ BA & \rightarrow & AB \\ A & \rightarrow & a \\ B & \rightarrow & b \end{array} \right\}$

(15) *Pumping-Lemma für reguläre Sprachen*:
Wenn L eine infinite reguläre Sprache über dem Alphabet Σ ist, dann gibt es Ketten x, y, z $\in \Sigma^*$, so dass y $\neq$ e und xy^nz $\in$ L, für alle n$\geq$0.

Bemerkung (s.o.):
Σ^* bezeichnet die Menge aller Ketten, die über dem Alphabet Σ gebildet werden können (der Abschluss oder 'Kleene-Stern' auf einer Menge von Ketten).

Beobachtung:
Mit dem Pumping-Lemma kann man nachweisen, dass eine Sprache *nicht* regulär ist. (Technik: Modus tollens)

(16) *Modus tollens*:

a. Wenn A, dann B.
b. Nicht B.
c. Es folgt: Nicht A.

Frage:
Ist L_1 in (17) eine reguläre Sprache? Hier müssen alle Ketten aus *n* Symbolen *a* bestehen, denen *n* Symbole *b* folgen. Falls ja, dann muss es für jedes *n* ein x, y, z geben, so dass xy^nz in L_1 ist.

(17) $L_1 = \{a^n b^n \mid n \geq 0\}$

(18) *Drei mögliche Belegungen für y*:

a. y = eine Anzahl von *a*'s, der eine Anzahl von *b*'s folgt.
b. y = eine Anzahl von *a*'s.
c. y = eine Anzahl von *b*'s.

(15) *Pumping-Lemma für reguläre Sprachen*:
Wenn L eine infinite reguläre Sprache über dem Alphabet Σ ist, dann gibt es Ketten x, y, z $\in \Sigma^*$, so dass y $\neq$ e und $xy^nz \in L$, für alle $n \geq 0$.

(17) $L_1 = \{a^n b^n \mid n \geq 0\}$

1. Fall:
(i) xyz $\rightarrow$ x = e, y = ab, z = e $\rightsquigarrow$ ab
(ii) xyyz $\rightarrow$ x = e, y = ab, z = e $\rightsquigarrow$ *abab
2. Fall:
(i) xyz $\rightarrow$ x = e, y = a, z = b $\rightsquigarrow$ ab
(ii) xyyz $\rightarrow$ x = e, y = aa, z = b $\rightsquigarrow$ *aab
3. Fall:
(i) xyz $\rightarrow$ x = a, y = b, z = e $\rightsquigarrow$ ab
(ii) xyyz $\rightarrow$ x = a, y = bb, z = e $\rightsquigarrow$ *abb

Resultat:
L_1 ist keine reguläre Sprache, weil y nicht hochgepumpt werden kann (und die resultierende Kette dann immer noch Teil der Sprache ist) – entweder wird beim Hochpumpen die Reihenfolge von *a,b* problematisch, oder die relative Anzahl von *a*'s und *b*'s.

(17) $L_1 = \{a^n b^n \mid n \geq 0\}$

Frage:
Ist Englisch (Deutsch, etc.) eine reguläre Sprache?

Beobachtung:
Der Schnitt einer regulären Sprache mit einer regulären Sprache liefert wieder eine reguläre Sprache.

Strategie:
Englisch wird mit einer bekannt regulären Sprache geschnitten; wenn das Pumping-Lemma die resultierende Sprache als nicht regulär erweist, ist bewiesen, dass Englisch nicht regulär ist.

(19) *Relativsätze im Englischen*:

a. The cat died.
b. The cat the dog chased died.
c. The cat the dog the rat bit chased died.
d. The cat the dog the rat the elephant admired bit chased died.

(20) *Form dieser Sätze*:
$(\textit{the} + \mathrm{N})^n\ (\mathrm{V}_{trans})^{n-1}\ \mathrm{V}_{intrans}$

(21) a. A = {the cat, the dog, the rat, the elephant, ...}
b. B = {chased, bit, admired, ate, befriended, ...}

Bemerkung:
L_2 ergibt sich aus dem Schnitt von Englisch und der regulären Sprache L_3 = A*B*{died}

(22) $L_2 = a^n b^{n-1} died \mid a \in A \text{ und } b \in B$

Frage:
Was sagt das Pumping-Lemma zu L_2?

(15) *Pumping-Lemma für reguläre Sprachen*:
Wenn L eine infinite reguläre Sprache über dem Alphabet Σ ist, dann gibt es Ketten x, y, z $\in \Sigma^*$, so dass $y \neq e$ und $xy^nz \in L$, für alle $n \geq 0$.

(22) $L_2 = a^n b^{n-1} died \mid a \in A \text{ und } b \in B$

Antwort:
Wie vorher lässt sich y nicht hochpumpen, ohne entweder die Abfolge oder die relative Anzahl zu zerstören, die von L_2 gefordert werden.

Konklusion:
L_2 ist nicht regulär, und damit auch nicht Englisch. Also müssen Grammatiken für natürliche Sprachen *mindestens kontextfrei* sein.

Stand der Dinge:
Genau dies ist die Annahme bei *Gazdar (1981)*: Grammatiken sind kontextfrei, aber nicht mächtiger. (Grammatiken mit Transformationen des klassischen, 60er-Jahre-Typs sind unbeschränkte Ersetzungssysteme.)

Ausblick:
Seit den Achtzigerjahren weiß man, dass natürliche Sprachen stärkere Grammatiken als kontextfreie benötigen. Der Beweis rekurriert auf ein *Pumping-Lemma für kontextfreie Sprachen*, und er basiert (u.a.) auf sog. ‘cross-serial dependencies’ im Schweizerdeutschen. Das Grundmuster ist

aber ganz ähnlich: Mehrere verschränkte Abhängigkeiten setzen komplexere Grammatiken voraus.

Streitpunkt:
Es ist umstritten, ob generative Kapazität aus linguistischer (nicht: mathematischer) Perspektive überhaupt ein taugliches Mittel zur Bewertung von Grammatiken ist.

Bibliography

Chomsky, Noam (1957): *Syntactic Structures*. Mouton, The Hague and Paris.

Gazdar, Gerald (1981): Unbounded Dependencies and Coordinate Structure, *Linguistic Inquiry* 12, 155–184.

Partee, Barbara, Alice ter Meulen & Robert Wall (1993): *Mathematical Methods in Linguistics*. Kluwer, Dordrecht.