

Syntax I Lecture Notes
Fall 2000
Anna Szabolcsi

Texts:

Fromkin, ed., *Linguistics. The Syntax Chapters*. 2000.
Roberts, *Comparative Syntax*. 1997.

There will be additional readings handed out every week.

Requirements:

Weekly assignments,
the presentation of an article that relates the contents of this class to an independent interest of yours (in a language, in psycholinguistics, etc.),
and a squib.

Topics:

1. Warmup: Categories and constituents
2. Theta-roles
3. X-bar Theory: IP, CP, verb placement
4. Theories of constituent structure
5. Subcategorization, thematic roles, and case
6. From Case theory to checking theory
7. Binding
8. Raising vs. control, PRO vs. pro
9. Wh-movement: General notions
10. Strong and weak islands
11. Theories with or without movement, theories with or without empty categories

Syntax I, 9-6-2000
Szabolcsi

1. Warmup: Categories and Constituents

Readings:

Fromkin, ed. Linguistics, Ch. 3: Argument structure and phrase structure [Stowell].

Radford, Syntax: A minimalist introduction, Ch. 2: Categories. -- **required**

Steedman, Structure and Intonation, Language, 1991/2. -- **recommended**

Assignment: Ex. 3.13 from Stowell, and Ex.III from Radford. Make your answers clear and concise. We'll discuss them in class next time.

Traditionally, a noun is the name of things, persons, ideas, etc. This definition suggests that you can only tell the part of speech of a word if you know its meaning. Is this correct?

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe...

Speakers identify the category of an expression on the basis of its distribution: the collection of properties it has in sentences of the language. A kind of meaning may be typical of a category, but that is not the test you instinctively use.

(1) Practice: Does any of these properties single out adjectives in English? Does any combination of them do?

- (i) Any word that can be preceded by very is an adjective.
- (ii) Any word that can be followed by dog (or some other noun) is an adjective.
- (iii) Any word that can take the ending -er is an adjective.
- (iv) Any word that can be preceded by the is an adjective.
- (v) ...

(2) Practice: Does any of these properties single out helping verbs (auxiliaries) in English? Does any combination of them do?

- (i) Any word that can be followed by swimming or swum is a helping verb.
- (ii) Any word that can be followed by swim but not swims is a helping verb.
- (iii) Any word that can be followed by never is a helping verb.
- (iv) Any word that can be followed by John is a helping verb.
- (v) ...

By default, we would expect items of the same category to have the same distribution (to be **systematically interchangeable** in sentences, preserving grammaticality). But are words that belong to the same part of speech systematically interchangeable? Often they are not! Consider donate and snore:

John donated money to charity.	*John snored money to charity.
John snored.	* John donated.

Parts of speech (large lexical categories) are best defined with reference to morphology. E.g., in English, only verbs take all of the inflectional endings -ed (past), -s (3pers.sg), and -ing (progressive). Lexical subcategories (e.g. main verbs versus the auxiliaries have/be) may again be definable with reference to true syntactic distribution.

NB in theories like categorial grammar, transitive and intransitive verbs would not be said to belong to the same category; they belong to different syntactic categories though they carry some features that encodes morphological commonalities.

Bigger (phrasal) categories on the other hand are distributionally coherent. Thus, the wet puppy belongs to the category Noun Phrase because it is systematically interchangeable with one-word representatives of that category, like it. In this domain systematic interchangeability with one word is also a **constituency** test: only strings that form a constituent belong to some category. Thus, the wet puppy is a constituent, because it can be systematically replaced by it, and conversely. Let us examine the notion a little further.

"Systematic" means that interchangeability goes both ways and holds in a great variety of environments. E.g., although there are environments in which and Bill can be interchanged with yesterday, this is in no way systematic:

I saw John *and Bill*. => I saw John *yesterday*.
 John *and Bill* left. =>* John *yesterday* left.
 Yesterday I danced. =>* *And Bill* I danced.

Unfortunately, we cannot require exceptionless interchangeability. For instance, animal and big animal are interchangeable in most environments, but not always:

I saw every *big animal*. => I saw every *animal*.
 The *big animal* bites. => The *animal* bites.
 I have a too *big animal*. =>* I have a too *animal*.

In such cases, we must examine carefully why the replacement gives a bad result. In the above case we may note that too big is in fact a constituent, one that is systematically interchangeable with big (except in environments like I saw a very ... animal).

Likewise, animal and big animal are not interchangeable in this environment:

I saw an *animal* =>* I saw an *big animal*.

Here we can show that the problem has to do with the fact that animal begins in a vowel, but big begins in a consonant, and not with their syntactic properties. But not all interchangeability problems can be explained away.

(3) Practice: Determine whether the italicized parts form constituents; argue using the

interchangeability test:

- (i) I saw *the book* that you gave him.
- (ii) I saw *the book that you gave* him.
- (iii) I saw *the book that you gave* him.
- (iv) I *saw the book* that you gave him.
- (v) I saw *the book that you gave* him.
- (vi) I *saw the book that you gave* him.

(4) Practice: Determine whether the italicized parts form constituents; argue using the interchangeability test:

- (i) *Not everyone* came in time.
- (ii) I invited *only Bill*.
- (iii) *More men than women* have left.
- (iv) I put *the books on the table*.
- (v) I put the books *on the table*.

The fact that a certain string of segments cannot be interchanged with a single segment does not necessarily mean that it is not a constituent. E.g., in *I saw the book that you gave him*, the italicized part is roughly equivalent to an adjective (*I saw the important book*), but while the adjective precedes the modified noun in English, the relative clause follows it, so they are not interchangeable.

Various grammatical processes (movement, replacement by a pro-form, deletion, and coordination) also turn out to be good diagnostics of constituency. Although their scope is narrower, their results converge with those obtained with the interchangeability test. We mention them here, although they really presuppose further linguistic analysis.

Movement:

- VP-Preposing:

John will *place the book on the table* <=> John will *swim*
Place the book on the table John (never) will.

- Heavy NP shift:

John put *the thickest book that he could find in the den* on the table
 <=> John put *it* on the table
 John put on the table *the thickest book that he could find in the den*

- Wh-fronting:

John read *two books about Madonna*.
 How many books about Madonna did John read?

- Relativization:

The government prescribes *the size of the lettering on the cover of these documents*.

These are the documents *the size of the lettering on the cover of which* the government prescribes.

Ellipsis, pro-forms:

- VP-ellipsis:

John *ran fast*. Bill (*did*) too.

John will *see the film*. Bill will --, too.

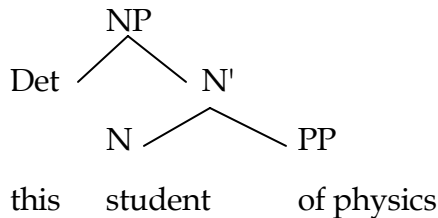
John will *have seen the film*. Joe will --, too.

John will *have seen the film*. Bill will have --, too.

- Evidence for N' being a constituent: one replaces N plus its Complement:

(a) this student of physics and that one `...and that student of physics'

(b) * this student of physics and that one of history

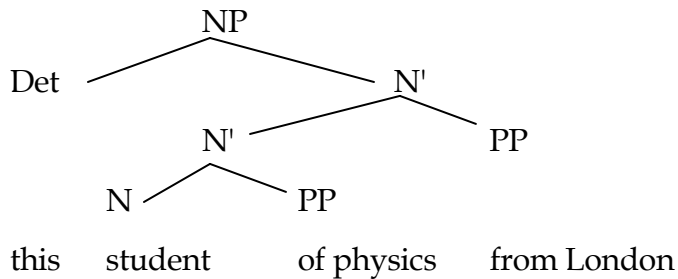


- Evidence for more than one N' layer (adjuncts):

(a) this student from London and that one from Paris

(b) this student of physics from London and that one from Paris

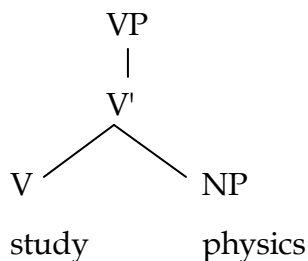
(c) * this student from London of physics



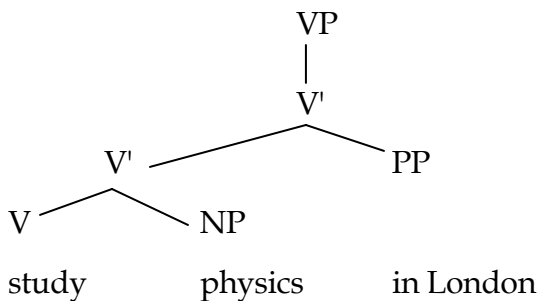
- Evidence for V' being a constituent: do so replaces V plus its complement:

(a) I studied physics and Mary did so (too) `... and Mary studied physics'

(b) * I studied physics and Mary did so history



- Evidence for more than one V' layer (adjuncts):
 - I studied in London and Mary did so in Paris
 - I studied physics in London and Mary did so in Paris
 - * I studied in London physics



Coordination:

[Mary] and [every man who has served in the army]
 every [man under 20] or [woman over 30]
 [smiled] and [put the book on the table]

This battery of tests raises questions. Why is it that some tests may show a string to be a constituent, while others do not? For instance, *man under 20* is shown to be a constituent by *one*-pronominalization and coordination, but not by movement. The answer is that each particular **movement is triggered by some particular property** of a unit. If the unit happens to have no property that triggers movement, it will not move. -- But more disturbing cases present themselves:

- [I gave Mary a book on Sunday] and [you showed Susan a record on Tuesday]
- I [gave Mary a book on Sunday] and [showed Susan a record on Tuesday]
- I gave [Mary a book] and [Susan a record]
- I gave [Mary a book on Sunday] and [Susan a record on Tuesday]
- I gave Mary [a book on Sunday] and [a record on Tuesday]

According to these coordination data, *on Sunday* may form a constituent with the whole sentence, or with its VP, or with the verb's two internal arguments, or with the verb's second internal argument! Not only does this make the position of *on Sunday* appear **ambiguous**, *Mary on Sunday* is clearly not a noun phrase systematically interchangeable with *Mary*:

*Mary on Sunday and Peter met at the airport.

How can these data be interpreted? One possibility is that coordination is after all not a good indicator of constituenthood. But coordination data square well with intonational data. As Steedman 1991 points out, the strings that can be coordinated correspond to those that can be intonationally grouped together.

coord: [Two doctors in ten] and [five nurses in twenty] prefer margarine.
 coord: * Two doctors [in ten prefer margarine] and [in twenty discourage smoking].
 inton: * [Two DOCTORS] [in ten prefer MARGARINE].

Or, Cambridge, Mass. optional stress retraction:

coord: [The Japonése] and [the English] met.
 coord: *The [Japanese won] and [English lost].
 stress: The [Jápanese cars] won.
 stress: * The [Jápanese won].
 stress: I will talk to [the Jápanese now] and [the French later].

Further, similarly puzzling data can be obtained from movement. Consider the topicalization of infinitives in German/Dutch, e.g.,

“walk has he not”

A verb that has two internal arguments can be topicalized either together with the first or together with the second, which in the above terms suggests that the constituent structure of the VP is ambiguous:

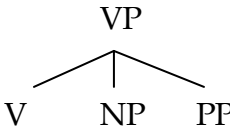
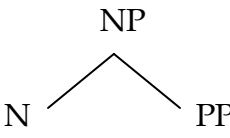
“a book give has he to John not”
 “to John give has he the book not”

The correct conclusion seems to be that constituency is not as simple a matter as it initially looked. I know of two theories that account for the problem cases. One is categorial grammar (assumed in Steedman’s article), which rejects the notion that constituent structure is rigid. Instead, words can be “bracketed together” in more than one way (though not necessarily in any old way). The **flexibility of constituent structure** is made possible by the fact that the concatenation (merging) of words is not interpreted only by functional application, but also type-lifting and functional composition can be invoked. The flexibility of constituent structure plays a major role in the categorial account of extraction without traces and many other areas.

Pretty similar results can be replicated by MIT-style syntax that recognizes **remnant movement** (originally from den Besten and Webelhuth 1989). For instance, the puzzling German examples would be derived by first moving one of the internal arguments out of the VP and then topicalizing the remnant VP. Likewise, to derive the ... *on Sunday* examples, *on Sunday* is attached to a single invariable XP, but various parts of that XP are subsequently removed. Coordination involves remnants so obtained. This explains why *Mary* and *Mary on Sunday* are not systematically interchangeable. The former is a noun phrase, but the latter is something like a remnant sentence (what is left over when everything but the direct object and the time adverb are removed from a sentence).

Structural ambiguity:

Apart from these intricate cases, there is a different and traditionally recognized notion of structural ambiguity. The sentence John hit people with sticks is semantically ambiguous, and the two meanings have different constituent structures associated with them:

- (i)  VP = instrument of hitting
 V NP PP
 hit people with sticks
- (ii)  NP = modifier of people
 N PP
 people with sticks

Let us consider a set of sentences that are, at least at first sight, quite similar, and find out whether they also have two such constituent structures.

- (1) John stole letters from children.
- (2) John hates people with sticks.
- (3) John sent people to AAA.
- (4) John brought glasses on trays.
- (5) John put books on shelves.

Ex. (1) is ambiguous: the reading 'John deprived children of letters' has an (i)-type structure, the reading 'John stole letters whose senders were children' has a (ii)-type structure. There are two reasons why the ambiguity exists. One: from children is a kind of PP that can modify letters. Two: the PP complement of the verb steal is optional.

Ex. (2) is not ambiguous: you can hate people who carry sticks (structure type (ii)), but with sticks cannot be an instrument of hating: the verb hate does not combine with an instrumental PP complement at all.

Ex. (3) is not ambiguous: you can arrange for people to go to AAA (structure type (i)), but to AAA is not the kind of PP that can modify people (although it could modify journey). Further, send is a verb that requires both an NP and a PP complement: John sent people can at best be an elliptical S.

Ex. (4) is ambiguous again: 'John used trays to transport glasses' (structure type (i)) or 'What John brought was glasses on trays' (structure type (ii)).

Ex. (5) is not ambiguous: put is subcategorized for both an NP and a PP complement. Thus, although books on shelves might be a reasonable constituent, in this context it

cannot be a single constituent.

(6) Practice: Determine the structure(s) of I warned the guy that I was going to kill. Use triangles for those categories whose internal structure is irrelevant.

Syntax I, 9-13-00
Szabolcsi

2. Theta-roles

Readings:

Fillmore, The Case for Case. In Bach-Harms, *Universals of Linguistic Theory*, 1968.
Dowty, Thematic Proto-roles and Argument Selection. *Language* 67/3, 1991.

Fillmore:

"Cases" are semantically relevant syntactic relationships between verbs and their noun phrase arguments. They are largely covert but are nevertheless empirically discoverable; they form a specific finite set; observations about them have cross-linguistic validity.

Agentive (GB th-role: Agent): typically animate instigator of the action

Instrumental (GB: Instrument): force or object causally involved in the action

Dative (GB: Experiencer): animate being affected by state or action

Factitive (GB: ?): being or object resulting from state or action

Locative (GB: Location): location or spatial orientation of action

Objective (GB: Theme/ Patient): semantically most neutral: affected by action, is in state, or simply, has a role in action identified by verb

Related principles of GB: Theta-criterion (Chomsky), Uniform Theta-role Alignment Hypothesis (UTAH, Mark Baker), etc.

Different Cases cannot be co-ordinated:

John broke a window & The hammer broke a window, BUT

* John and the hammer broke windows

John painted the fence [red] & John painted a nude, BUT

*John painted the fence and a nude

John opened the door with a key.

John opened the door.

The key opened the door.

The door opened.

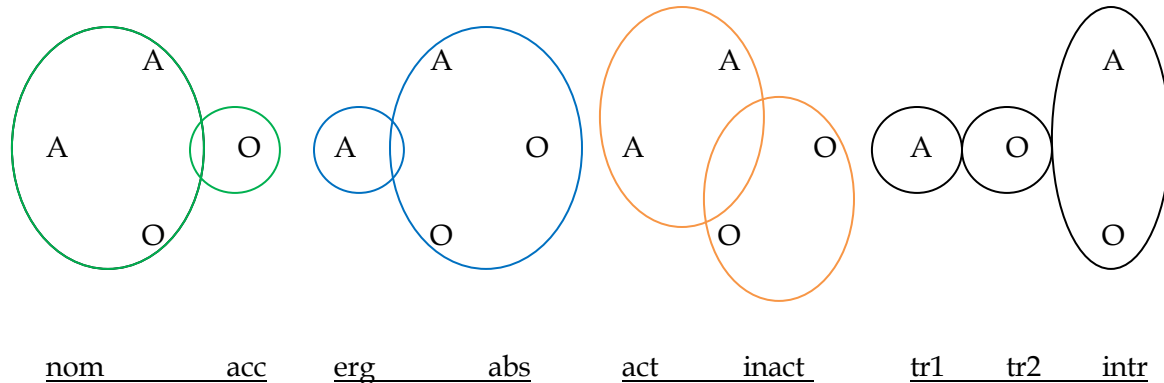
The door was opened by John.

Lexical entry: OPEN +[___O(I)(A)]

Subject selection (unmarked): If there is an A, it becomes a subject; otherwise, if there is an I, it becomes a subject; otherwise, the subject is O. (Passive sentences have marked subject selection.)

Crosslinguistic patterns of morphological marking:

- (a) John turned around (because he wanted to)
A
- (b) Mary turned John around (by pulling his sweater)
A O
- (c) John turned around (due to Mary pulling his sweater)
O



In subsequent work, Fillmore tried to push the systematic application of Case Grammar beyond the domain of physical action verbs, but it somehow did not seem possible. Why?

Dowty

Problems with the traditional notion of thematic role:

- (a) Unclear boundaries: *I walked a mile* VS *This book cost \$5*.
- (b) No motivatable role distinguishes two arguments: *Mary resembles Sue*.
John sold the piano to Mary VS *Mary bought the piano from John*.
- (c) Allegedly thematic generalization is in fact purely syntactic: *unsold car* VS **unsold customer*;
unserved soup VS *unserved customer*
- (d) Subject selection is determined by perspective: *The lamp (Figure) is over the table (Ground)*
VS *The table (Figure) is under the lamp (Ground)*

Proposal: Focus on all and only those semantic distinctions that determine argument selection

A new role: Incremental Theme (I.T.)

The (structured) Theme argument of that kind of verb which lets the event inherit the object's articulation

Definite change of state: *build houses/a house*; *drink beer/a glass of beer*; *polish a shoe*

Incremental Path Theme (I.P.T.): *cross the street*

Representation-Source Theme: *copy a file*, *memorize a poem*

versus Only indefinite change of state: *push a cart; dim the lights* (not I.T.)

No articulation inherited: *recognize a face* (not I.T.)

I.P.T. is not a syntactically realized argument: *walk from the bank to the hospital* (VS *cross the street*)

Thematic Roles as Prototypes

Proto-Agent properties:

- a. volition
- b. sentience (wrto event)
- c. causing event or change in other participant
- d. movement (relative to other participant)
- (e. existence independent of event)

Proto-Patient properties:

- a. undergoes change of state
- b. incremental theme
- c. causally affected by other participant
- d. stationary (rel. to other participant)
- (f. doesn't exist independently of event)

Argument Selection Principle for predicates with subject and direct object: The argument for which the predicate entails the greatest number of Proto-Agent properties will be the subject and the one with the greatest number of Proto-Patient properties will be the object.

e.g., *build, murder, eat, wash*

Corollary 1: If the two arguments have roughly the same # of Proto-Agent and Proto-Patient properties, either or both may be subject/object

e.g., *resemble, borrow/lend, like/please, regard as/strike as*

Corollary 2: With 3-place predicates, the non-subject with the greatest # of Proto-Patient properties will be the direct object, etc.

e.g., *put the lamp on the table*

Non-discreteness: Proto-roles don't classify arguments exhaustively, uniquely, or discretely.

Assignment:

- (1) How does Fillmore's subject selection rule (54) explain the following examples? Extend the rule if necessary, and say if an example is unaccounted for (and why). Analyze the same sentences according to Dowty's theory and point out why it fares better (if it does).
 - (a) John dropped the coin into the box.
 - (b) The coin dropped into the box.
 - (c) The rock killed the cow.
 - (d) John knows Mary.
 - (e) John likes this music.
 - (f) This music pleases John.
- (2) Ex. 9.17 (p.447) from Fromkin, ed. Comment on the significance of the findings for theories of th-roles.
- (3) Read Syntax III (Koopman), in preparation for next week's lecture. It'll make your life much easier.

Syntax I, 9-20/27-00
Szabolcsi

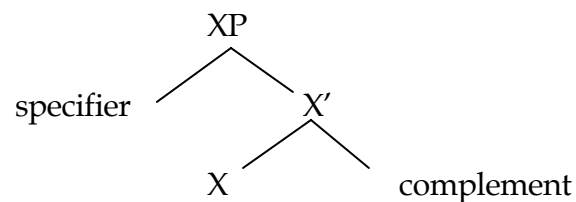
3. X-bar Theory: IP, CP, verb placement

Readings: Fromkin, ed. Ch. 5: Syntax III (Koopman).

Roberts, Comparative Syntax, Ch. 1. Categories and constituents, minus Appendix (we discuss the head initial/head final issue and Kayne's LCA in detail a bit later)
Chomsky, Syntactic Structures, up to 7.2 (p. 69).

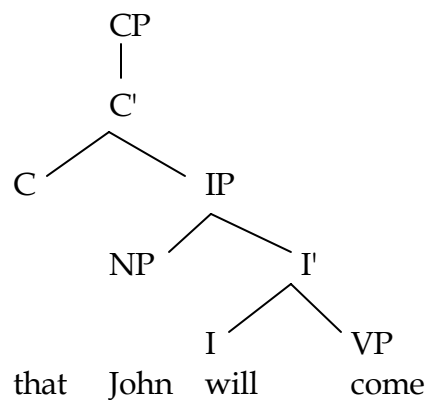
Pollock, Verb movement, universal grammar, and the structure of IP. Linguistic Inquiry 20 (1989). – **recommended, esp. if you know French**

The X-bar schema:



Are traditional S and S' also projections of some head?

Subordinate clauses with a modal auxiliary easily allow us to analyze S and S' in X-bar theoretic terms:



The constituent structure above is uncontroversial. As regards labels, let C (complementizer) be the category of that, and I (inflection) be the category of will. More on I later. This affords an analysis whereby S'=CP and S=IP. For the time being, we mention just one kind of justification.

Some verbs take an S' complement, some do not. But verbs may also be picky about whether they take that-clauses or infinitival clauses as complements:

- (7) I realized that John was home. realize only takes a that-clause
 * I realized John to be home.
 * I realized for John to be home.

- * I realized to be home.
- (8) * I started that John/I go(es) home. start does not take a that-clause
- * I started John to go home.
- * I started for John to go home.
- I started to go home.

The lexical entry of the verb can only specify the category of its complement and perhaps some feature. It cannot specify the internal structure of the complement in further detail. X-bar theory entails that specifying a category amounts to specifying what its head is. If S' is indeed the projection of the complementizer, then realize may easily require that its complement be a that-clause, and start may require otherwise.

Going one step further, we observe that that-clauses always contain either a finite verb or a modal auxiliary: they never contain an infinitive. How is this guaranteed? If that is the head of the clause, its lexical entry may specify what kind of complement it takes. If the category I, which unites inflection and modals is indeed the head of S, then the entry of that may select it.

realize, V, __ CP (where CP is headed by that)
that, C, __ IP (where IP is headed by finite inflection or a modal)

There is no other way in our grammar for these selectional properties to be captured.

Agenda:

1. Justify S=IP (I for Inflection):
 - Modal auxiliaries are elements of the category finite I; to is an element of the category non-finite I. Elements of I take bare VP complements.
 - The auxiliaries have and be are special verbs that take participial VP complements.
 - The subject NP is the specifier of I; they stand in an agreement relation.
2. Justify S'=CP (C for Complementizer):
 - That, whether, if belong to the category C.
 - Insofar as C indicates clause type (interrogative, declarative, etc.), it may occur in main clauses. In English, main clause C has no morpheme.
3. Exploit the above in the analysis of yes/no questions:
 - Subject--auxiliary inversion occurs whenever no complementizer is present in C. Inversion is I-to-C movement.
 - The notion of a movement transformation.

4. Contrast have/be with main verbs:
 - Have/be are not elements of I but they move to I: V-to-I movement.
 - Main verbs in English do not move to I: they pick up their inflectional -s or -ed by Affix Hopping.
5. Theory: what drives these movements?
6. Cross-linguistic variation in the position of verbs.

1. We begin with S=IP.

As Radford's "Categories" explained, all English auxiliaries share certain properties that set them apart from main verbs, e.g.:

John has left -- Has John left? -- John has not left
 John may leave -- May John leave? -- John may not leave
 John snores -- *Snores John? -- *John snores not

We now begin by noting that auxiliaries fall into two subsets: (i) modals and (ii) have, be.

Modal auxiliaries are can, could, shall, will, should, would, may, might, must, etc. All modals are characterized by the fact that only one of them can occur per sentence:

John may sink
 * John may will sink

they do not take 3pers.sg (-s) or past (-ed) morphology:

* John mays sink
 * John mayed sink

their complement is in the bare infinitival form, not an inflected form or participle:

John may sink
 * John may sinks/sank
 * John may sunk/sinking

and they can be followed by the auxiliaries have and be:

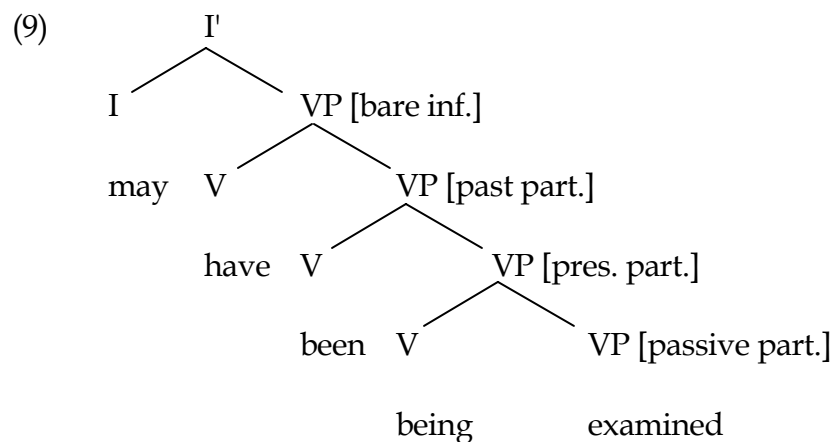
John may have sunk
 John may be examined
 John may be sinking

John may have been examined
 John may have been sinking
 John may have been being examined

Perfective have, passive be, and progressive be, however, combine with each other (in a specific order), follow modals, can be inflected for 3pers.sg or past, and take participles as complements:

John has/had sunk	perfective <u>have</u> + past part.
John is/was examined	passive <u>be</u> + passive part.
John is/was sinking	progressive <u>be</u> + present part.

This suggests that modals and have/be are really two different kinds of beast. The proposal in the literature is that have and be are a special kind of V that selects VP-complements of a particular participial format, while modals are instances of the category I(nflection). We omit the V' levels from the diagram for simplicity.



We are primarily concerned with I(nflection). We are going to assume that segments that are traditionally relegated to morphology can in fact belong to syntax. The virtue of this assumption is that it enables us to state various syntactic generalizations that would be missed otherwise.

Assume that I is a head that is primarily a cluster of features, comprising person-number and tense. Inflection can be [+finite] or [-finite], i.e. infinitival. The simplest segments that belong to the category I [+finite] are -s and -ed. The phenomenon of do-support corroborates the idea that tense and agreement information is essentially separate from the contentful verb itself:

(Bill may not watch the news but) John does.
 (Bill may not have watched the news but) John did.

Now, while modals do not carry the -s or -ed endings, they are clearly [+finite]. Both modals and -s/-ed marked verbs occur in that-clauses, are unique per clause and cannot

(12)	C[+finite]	C[-finite]
	<u>that</u>	[+declarative.subordinate]
	<u>whether</u>	[+interrogative.subordinate]
	<u>if</u>	[+interrogative.subordinate]
	<u>if</u>	[+conditional.subordinate]
	etc.	
		<u>for</u>
		etc.

Are main clauses also CPs? If we think of C purely as something that turns a sentence into a subordinate clause, then they are not. But the [+interrogative] feature of whether can certainly be shared by a main clause.

C in English conflates two functions: complementizing and indicating clause-type. There are languages, e.g. Korean, in which the two are carried by two separate morphemes, and clause-type indicators (-da, -ni, etc.) occur in main clauses, too. In subordinate clauses, an additional subordinating morpheme (-ko) is added. (See Bhatt-Yoon, WCCFL 10.)

(13)	John-i	o-ged-da
	John-nominative	come-conditional-declarative
	`John would come'	
	John-i	o-ged-ni
	John-nominative	come-conditional-interrogative
	`Would John come?'	

So assume that (at least some) English main clauses are also CPs.

3. Yes/no questions

We first exploit the assumption that (some) English main clauses are CPs in the analysis of yes/no questions. At this point, we examine the simplest example.

May John leave?

This sentence is very difficult to accommodate if the grammar has nothing but phrase structure rules (or merger), as our grammar currently does. If we did not have X-bar theory, the situation would not seem so bad. We could have two S-rules:

S → NP Mod VP
S → Mod NP VP

But the second rule gets us into trouble with subordinate clauses: whether can only combine with an S of the first type:

I wonder whether John may leave. whether NP Mod VP

* I wonder whether may John leave.

* whether Mod NP VP

Maybe the solution is that Mod NP VP is an S', not an S? Subordinate clauses are still a problem:

* I wonder may John leave.

The PS rule treatment of subject/auxiliary inversion runs into descriptive problems. It is also conceptually suspect. $S \rightarrow \text{Mod NP VP}$ "flattens out" the constituent structure in a crucial way. And there is no hope to marry it with X-bar theory.

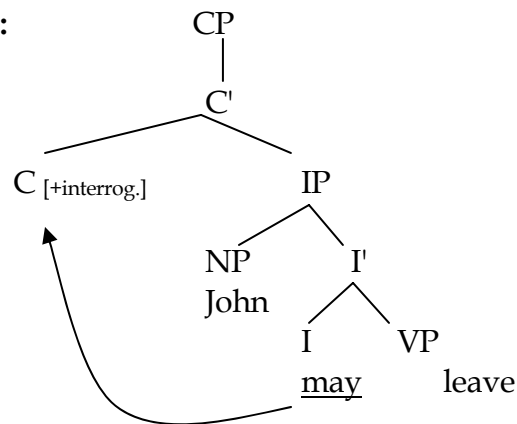
The alternative is to introduce an entirely new device: **transformations**. A transformation may rearrange the structure produced by phrase structure rules.

Should we use a transformation that lowers the subject NP, or one that raises I higher? There is a good descriptive reason for assuming the latter. The NP-lowering assumption will not help to capture following descriptive generalization:

Inversion never takes place if there is an overt complementizer.

If, on the other hand, we assume (i) that C in English can only accommodate a single item and (ii) that the content of I does not just move forward but moves specifically to C, we capture that inversion and the presence of a complementizer exclude each other:

(14) **I-to-C:**



If C[+interrog.] contains if or whether, there is no inversion. If it does not, the content of I moves to C. This is supplemented by the observation that if and its brothers have a [+subordinate] feature, wherefore they do not occur in main clauses.

A similar complementarity (mutual exclusion) exists between inversion and the if of conditionals:

If Bill calls, ...

Should Bill call, ...

* If should Bill call, ...

and between inversion and the that of subjunctives:

I wish that he live long.
 May he live long.
 * I wish that may he live long.

Let us take stock. We substantiated the claim that $S=IP$ and $S'=CP$, and introduced a new theoretical device: transformations. Some of the questions that arise from what we have done are as follows:

Theoretical questions:

How do heads like V and heads like I differ?

What units can transformations move, where to, why, and under what conditions?

Analytical question:

How does the content of I get attached to V, when I is not a modal?

We give preliminary answers to the theoretical questions now, so that we can turn to the analytical one before plunging into a new set of facts.

The heads N, V, etc. are called **lexical** heads, and their projections lexical projections. The heads I and C are called **functional** heads, IP and CP are functional projections. The inhabitants of I and C do not have much descriptive content; they primarily have grammatical content (they are "function words"). The two kinds of heads participate in the same X-bar schemata, but they differ as to what the specifier--head and the complement--head relations consist in.

The most important types of movement are (i) movement of a maximal projection to a specifier position (**XP movement**), and (ii) movement of a head to another head position (**head movement**). I-to-C is an example of head movement.

4. Have/be versus main verbs

The auxiliaries have and be have the same distribution as modals: they precede negation and invert with the subject in yes/no questions:

- (15) John may not leave
 John has/had not left
 John is/was not leaving
 John is/was not abandoned
- (16) May John (not) leave?
 Has John (not) left?
 Is John (not) leaving/abandoned?

Should they be inserted into I, too? This would contradict the earlier claim that they are Vs that have inflectional morphology attached to them. Plus, in infinitival clauses they are preceded by to, which is allegedly I[-finite]:

- (17) to have left
to be leaving/abandoned

The easy solution is that have and be raise to I[+finite], rather than get inserted into I. Assuming that not is the head of a new functional projection, NegP, the structure is this (the presence of NegP is optional):

- (18)
- | | | | | | |
|----|------|--------|---------|------|---------------------|
| CP | | | | | |
| | C' | | | | |
| | C | IP | | | |
| | | NP | I' | | |
| | | | I[+fin] | NegP | |
| | | | | Neg' | |
| | | | Neg | VP | |
| | | | | V | VP |
| | John | may | not | have | left |
| | John | -s | not | have | left |
| | | ↑----- | | | = John has not left |

But main verbs have a very different distribution:

- (19) * John smiles not -- John does not smile
* Smiles John (not)? -- Does John (not) smile?

Conclusion:

Auxiliary have and be raise to I[+finite] in English. (**V-to-I head movement**)

Main verbs (smile, etc.) stay in VP. Following the Chomskyan tradition, we will for the time being say that inflectional morphology lowers onto them. (**Affix Hopping**)

Do the same rules apply when there is no overt suffix? Clearly yes, because 2pers verbs have the same distribution as 3pers.sg verbs. This suggests that the features, not the actual suffixes are relevant.

- (20) You have not left. and not: * You not have left.

Affix hopping is apparently **blocked by the intervention of negation** between I and V. Thus, do support kicks in: a dummy verb is inserted:

- (21) * He not smiles.
He does not smile.

Likewise, the fact that main verbs in English do not leave VP forces do-support for the sake of I-to-C:

- (22) He smiles.
* Smiles he?
Does he smile?

The above structures also nicely **predict the position of adverbials** like often, always, etc. Assume that these are VP-adjuncts (roughly the same position as NegP):

- (23) John may always leave. * John always may leave.
John is always sad. * John always is sad.
John has always helped. * John always has helped.
* John smiles always. John always smiles.

Always follows modals and finite have/be, because these are in I. It precedes finite smiles, because the latter is in VP.

Unlike negation, the VP-adjunct does not block I-lowering. We do not get * John does always smile, in analogy to John does not smile. Why? We may capitalize on the fact that the inflectional affix is a head, while the VP-adjunct is an XP.

5. What triggers all these movements?

First, **verbs need to pick up their inflection** (tense and agreement). As we have seen, they may do so either by raising to the functional heads in which those features are located, or by the features lowering to them. The first is an entirely well-behaved transformation; the latter is an ad hoc one (we will see that lowering is generally not possible), but we adopted it for convenience. In more recent literature, it is assumed instead that in English, the verb moves to inflection covertly (Chomsky), or that the reason why it looks as if the verb wasn't moving is that it pied pipes a larger phrase into inflection, and this larger phrase also contains the adverb, for instance (Hallman <http://www.leidenuniv.nl/hil/hilnews/hnews51/hmatters.htm>, Sportiche).

Second, functional heads that contain merely features need to be associated with lexical material (some word or other). Features need a carrier: we may call this the "**No Stray Feature**" principle. Thus, the person-number.tense features of I acquire a carrier in English either by having have/be raise to them, or by hopping onto the V in VP, or by the

insertion of dummy do. Although the first two operations are now doubly triggered (by the needs of both the verb and the inflectional feature), the third, do-support requires "No Stray Feature" to be a separate principle.

"No Stray Feature" may be responsible for **I-to-C** movement as well. It appears that the [+interrogative] and the [+conditional] features of C need a carrier. The [+declarative] feature may not need one (at least in English main clauses).

Person-number agreement between the subject and the verb requires a specifier--head configuration. Does this requirement trigger movement, too? We assumed that [person.number] is a feature of I, so as far as the verb is concerned, its presence in I is not necessary. If the subject NP enters the structure directly under IP, then it is in the right place, too. If, however, we assume that it originates lower (as we will, later), then it needs to raise to agree with I.

6. Cross-linguistic variation in the position of verbs

So far we have limited the discussion to English and the theoretical assumption/explained fact ratio is not very impressive. The more persuasive arguments in favor of the proposed analyses come from the insights they allow into the differences and similarities between languages. Let us start with comparing the following **English and French** sentences. The English order is ADV V_{finite}, while in French the order is V_{finite} ADV.

- (1) E. John often visits his parents
 F. * Jean souvent visite ses parents
- (2) E. * John visits often his parents
 F. Jean visite souvent ses parents

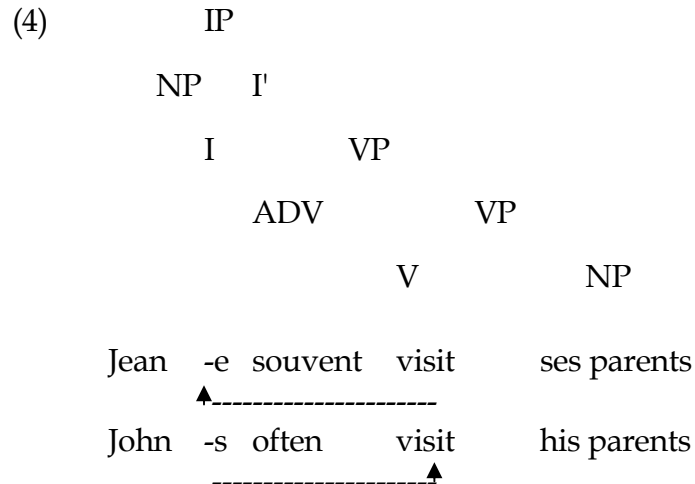
A first step in understanding the issue is to see that the "English" order is also good in French if the verb is non-finite, either a participle or an infinitive. (The finite auxiliary a precedes souvent, exactly like finite visite does.)

- (3) (a) Jean a souvent visité ses parents
 J. has often visited his parents
- (b) (Jean a décidé de) souvent visiter ses parents
 (J. has decided to) often visit his parents

So, a proper account of the difference between French and English should also account for the difference internal to French between the non-finite ADV-V order (3) and the finite V-ADV order (2F).

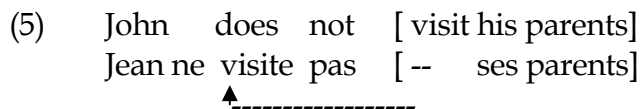
Let us look more closely at (2F): the verb and its complement are separated here. Assuming that the principles governing X-bar structure do not vary across languages, but

are rather part of Universal Grammar, this separation should have come about as a result of movement. We can then assign the same structure to the sentences in (1) and (2), i.e. the structure in (4):

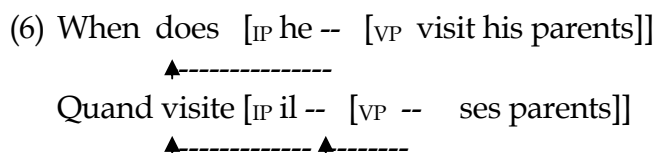


The word order difference is immediately captured if **the inflectional morphology in English lowers to V, while V raises to I in French. V thus remains in its base position in English, while it occurs in the I-position in French.** As the verb is not inflected for tense and agreement in (3), it remains in its base position as well, yielding the same order as in English.

This analysis is corroborated by negative sentences. As shown above, the presence of *not* requires, in sentences without auxiliaries, the presence of *do* to carry the I-information. Apparently, I cannot lower to V if *not* intervenes. The raising of V-to-I is not blocked by negation in French, as is evident in (5). The -- indicates the original position of V.



Note that the difference we have set up between French and English so far makes the prediction that also in questions, the positions of the verb may differ. In French, since the verb moves to I, it may move along with I to C. This is indeed the case, although there is a complication in French questions, due to which we must illustrate this with a pronominal subject:



When auxiliaries are involved, no difference between English and French is expected, and this again is correct:

(7) When can [IP he -- [visit his parents]]
 ↑-----
 Quand peut [IP il -- [visiter ses parents]]
 ↑-----

So, the clause structures of French and English are basically identical, differences in word order patterns coming about as the result of a difference in the linking of I and V.

Nota bene: we have not mentioned all the possible word orders in French. E.g., ne pas visiter souvent ses parents and ne pas souvent visiter ses parents are both good. This can be accounted for if I is split into Agr and T, and visiter can either stay in VP or raise to T, between negation and the adverb.

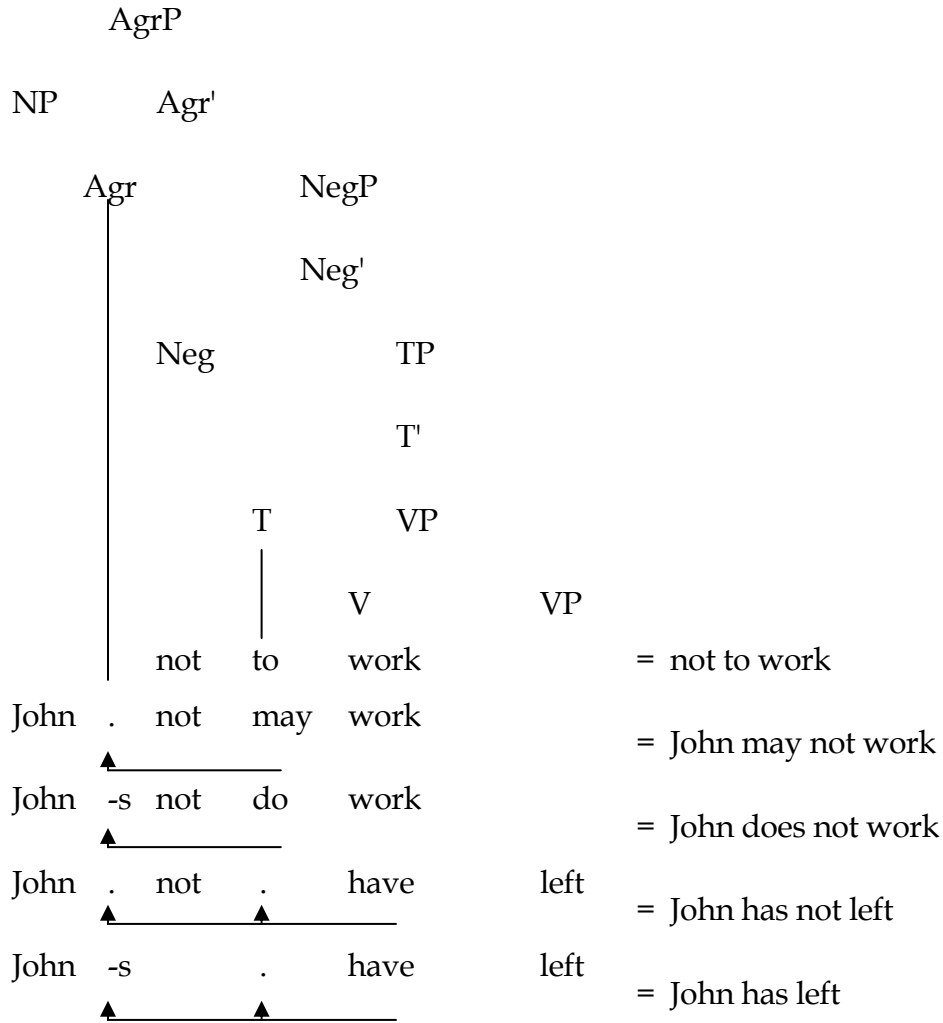
Splitting I into Agr and T

Earlier we suggested that to is I[-finite], but this predicts the less preferred order with respect to negation:

?? To not work (is pleasant)
 Not to work (is pleasant)

Finite I has at least two kinds of feature, person-number agreement and [+tense/mood]. It makes sense to say that to is a carrier of [-tense] only. We may then split I into two functional projections: AgrP and Tense/MoodP (TP), with AgrP above and TP below NegP.

The subject is naturally the specifier of Agr, not T. V-to-I is now broken down into two-steps: V-to-T and T-to-Agr. It makes sense to assume that modals are generated in, and do gets inserted into T; then they both move on to Agr. Affix hopping is also two hops (Agr-to-T-to-V). The split allows the correct placement of to.



Perhaps to can optionally raise to Agr, giving rise to the less preferred to not work order. (It is not clear though what might trigger to-to-Agr movement.)

Assignment #1:

- Fromkin: Syntax III, Exx. 5.10-11
- Roberts: Ch. 1, Ex. 5. (do ex. 4 or recap Radford for the F/E comparison).
- Fake Hungarian #1

The following set contains grammatical Hungarian sentences. Several descriptive generalizations concerning word order can be recovered from this set. Figure these out and propose a mini-grammar. Propose an initial structure and specify rules that derive all the order-variants. – If you feel you need more data, you can construct up to five Hungarian sentences and I'll give you the judgments.

Make the structure as simple as possible; feel free to leave (some of) the nodes unlabeled. Demonstrate that you understand the patterning of the data; use fancy theoretical tools only insofar as they are clearly needed to account for the data.

Caveat: This is a simplified set of data; the picture it suggests is in some crucial respects false of Hungarian. But it is altogether logically possible for a language to work the way it suggests.

Glossary: fog = will, olvas=read, akar=want, kezd=begin, szalad=run, el=away, a(z)=the, udvar=yard, -ni=infinitive, -t(t)=past, -on=locative.

Kati fog olvas-ni.
 Kati olvas.
 Kati olvas-ott.
 Kati fog akar-ni olvas-ni.
 Kati fog olvas-ni akar-ni.
 Kati akar-t olvas-ni.
 Kati olvas-ni kezd.
 Kati fog akar-ni olvas-ni kezde-ni.
 Kati olvas-ni kezde-ni akar-ni fog.
 Kati fog akar-ni kezde-ni olvas-ni.
 Kati el-szalad.
 Kati kezdett el-szalad-ni.
 Kati fog kezde-ni el-szalad-ni.
 Kati el-szalad-ni akar-t.
 Kati el-szalad-ni kezde-ni akar-t.
 Kati szalad az udvar-on.
 Kati fog akar-ni szalad-ni az udvar-on.
 Kati fog akar-ni kezde-ni szalad-ni az udvar-on.
 Kati fog szalad-ni az udvar-on.

Does your analysis predict the following orders to be grammatical? Why? The answer should be of the following format: This order is (is not) predicted to be grammatical, because it is (is not) derived by a simple application of movements that I needed to postulate in order to account for the data in the handout.

- (1) Kati kezdeni olvasni fog akarni.
- (2) Kati olvasni akarni fog.
- (3) Kati az udvaron szaladni akart.
- (4) Kati elszaladni fog akarni kezdeni.

Assume that some of the examples represent the basic order and the rest are derived by movement. Draw a binary branching tree diagram for what you wish to take to be the basic order. Leave the nodes unlabeled, or label them purely mnemonically, e.g. 1/1P, 2/2P, 3/3P, etc. In this way you will be focusing on the patterns you discover, rather than trying to guess some theoretical analysis. Show how the other orders can be derived by simple mechanical movements. Do not worry about inflection; stay at the word level.

Assume that movement is always upward in the tree, and that it can be either (i) head movement, which can (ia) fill a vacant head position, like T-to-C movement in Koopman or (ib) left-adjoin a head to another head, like V-to-T movement in Koopman, or (ii) phrasal movement, which moves an XP into a vacant specifier position, like the movement of the subject to specifier of TP in the Q-float exercise. (I am not saying that you necessarily need all these tricks, this is simply the menu you can choose from. Indeed these are the kinds of movement we are going to assume in this semester, although we will impose further constraints on their operation.)

Assignment #2

Read Syntactic Structures at least up to p. 69.

- (i) Lay out how SynStr derives the following sentences: John has been reading this book and Has John left?, and how we did it in class. Compare the two sets of derivations. Discuss how each theory captures the fact that have requires a past participle, that auxiliaries line up in a specific order, that have combines with the affix -s, and that the order of the subject and the auxiliary is different in declaratives and interrogatives.
- (ii) After this technical discussion, discuss what is the significance of these data for each of the two theories, what interesting insights or challenges they offer for them, what strategic differences you find in the way the two theories approach them, etc.

Thus, the first goal of this assignment is to see how the same data and generalizations can be captured in two different frameworks. The second goal is to raise beyond this (extremely important) technical level and place the issue within the two different theories. Note that there is no expectation for you to conclude that in roughly 40 years, we have become much smarter or something; in general, this exercise is not about ideology.

You are encouraged to proceed as follows:

Step 1: everyone reads the textbooks and Syntactic Structures on his/her own.

Step 2: you get together and discuss the comparisons.

Step 3: everyone writes up the results on his/her own.

We will review the results together in the next class.

Syntax I, 10-4-00
Szabolcsi

4. Theories of constituent structure

Readings:

Roberts, Ch. 1, focus on word order typology and the LCA and read Appendix

Kayne, *The Antisymmetry of Syntax*, Part I -- **recommended**

Greenberg, *Some universals of language...* -- **recommended**

- *Comparison of the treatments of verb placement, subcategorization, and inflection in Syntactic Structures and in Pollock 1989 and after*
- *Syntactic Structures: finite state automata vs. PS rules vs. (linearly defined) transformations*
- *Top-down PS rules (i) vs. bottom-up merge (ii):*
 - (i) Lexical subcategorization information serves to filter the output of the abstract PS component (from Syntactic Structures up to Minimalism);
 - (ii) Lexical subcategorization information drives the construction of legitimate phrase structures (categorial gr., HPSG, Minimalism).
- *How to derive the desired properties of X-bar theory?*

Some order facts (V DP PP, *V PP DP, V CP PP, V PP CP in English) to be derived from Case theory (Stowell 1981)

Binary branching: unambiguous paths (Kayne 1983)

Various (no head-to-XP or XP-to-head adjunction, no head takes head as complement, no multiple adjunction, no spec and comp on same side of head, etc.): Linear Correspondence Axiom, LCA (Kayne 1994; Appendix to Ch 1 in Roberts)

LCA: (informally) Hierarchy and linear order match up. (Whether “higher stuff precedes lower stuff” or “higher stuff follows lower stuff” is determined independently.)

LCA: (less informally) Non-terminal (=category symbol) A asymmetrically c-commands non-terminal B if and only if all terminals (=morphemes or features, overt or phonetically null) dominated by A precede all terminals dominated by B, OR iff ... follow ...

where node A c-commands node B iff every node dominating A dominates B.

LCA: (formally) $d(A)$ is a linear ordering of T

where A is the maximal set of ordered pairs $\langle X, Y \rangle$ of non-terminals such X asymmetrically c-commands Y, $d(A)$ is the set of terminals these dominate, and T is the set of terminals

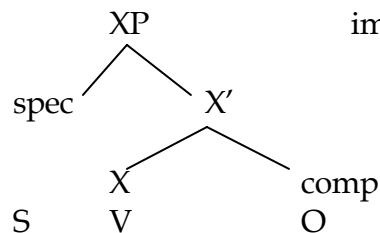
- Word order universals vis-a-vis X-bar theory

Greenberg 1963: OV – OP – PossN – AdjN versus VO – PO – NPoss – NAdj
 Tentative X-bar parameters: X precedes/follows complement within X'
 spec precedes/follows X'

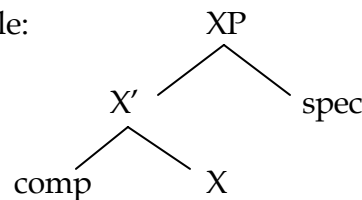
Problems: Poss and Adj are not complements; OV-OP don't always go together; specifier final languages are extremely rare

Kayne 1994: Fix the ordering relation in the LCA so that higher stuff precedes lower stuff.
 Then:

possible:



impossible:



cf.

S V O

Apparent counterexamples need to be accounted for using movement. We'll come back to this many times during the semester.

- *Universality of phrase structure? A separate question.*

Empirical hypothesis (Cinque 1999): the order of XPs is universally fixed. That is, not only do languages not vary as to whether complements follow or precede heads in them, they also do not vary as to whether NegP is below or above TP, for example.

Assignment (whether all exx are assigned will be specified by email)

Roberts, Categories and constituents, Exx. 1-2-3, Ex. 6.

NB the "Try to put together..." part of Ex.1 is the most interesting task. Don't miss it! A certain section in Koopman's chapter will give you a useful hint.

Syntax I, 10-11-00
Szabolcsi

5 Subcategorization, thematic roles, and case

We have taken the position that phrase structure, i.e. the internal structure of phrases, is uniform, and relatively simple. Each phrase is **headed**, i.e. phrases are groups of words built around a head. The head determines the category of the group (i.e. the category of the head is projected to the level of the phrase) and the head enters into two relations within the group: the head-complement relation and the head-specifier relation. This gives us the following basic X-bar schema.



X, W, Y are variables ranging over syntactic categories, so far: V(erb), N(oun), A(djective), P(reposition), I(nflection), C(omplementizer).

Each member of these categories will project a phrase built in accordance with the X-bar schema. In addition to these two head-dependent relations, we have assumed that phrases may be modified through adjunction. As SPEC and COMP are head-dependent, while adjuncts are not (they are optional), we shall in what follows only represent SPEC and COMP, adding adjuncts only when required.

Subcategorization

The lexical entry of each head specifies whether it takes complements and if yes, what the category of the complement should be. All lexical categories have representatives that can occur without complements, e.g.

- | | | |
|------|--------------------|---------------------|
| (25) | John runs. | V has no complement |
| | John ran in. | P has no complement |
| | John was faithful. | A has no complement |
| | John saw the book. | N has no complement |

When a complement is required or allowed, there are various category options:

- | | | |
|------|--------------------|---|
| (26) | subcategories of V | |
| | -- DP | read a book, buy the cookies, visit the museum |
| | -- PP | depend on your parents, appeal to the court, long for peace |
| | -- AP | become rich and famous, be happy, seem nice |
| | -- CP | say that it is late, wonder what happened |

Which category will function as the complement of V depends on the particular V chosen: this is what is called subcategorization. Not only verbs are subcategorized for particular types of complements, members of N, A and P are similarly subject to subcategorization wrto their complement category:

(27) subcategories of P:

- DP in / under/ near/ behind *the box*
- PP from *before the war*, until *after dinner*
- CP before *John came in*, after *the bomb fell*

(28) subcategories of A:

- PP faithful to one's principles, dependent on the job
- CP afraid that the bomb will drop, sure that the world is round
- *-- DP *afraid the dog (cf. fear the dog)

(29) subcategories of N:

- PP trust in the economy, shame over one's past
- CP announcement/claim/fear that the bomb will be dropped
- *-- DP *(John's) fear the dog, the destruction the city (cf. John fears the dog, the army destroys the city)

(3)-(6) show that each of the categories V, N, P, and A may take various types of categories as their complement. We also observe a gap: A and N cannot take DP complements (without an accompanying of). This is something that our theory will explain later.

What about subjects? In sentences like John arrived, John occurs in the specifier of IP, not inside VP. But do only IPs have subjects?

- (30) We saw [John pick up the knife]
 We want [John off my ship]
 We find [John smart]
 We elected [John president]

It makes sense to say that John is the subject of the bracketed phrase [...] in each example. For simplicity, assume that these phrases are just VP, PP, AP, and DP. If so, then subjects originate as specifiers of some lexical category. They may end up elsewhere (in the specifier of a functional projection, like IP); we will return to this but it does not concern us now.

Does the lexical entry of the head need to specify what category the subject (its specifier) will be? In the examples above, the subject was always an DP. Could it be otherwise? Some potential counterexamples:

- (31) [_{CP} That the Earth is round] surprised John.

[_{PP} Under the bed] is a good place to hide.
 [_{AP} Free] is what you'll never be.

If these counterexamples can be explained away, subjects are always DPs, and the verb's entry does not need to specify its category.

Thematic roles (recall Dowty for a more nuanced picture)

There is another respect in which the head certainly needs to take care of both its subject and its complement(s).

The meaning of hit is such that it describes an event that cannot occur without at least two participants: the one who hits and the one who is hit. Sometimes the first is called the Agent and the second the Patient of hit. (Haegeman lists further roles.) We say that hit has two **roles** to assign: the hitter (agent) and the hittee (patient) role. Thus, the lexical entry of hit may be as follows:

hit: V, <agent DP, patient DP>

The information enclosed in angled brackets is called the verb's **theta-grid**.

The meaning of run is such that the event it describes can occur with a single participant: the one who runs. Thus, one entry for run will be as follows:

run: V, <agent DP>

A fundamental requirement in syntax is that there be an exact match between roles and role-players.

Theta-criterion:

Every theta-role is assigned to one and only one potential role-player.

Every potential role-player receives one and only one theta-role.

The Theta-criterion excludes the following ungrammatical examples:

- (32) * John hit. (the patient role is not assigned to any DP)
 * John hit the ball the boy. (one of the DPs is left without a role)

At this point, theta-roles and the theta-criterion seem somewhat superfluous. Why isn't it enough to specify that hit needs one DP specifier and one DP complement? The real significance of the theta-criterion will become obvious when we introduce XP-movement. There will be no one-to-one relation between DP-positions and theta-roles, and the theta-criterion will do a very useful job. It will be important to state that transformations (movement) must not obscure the syntactic representation of lexical properties. This is ensured by the **Projection Principle**.

The term "**potential role player**" in the Theta-criterion is obviously metaphorical. To see what it should stand for, we need to consider at least three things.

(A) Main verbs versus auxiliaries and functional heads. The lexical entry of auxiliary be must mention that be requires a VP complement, and the entry of the inflectional head -s must mention this too. But whereas main verb hit assigns the patient role to its complement, the auxiliary be in be talking and the I -s in talks do not assign any comparable role to their complements. This is acceptable only if VP does not qualify as a "potential role player". Likewise, it is doubtful that the complementizer that assigns a role to its IP complement; we must assume that IP is not a "potential role player", either.

(B) Referential expressions versus expletives (pleonastics). Consider these:

- (33) It seems that the Earth is flat.
There remained no crumbs on the floor.

What role might it and there have here? Nothing obvious comes to mind! The verbs seem [__ CP] and remain [__ DP PP] do not have theta-roles for their subjects. Thus, the Theta-criterion cannot require the presence of these subject DPs, but then what does? It seems that a lexicon-independent principle does:

Clauses have subjects. (=The Extended Projection Principle)

Now note that not any old DP can be the subject of seem and remain:

- (34) * John/the situation seems that the Earth is flat.
* John/cleanliness remained no crumbs on the floor.

The reason must be that the above DPs are potential role-players. The expletive it and there are not, wherefore they can survive without a theta-role. (These DPs have non-expletive homonyms, cf. I like it and I put the book there.)

"Potential role-players" are officially called **arguments**. They are referential (contentful) expressions occurring as specifiers and complements of **lexical** heads. Expletives, VP, and IP are not arguments; adjuncts are not arguments. They are not mentioned in the verb's lexical entry. They cannot and need not receive a theta-role.

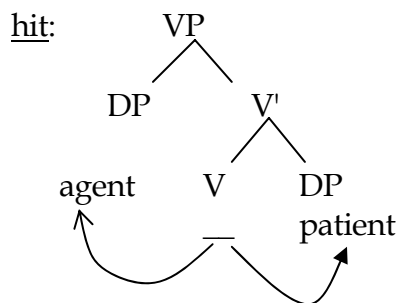
(C) DPs versus PPs. Referential DPs are always arguments. What about PPs? PPs can appear either as complements or as adjuncts. Adjuncts are optional modifiers. They are not mentioned in the verb's lexical entry, wherefore they clearly do not receive a theta-role from the verb. This is the case in John can read in this room. Now what about comparable PPs in complement position, e.g. John put the book on the table? Does on the table receive a theta-role from put? We might say so, but then why is it that one PP needs a theta-role and another does not? The standard answer to this dilemma is that PPs never get a theta-

role. The DP inside the PP needs and gets one. If the PP is an adjunct, P assigns the role to DP (in assigns a role to the room). If PP is a complement, the verb and P jointly assign the role to DP (put and on join forces to assign a role to the table).

The reasoning in (C) immediately explains why DPs are never adjuncts. They need a theta-role, the verb cannot assign them one, and there is no P to assign them one, either. (Rare adjuncts of the type this way, this year require special considerations.)

The alignment of theta roles (see Mark Baker's UTAH and its critics)

Finally, does the verb's entry specify that the agent role is assigned to the subject (specifier) DP and the patient role to the complement? As a first approximation, we may say it does. If so, the lexical entry of hit will be something like this:



The theta-role the verb assigns to its complement (DP, PP, or CP) is called an internal theta-role, the one it assigns to the subject is an external theta-role. It makes sense to assume that the subject originates in Spec, VP and moves to Spec, IP for some independent reason, to which we turn shortly.

It is interesting to ask whether we can be less specific about the lexical entries of individual verbs and state some generalization regarding which DP each role is assigned to. We do not decide this issue here, but mention two relevant issues.

(α) Different frames for the "same" verb?

John opened the door.
The door opened.

How many verbs open are there in English? Two, where *open*₁ assigns the patient role to the complement and *open*₂ to the specifier? Or just one, which always assigns the patient role to the complement? The latter is an interesting hypothesis, but then we must explain how that patient winds up as a subject!

(β) Generalizations across verbs and languages? Interestingly, we know of no transitive verbal root in any human language with subject=patient and complement=agent. ("Verbal root," because derived verbs may appear in this configuration.)

- * John blicks Peter
 `John undergoes a blick-type change through Peter's agency'

This suggests that perhaps the patient role is necessarily assigned to the complement and the agent role to the specifier. (This assumption may be tenable, if we can successfully explain The door opened.) On the other hand, there exist verbal pairs of the sort like--please:

John likes this music
 This music pleases John

The subject of like has an experiencer role, its object a stimulus role; this is the exact inverse of please. This indicates that a particular role is not always assigned to the same position. Importantly though, these latter roles are not agent and patient. Thus, such generalizations must be carefully stated.

The structural conditions on theta-role assignment

We said that a head assigns a theta-role to its complement and maybe to its specifier. On the other hand, we want to exclude the possibility of a head assigning a theta-role to some other head's specifier or complement. We will need a general notion of under what structural conditions a head assigns a theta-role to an argument.

Case assignment

Besides subcategorization and theta-role assignment, case assignment is an important aspect of the head--argument relation. We are going to explore what heads assign case and under what structural configuration.

In English, only some pronouns show case distinctions. The case of the subject is nominative; the case of the direct object is accusative.

He likes him, She likes us, We recognize her, They left them

In many other languages, all nouns and pronouns exhibit case distinctions. We will assume that also in English, all of them have abstract (non-audible) case.

It has always been recognized that the presence of accusative case on the direct object is related to the verb of the sentence. In our terminology, we will say that **the transitive verb assigns accusative case to its direct object.**

What else assigns accusative? Prepositions: about him, behind her, to us, etc.

How about nouns and adjectives? The following contrasts are telling:

- (a) John announced the news
- (b)* the announcement the news
- (c) the announcement **of** the news

John fears dogs	John fears dogs
* his fear dogs	* He is fearful/afraid dogs
his fear of dogs	He is fearful/afraid of dogs

Fear (noun), fearful (adjective) and afraid (adjective) take the same kind of direct object complement as the verb fear, nevertheless, they do not assign accusative case to it. Two analyses have been proposed for the role of of here. One is that it is a dummy case marker that assigns accusative to the nouns's or the adjective's complement. Another is that nouns assign **genitive** case, and of is a genitive case-marker (as opposed to a case assigner).

It is cross-linguistically true that **nouns and adjectives do not assign accusative**, but not all languages have a device like dummy of. In Hungarian, for instance, nouns and adjectives never have a direct object in complement position. (John's announcement of the news is not directly translatable into Hungarian.)

Do all verbs assign accusative? * John raves him

The ability to assign case, exactly like the ability to assign a theta-role, is specified in the lexical entry of each head.

Where does nominative come from? Are all subjects in the nominative?

- She walks.
- * She to walk would be surprising.
- * It would be surprising she to walk.
- * I want she to go.
- * I want she off the ship.

We see that only the subjects of inflected verbs are in the nominative. This suggests that it is [+finite] I that assigns nominative case. In fact, it seems that it is the agreement part of it that does:

Portuguese:	Era	importante	eles	sair-em
	was	important	they	leave.infinitive-3pl

Let us now try to make sense of the fact that subjects receive their theta-role from the verb and their nominative case from I. Assume that just like DPs need a theta-role, they also need case.

Do all verbs have the exceptional case marking ability?

* We decided her to go.

* It seems her to go.

#4 Direct objects of nouns and adjectives are assigned case by dummy P of (unless of is genitive, assigned by A/N, in which case there is no theta/case discrepancy).

To summarize,

all arguments (=referential DPs and CPs) need a theta-role, and all (overt i.e. phonetically realized) DPs need case.

A head may assign both a theta-role and case to some DP, as verbs and prepositions do to their complement DPs.

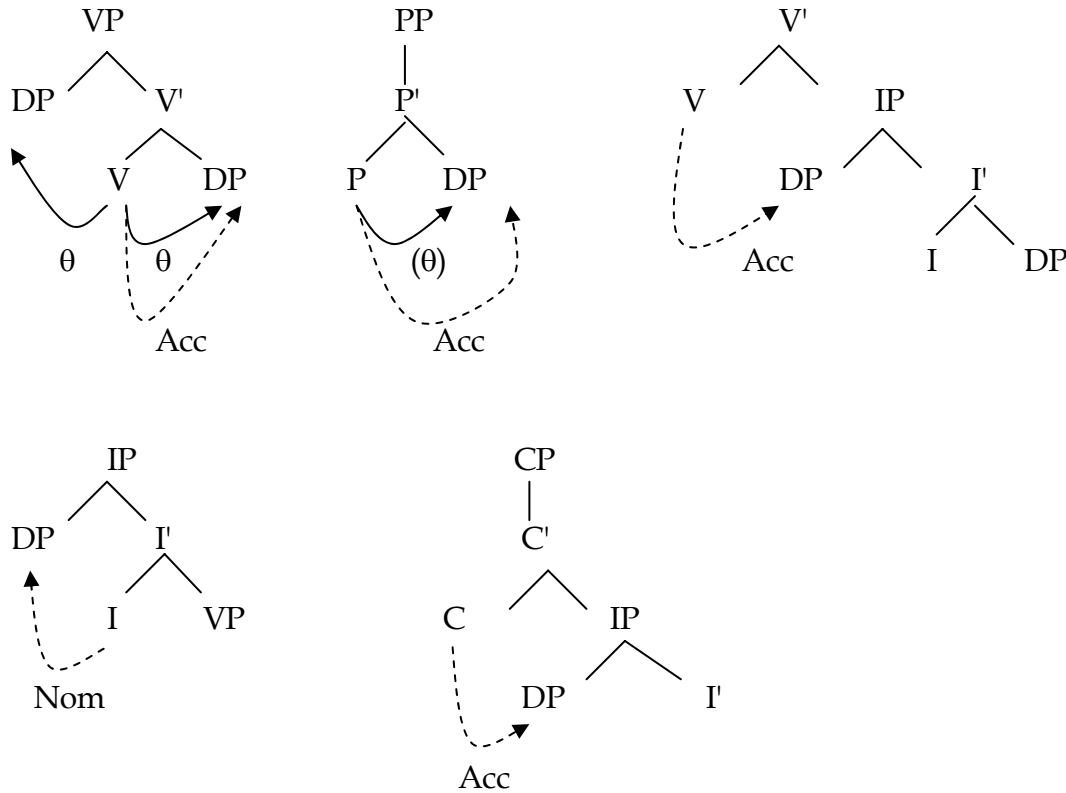
It is possible that a head assigns a theta-role but not case to some DP: lexical heads to their subjects, and possibly nouns/adjectives to their complements.

A head may also assign case to an DP it does not assign a theta-role to, as inflection, dummy of, complementizer for, and exceptional case marking verbs do.

The conditions are necessary, not sufficient conditions: X assigns a theta-role or a case to YP if they are in the right configuration **and** X has the requisite role or case in its lexical entry.

NB The definitions to follow count the complementizer for as a lexical head, i.e. P.

In what structural configuration can a head assign a theta-role and case a DP? The following diagrams summarize the legitimate possibilities in `Eighties talk. I leave out dative and genitive; these can be assigned in some of the environments accusative is.



Theta-role assignment and case assignment differ somewhat. They are both very local, but with slightly different upper/lower limits. A head can assign a theta-role to its own specifier and its own complement, and case to its own specifier, its own complement and also to the specifier of its complement.

Cases come in two flavors: structural and inherent. (By “case”, we mean case-type, not the case assigned by a particular head.) Nominative and accusative are structural cases, because they can be assigned by a head under particular structural conditions irrespective of whether the head also assigns a theta-role to the same position. The assignment of an inherent case is always accompanied by theta-role assignment. Dative is inherent and genitive can be either structural or inherent. Lexical categories can in principle assign either structural or inherent cases, but functional categories can assign only structural cases (since they never assign theta-roles).

Structural:	Nominative is assigned by	I (Agr)	under spec-head agreement.
	Genitive (str) is assigned by	D	under spec-head agreement.
	Accusative is assigned by	P or V	under government.
Inherent:	Dative is assigned by	P or V	under minimal m-command.
	Genitive (inh) is assigned by	A or N	under minimal m-command.

Definitions, FYI (plenty of work to define roughly three relations, with little insight into their nature and relationship -- AS)

C-command (p. 27) for anaphora, scope, etc.

A c-commands B iff
A doesn't dominate B and
every category dominating A dominates B.

X c-commands complement but not specifier (for X' doesn't dominate spec);
X' c-commands specifier;
XP c-commands things outside it.

Minimal c-command (p.39) for head movement, with a twist

A minimally c-commands B iff
A c-commands B and
there is no C such that C c-commands B but not A.

Government (p.64) for Accusative (can be ECM) by lexical head

Head A governs any B iff
A c-commands B and
no barrier dominates B but not A.
Barriers: any XP except IP

M-command (p.59)

A m-commands B iff
A doesn't dominate B and
some projection of A dominates B.

X and X' both m-command both complement and specifier (XP dominates them);
XP doesn't m-command anything.

Minimal m-command (p.59) for theta-role assignment and inherent Case ass.

A minimally m-commands B iff
M m-commands B and
there is no C such that C c-commands B but not A.

No other YP can intervene between X and minimal m-commandee.

Spec-head agreement (p.66) for Nominative/Genitive by functional head and many other things to come!

Head A agrees with B iff
A minimally m-commands B and
A doesn't govern B.

X agrees only with its specifier.

Movement motivated by case: Passive and bros.

Suppose there was a transformation that turned the active sentence He lifted her into the passive sentence She was lifted (by him) in one fell swoop:

He lifted her ==> She was lifted by him

This transformation would involve complex changes: (i) object becomes subject, positionally and case-wise, (ii) subject becomes optional or part of a by-phrase, (iii) verb changes into participle, (iv) be is inserted.

Instead, derive the two constructions separately, and factor things out: a modular approach. In this way, we are using modules that are already familiar from other areas or will prove to be useful in others in further work.

She was lifted (by him):

- Auxiliary be is chosen. The passive participle is selected by be.
- Participles, like adjectives, assign a theta-role but do not assign accusative case to their direct object.
- Of insertion is not available in participial VPs for some reason.
- The Case Filter now forces the direct object to move somewhere where it can receive case: specifier of IP.

Why is this movement possible? Why is the subject not competing for the same position? Assume that its theta-role (here: agent) is "absorbed" by participial morphology. The agent is either "implied" (as in She was lifted), or is expressed in an optional adjunct (She was lifted by him), where the agent role is assigned to DP by by. Then, the subject position is vacant and the direct object is free to move in.

This proposal receives support if its ingredients turn out to be useful in further cases that are only partially similar to passive. Here are some:

IP		The so-called "unaccusative" version
	I'	of <u>open</u> does not assign case to its
		complement.
I	VP	result: <u>The door opens</u> .
-s	V'	
	V	
	DP	
	open	the door

John turns the chair around -- The chair turns around
 John rolls the keg into the storeroom -- The keg rolls into the ...

"Existentials"

There is a man at the door (no theta-role for subject)
A man is at the door

"Raising"

It seems that John is a fool (no theta-role for subject)
* It seems John to be a fool (no exceptional case marking)
John seems to be a fool

Burzio's generalization:

A verb with a direct object does not assign accusative case iff
it has no theta-role for its subject.

(Of course, if V has no complement or its complement is not a DP, then not assigning accusative does not imply not having a subject theta-role, e.g. talk. But in fact, Vs that have no subject theta-role do not even participate in ECM, e.g. seem.)

In all cases that fall under Burzio's generalization, the accusative-less DP must make extra effort to receive some case. Normally, the specifier position of IP, where nominative can be received, is occupied by the subject. But these verbs have no theta-role for the subject, whence there is no DP obviously heading for the specifier of IP, and the case-less DP can move there. (If it does not, the specifier of IP will be filled by an expletive.)

NB A passive participle may well assign inherent case to a complement. I'm not sure why that complement tends to be preposed (topicalized?), as in Mihi invidetur.

Two theoretical issues concerning the properties of XP-movement:

1. The Projection Principle requires that lexical properties are syntactically reflected throughout the derivation. How do we know that this requirement is satisfied if the direct object moves away from its original position?

We assume that movement leaves a **trace**. Trace and moved expression form a **chain**, e.g. <she, trace>. Now the head of the chain has case and the tail has a theta-role. **The chain counts as one element.**

The Theta-criterion requires that each argument have one and only one theta-role, and no non-argument have any. Since each argument starts out in a position where it gets a theta-role (or else the Th-crit wouldn't be satisfied in initial structure), we can deduce a constraint on XP-movement. **Movement is possible only to a position where no theta-role is assigned.**

2. Can an argument move to any theta-free position? To see one aspect of what is at stake, consider the following sentence:

I persuaded John that there was a coin in my pocket.

Suppose we are trying to derive a variant of this sentence in which the main clause verb, persuade, is passivized. The initial structure is something like this:

___	was persuaded	<u>John</u>	that	___	was a coin in my pocket.
+case		-case		+case	
-theta		+theta		-theta	

In this structure, there are two spec of IP positions where John might escape to get case. One is in the main clause, the other is in the subordinate clause. (It is usually assumed that expletive it or there is only inserted at the end of the derivation, so these positions are vacant in the beginning.) We get the following results by exploring each of these possibilities:

(a) John was persuaded trace that there was a coin in my pocket.



(b)* It was persuaded trace that John was a coin in my pocket.



Why is (b) bad? It does not violate anything we said so far. The striking difference between (a) and (b) is that in (a), John moves upwards, but in (b), downwards.

Why is downward movement bad? We come back to this. One explanation can be in terms of semantically asymmetric dependencies, another in terms of the strict cycle.

Assignment, 10/11

Problem 1: Practice

- (1) It is unlikely for John to have been elected.
- (2) John is unlikely to have been elected.

Give very brief answers to the following questions and derive both sentences.

- (a) What head does John receive its theta-role from?
- (b) Does the same head assign a theta-role to its specifier? If yes, what, if no, why.
- (c) Does the same head assign case to John?
- (d) What is the complement of unlikely?
- (e) What word is the I in the infinitival subordinate clause?
- (f) Does this word assign a theta-role to its specifier?
- (g) Does this word assign case to its specifier?
- (h) What case does John have?
- (i) What head might assign that case to John?
- (j) In what syntactic position is John when it receives its case?
- (k) Why is that head able to assign case to John?

Problem 2: Research

- (3) There is/remains a book [on the table].
- (4) There are/remain books [on the table].

These sentences raise at least three interesting questions: What case does a book/books have? How does it get case? How do the verbs get to agree with a book/books?

First, write out clearly why these are problems.

Second, try to offer answers. On the table should figure in the solution only if it helps you; otherwise, feel free to ignore it.

The answers to the three questions should be compatible with each other.

In devising your answers, you can make theoretical innovations, but you should by default be conservative: conform to all the rules of the game we have so far set up (or Roberts discusses in Chs 1 and 2), and introduce as little new stuff as possible. If you introduce new ideas or revise existing ones, provide clear justification. – Note that in an empirical science, new facts do lead to theoretical innovations, but coherence of the grammar remains extremely important.

It is okay if you cannot solve the problem, but make sure that the outline of the issues and the preliminary considerations are of high quality.

Syntax I, 10-18-00
Szabolcsi

7. From Case Theory to Checking Theory

Rouveret and Vergnaud (1980) originally invented abstract Case to account for the fact that infinitival clauses cannot have phonetically overt subjects. More recently, Case theory has become the model of what drives movement: the moving element's need to be licensed, by "checking" some syntactic feature against a functional head. Hence Checking Theory in Minimalism.

- **All movement is motivated by licensing; all lexical categories need to be licensed. Licensing is feature checking. (The core idea, held by many, even if they do not believe in weak vs. strong features or economy, for instance.)**
- Features need to be checked and thereby eliminated, so that they do not interfere with phonological or semantic interpretation.
All non-interpretable features need to be eliminated by Logical Form (=input to semantics).
Strong features need to be eliminated before Spell-Out (=input to phonology). In other words, they trigger overt movement.
Weak features might be eliminated either before Spell-Out (overtly) or after, by LF (covertly). What forces them to be checked covertly is a principle of economy:
Procrastinate: Delay movement whenever possible.
- Checking domain of functional heads: specifier (for XPs) or adjunction (for Xs). The two might be unified either by (i) assuming that specifier = unique XP-adjunct (Kayne), or (ii) observing that in both, head X minimally m-commands but does not govern YP/Y (Chomsky/Roberts).

I strongly disapprove of option (ii). It uses very complex notions to define two extremely simple relations. Even more importantly, this "checking domain = minimal m-command minus government" definition is much like that of humans as featherless bipeds. Minimal m-command is the condition on theta-role assignment; government is the condition on e.g. accusative assignment. But checking is not "theta-role assignment, discounting accusative assignment." There is no content to defining checking domains using these two properties.

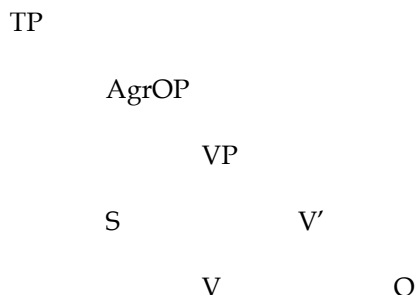
Case assignment under government is eliminated. New functional heads needed, e.g. AgrOP. Some gaps remain. Non-Case features treated the same way: stay tuned.

Inflection (I, short for the gang of AgrO, T, and AgrS) has both V-features (checked by verb movement) and N-features (checked by DP-movement).

The strength of each feature can be a matter of parametric variation. Feature strength tends to correlate with overt morphology in head position: overt morphology in head tends to trigger overt movement to it. (NB this correlation doesn't hold up with operator features.)

SVO, English: Finite I has all weak V-features \Rightarrow V doesn't move overtly.
 a strong N-feature for Nominative \Rightarrow S moves overtly.
 a weak N-features for Accusative \Rightarrow O doesn't move overtly.

AgrSP



How to get VSO and SOV from Kayne's universal initial SVO?

VSO, Welsh (see details in book)

SOV, German: Both S and O move out of VP to get licensed (hence Kayne's initial SVO ends up as SOV). CP and the object of prepositions do not move to get licensed, hence the "mixed" typological character.

SOV, Japanese: Do we have evidence that S and O move at all? How do tense, clause-type indicators, and subordinators all get suffixed onto the verb (does V move to C, getting to precede all dependents? do all affixes lower?) – **Ken Lacy presents Honda.**

Erez Levon presents Schuetze's template theory of Case.

Assignment: Finish Roberts' Ch. 2.

Read either Belletti, *The Case of Unaccusatives* (*Linguistic Inquiry* 19: 1-35, 1988) or Koopman and Sportiche, *The Position of Subjects* (*Lingua* 85: 211-58, 1991), and write a NELS/WCCFL abstract of it. The abstract can be one page (12 pt. font, 1 inch margins). You can put additional examples and references on a second page. See the current WCCFL call on LINGUIST. For models, you can check the web sites of previous NELS/WCCFL conferences; they may still post the abstracts. (If the paper has an abstract in the journal, cover the abstract before you write your own.)

Start out by stating the problem and the essence of the proposal. Then go on to lay out the proposal. Put the crucial examples on the first page where they play a role in the argument; only "additional" exx go on the 2nd page. You don't need to use a second page. (In real life, people often use the second page.) Never put your rules or generalizations on the second page. Use several paragraphs, ideally separated by blank lines.

Don't try to put everything you have to say about the subject into the abstract, but put in everything that is necessary for the idea to be understandable and convincing. An abstract is a sonnet: it has to be self-contained and sound good on its own.

Syntax I, 10-25
Szabolcsi

7 Binding

(we are doing only a fraction of Roberts' Ch. 3)

Skim Ch. 3 of Roberts (except for PRO/pro, which we come back to; pay attention to the three locality conditions and to 3.4 on long-distance anaphora).

Reinhart, Anaphora and Semantic Interpretation, 1983 -- **recommended**

Barss – Lasnik, A note on anaphora and double objects, LI 17: 347-354 (1986). -- **required**

Larson, On the double object construction. LI 19: 335-391 (1988) – **recommended**

We have proposed structures in which various DPs and other elements occupy specific hierarchical positions with respect to each other. In this section, we are investigating a set of phenomena that provide direct evidence for those **hierarchical relations**.

Reflexives and reciprocals (together: anaphors):

Anaphors require the presence of another DP they can co-refer with (=refer to the same person or thing). This is called the binder/antecedent:

John admires himself

John talks about himself

She turned John against himself

* Himself arrived

* She saw himself

We admire each other

We talk about each other

She turned them against each other

* Each other arrived

* She saw each other

The position of the antecedent DP matters:

* A sister of John(s) admires himself

* A woman that John met talks about himself

* John thought that Mary admired himself

* A sister of the boys admires each other

* A woman that the boys met talks about e.o.

Negative Polarity Items (NPI):

NPIs require the presence of a negative element that "licenses" them:

Nobody has ever been here

* Someone has ever been here

Nobody gives a damn about this

* Someone gives a damn about this

Nobody wants any more of the cake

* Someone wants any more of the cake

John has not ever been here
 John does not give a damn about this
 John does not want any more of the cake
 John does not like it much

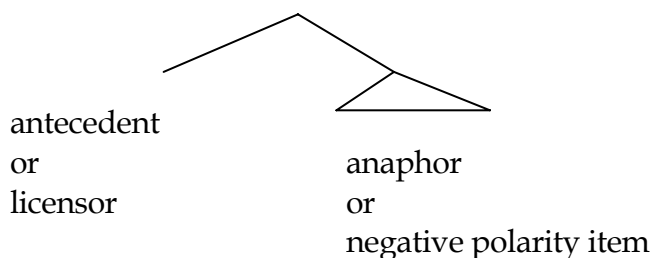
The position of the licenser negative element matters:

- * I have ever seen nobody
- * I have ever said that you did not show up
- * Any more of the cake is not on the table
- * Any more of the cake would interest nobody

There are at least two things that the antecedent--anaphor relation and the licenser--negative polarity item (NPI) relation have in common.

In terms of meaning, they are both **asymmetrical relations**: there are a dominant member and a dependent member in the relation. The antecedent is the dominant member with respect to the anaphor (it determines what the anaphor refers to), and the licenser is the dominant member with respect to the NPI: it enables the NPI to occur (be interpretable).

In terms of structure, they are characterized by the following **configuration**:



“First branching node c-command” (Reinhart 1976): Node A c-commands node B if and only if the first branching node above A is also above B (=dominates B).

When A c-commands B, A is at least as prominent in the structure as B (it is sister to, or higher than, B). Moral: The semantically more prominent member of a relation must be syntactically at least as prominent as the other member. This is Reinhart’s main insight.

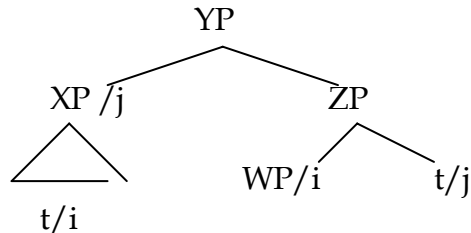
There are many other relations in grammar that require a c-command configuration; we will explore some of them. Here we point out just one:

When XP is moved and leaves a trace, the moved XP is obviously dominant with respect to the trace. As we saw last time, XP-movement is always "upwards". We can now make this requirement more precise:

Movement is to a c-commanding position:

Because the moved element is semantically more prominent than its trace, the moved element must c-command its trace.

This predicts that **remnant movement** is not possible. Here WP first moves out of XP, then the remnant XP moves to YP. WP does not c-command its trace inside XP after XP is moved up to YP:



There are different ways to remedy this situation. (i) By requiring that the remnant-moved category be put back into its original position for binding purposes (“reconstruction” or “the copy theory of movement”). In this case plain downward movement will be excluded, not with reference to binding, but with reference to the Strict Cycle. This is Gereon Mueller’s 1998 proposal. (ii) By denying that movement leaves a trace that counts as a free variable. (Traces, if they exist at all, count as locally lambda-bound variables.) If so, remnant movement is predicted to be unproblematic, and movement into a non-c-commanding position must be excluded with reference to some principle of tree-building; this is in a sense similar to appealing to the Strict Cycle. This is what categorial grammars as in Steedman/Szabolcsi/Jacobson/others say. – We may come back to these issues when we talk about wh-movement.

The fact that two expressions have the same reference is often notated by assigning the same index to them (**co-indexing**).

Generally, when a XP_1 c-commands another XP_2 and they refer to the same entity, we say that XP_1 **binds** XP_2 .

Besides government, binding is a pervasive and fundamental notion in the theory of syntax of the ‘Eighties. Hence its name: **Government--Binding Theory**.

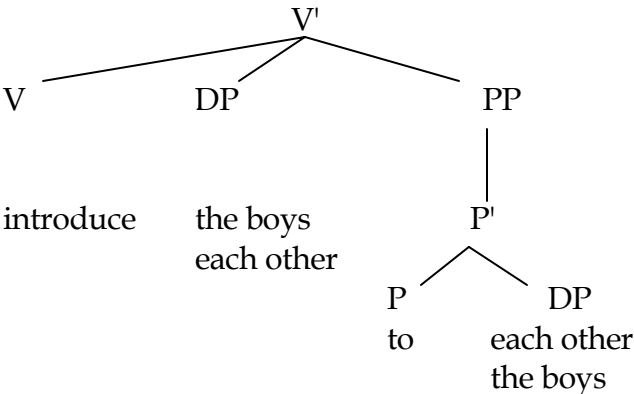
The c-command condition is often accompanied by some other condition, typically: **locality**. E.g. we have seen that the antecedent of a reflexive has to be (roughly) in the same clause as the reflexive, in addition to c-commanding it. The Binding Principles specify within what domain an expressions can find an antecedent. These are often referred to Principles A (for anaphors), B (for pronouns), and C (for referring expressions, names) -- see Roberts for details.

Binding as a test for hierarchical relations

Once we established that an anaphor must be c-commanded by its antecedent, we can use the antecedent--anaphor relation to test c-command relations. One area where this can be utilized is the structure of two-complement VPs. We already know that there are verbs that take two complements: put DP PP, introduce DP PP, show DP DP, show DP PP, compare DP PP, etc. Should they be represented with ternary branching?

The following contrast is already explained: when the antecedent is buried inside a PP, it cannot c-command an DP outside the PP:

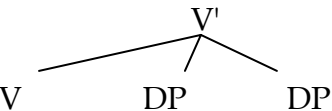
- (a) I introduced the boys to each other
- (b)* I introduced each other to the boys



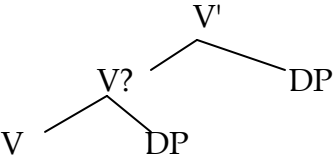
But what about the following?

- (a) I showed the boys each other (in the picture)
- (b)* I show each other the boys (in the picture)

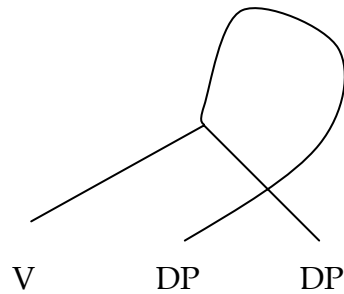
According to the structures we have so far, either of the two DP complements of show should be able to antecede the other: both DPs c-command the other:



What if we replace this ternary branching structure with the following?

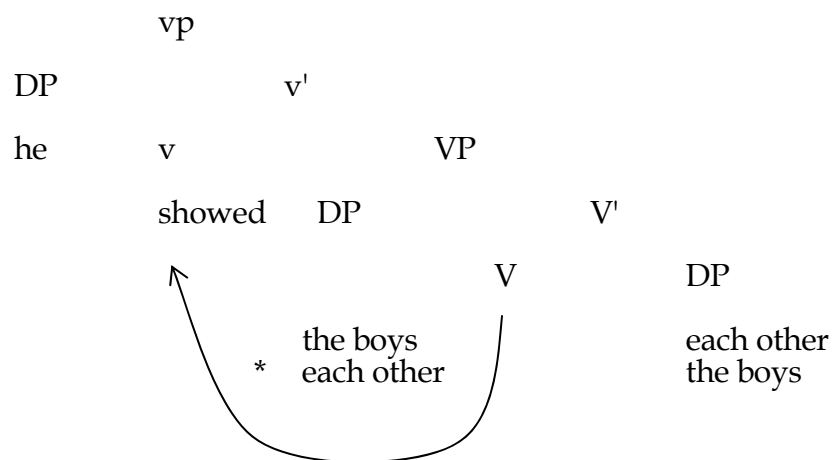


This revision predicts that the (a) sentence is bad and the (b) sentence is good; again an incorrect prediction. In reality, the first DP is more prominent than the second. This observation might be implemented by assuming discontinuous constituents, for instance:



One might assume that c-command and linear precedence together form a condition on binding (Barss and Lasnik 1986 entertain this as one possibility).

But in Larson 1988, a different structure is proposed: the VP looks something like this, with the verb moving from the lower VP to the higher vp:



With further arguments from the position of manner adverbs, even VPs with [V DP PP] are assigned a similar structure. In current work, refinements of Larson's analysis are used.

===

A summary of the conditions on XP-movement so far:

- XP-movement moves a constituent.
- If the XP is an argument, it must originate in a position where it receives a theta-role, and it can only move to another position where it does not receive any further theta-role.
- XP-movement is driven by some need, e.g. the need for DP to get case.
- XP-movement is to a c-commanding position.
- DP-movement must proceed in small steps: DP-trace is an anaphor.

Assignment

Coreference is between a name and a pronoun: John lost his hat `John lost John's hat'.

A quantifier can only *bind a pronoun as a variable*: Every boy lost his hat `every boy x, x lost x's hat' and not `every boy lost every boy's hat'.

Consider the following data and refute or defend each of the following hypotheses. Please mention in each case which example sentences are relevant, or particularly relevant to that hypothesis. In solving the above problem, you may rely on the insights the readings' discussion of pronouns, but please state your solution fully with reference to the examples and the hypotheses as I gave them.

Hypothesis One:

DP does not need to c-command a pronoun for the two to **corefer**. C-command is irrelevant to whether DP and pronoun corefer or not. Their coreference is not constrained by **any aspect** of sentence structure.

Hypothesis Two:

DP does not need to c-command a pronoun for the two to **corefer**. C-command is irrelevant for whether DP and pronoun corefer or not. Their ability to corefer depends solely on whether they are in the **same clause**.

Hypothesis Three:

DP does not need to c-command a pronoun for the two to **corefer**. But both c-command and clause-mateness are relevant for whether DP and pronoun **can** co-refer, **namely ...**

Hypothesis Four:

The **bound variable** interpretation of pronouns is possible under the **same** conditions as the coreferential interpretation.

Hypothesis Five:

The **bound variable** interpretation is available in a **proper subset** of the environments where the coreferential interpretation is possible, **namely...**

Hypothesis Six:

The **bound variable** interpretation is available (according to the data given) in the following configuration (not excluded above): ...

Data:

Coreferential: (please refer to these as (Ca), etc.)

- | | |
|---|-----------------|
| a. John admires him | (* John = him) |
| b. John talks about him | (* John = him) |
| c. That Mary saw John upset him | (ok John = him) |
| d. That John left cast a bad light on him | (ok John = him) |
| e. A woman that John met talks about him | (ok John = him) |
| f. My conversation with John shocked him | (ok John = him) |
| g. He admires John | (* John = him) |
| h. He talks about John | (* John = him) |
| i. John thinks that Mary saw him | (ok John = him) |
| j. He thinks that Mary saw John | (* John = he) |
| k. John thinks that he is a genius | (ok John = he) |
| l. John lost his hat. | (ok John = he) |

Bound variable: (please refer to these as (Va), etc.)

- | | |
|--|----------------|
| a. Every boy admires him | (* bound him) |
| b. Every boy talks about him | (* bound him) |
| c. That Mary saw every boy upset him | (* bound him) |
| d. That every boy left cast a bad light on him | (* bound him) |
| e. A woman that every boy met talks about him | (*bound him) |
| f. My conversation with every boy shocked him | (*bound him) |
| g. He admires every boy | (* bound him) |
| h. He talks about every boy | (* bound him) |
| i. Every boy thinks that Mary saw him | (ok bound him) |
| j. He thinks that Mary saw every boy | (*bound he) |
| k. Every boy thinks that he is a genius | (ok bound he) |
| l. Every boy lost his hat. | (ok bound he) |

Syntax I, 11-1-00
Szabolcsi

8 Raising vs. control, PRO vs. pro

Read the pertinent parts of Roberts' Ch.3 (finish the binding parts too if left over)
Rizzi, Null objects in Italian and the theory of *pro*. LI 17:501-57 (1986) -- recommended
Cinque, On *si* constructions and the theory of *arb*. LI 19:521-583 (1988) -- recommended

Seem, be agreed, and be likely may take a finite clausal complement, see (a). They also occur with an infinitival complement (c), but this may not have an overt subject (John), see (b). Where does John originate in (c)?

- (a) It seems that John has a fever.
- (b) * It seems John to have a fever.
- (c) John seems to have a fever.

- (a) It is thought that John is a good violinist.
- (b) * It is thought John to be a good violinist.
- (c) John is thought to be a good violinist.

- (a) It is likely that John comes late.
- (b) * It is likely John to come late.
- (c) John is likely to come late.

We assume that the infinitival complement is a clause and John is its subject. Seem, passive participle thought, and likely conform to Burzio's generalization: they do not assign a theta-role to their own specifier and they do not assign accusative case to any DP (notably, to the subject of their infinitival complement). John does not receive nominative case in the complement, because infinitival I (to) does not assign nominative. It is then natural to derive (c) by moving John from the complement subject to the matrix subject:

John seems trace to have a fever.

This is exactly like a chain formed within one clause, although it spans two clauses. This kind of case-related movement is called **(subject-to-subject) raising**.

Raising vs. control

Initially, the following two structures look alike:

John seems to have a fever.
John likes to have a fever.

But there are striking differences. There is never an expletive in the subject position of like:

- It seems that John has a fever.
 * It likes that John has a fever.

Likewise, whenever DP is a reasonable subject of a particular VP, the construction DP seems/is likely to VP has a coherent interpretation. But DP likes/tries to VP may not have a coherent interpretation:

The weather is good.	The weather is good.
The weather seems to be good.	?? The weather likes to be good.
The weather is likely to be good.	?? The weather tries to be good.

In other words, in DP seems/is likely to VP, it is enough if DP is compatible with the infinitival VP. In DP likes/tries to VP, DP needs to be compatible **both** with the infinitival VP and with like/try.

These observations suggest that we are dealing with two different constructions. In the raising construction, the predicates seem and likely do not assign a theta-role to the DP that appears as their subject; in the other construction, which is called control, the predicates like and try do assign a theta-role to that DP.

It is unproblematic to say that in the raising construction, DP got into the subject position of seem, etc. by DP-movement motivated by the need for DP to receive case. In what follows we will focus on how the control construction works.

Control

What is the category of the complement of like in John likes to have a fever? We will argue that, the differences from raising notwithstanding, this complement is a clause which has an invisible DP subject called PRO (read: "big pro").

	IP		==>		IP
	I'			DP	I'
I	VP			PRO/i	I
to	DP	V'		to	DP
	PRO	have a fever		t/i	have a fever

The reason to have IP is this. We have seen that it is useful to assume that infinitival to is

I(nfl). Then a phrase containing to cannot be smaller than IP; it cannot be a plain VP, for instance.

Next, we turn to justifying the non-overt pronoun PRO. We have three questions to answer. First, what motivates PRO? Second, to capture the distribution of PRO: in what positions may PRO be found? The third requirement concerns the interpretation of PRO: how is PRO interpreted. The first and the third questions will be dealt with in these notes. As regards the second question, we will simply state this:

PRO can only occur as a subject (but never as a complement). It moves from specifier of VP to the specifier of infinitival inflection because of the extended projection principle, which requires that each clause have a subject.

The motivation for PRO

The motivation for the existence of PRO can be obtained from several of the modules we have so far developed.

The theta-criterion requires that a lexical head assigns each of its theta-roles to an argument, and that each argument end up with exactly one role. Consider now the examples in (1):

- (1) a. John knows [how he should behave himself]
 b. John knows [how PRO to behave himself]

In the finite embedded clause in (1a), *behave* has a theta-role to assign, and this role is assigned to *he*. Note that this role cannot be assigned to *John*, as *John* already bears a theta-role provided by *know*, and is too far up anyway. In (1b), then, the predicate *behave* likewise has a theta-role to assign, which it can again not assign to *John*. If we assume PRO, PRO can get this role, in accordance with all the rules of our game. **In sum, PRO is needed so that behave can assign its agent role to an argument.**

Binding theory deals with possible dependencies between the interpretations of DPs. In particular, the binding theory as fleshed out for English stipulates that anaphors must be bound within their own clause, and that pronouns must not be so bound. This accounts for the following facts:

- (2) a. John_i washes himself_i / *him_i
 b. John_i thinks that I like him_i / *himself_i

Let us next review the following facts;

- (3) a. John_i promised you to take better care of himself_i
 b. You persuaded John_i to take better care of himself_i

At first blush, it seems that John is linked to himself directly. This contradicts the assumption that to take better care of himself is a clause -- himself seems to be bound by an antecedent outside its clause. This problem would not arise if the infinitival clause had a silent PRO subject, because PRO could serve as the antecedent for himself.

- (3) a. John promised you [PRO_i to take better care of *himself*_i]
 b. You persuaded John [PRO_i to take better care of *himself*_i]

If so, the relation between John and himself can at best be indirect. What we need to account for is how John and PRO are related. The suspicion that PRO mediates between John and himself is corroborated by the following examples:

- (4) a. *John_i persuaded you to take better care of *himself*_i
 b. *You promised John_i to take better care of *himself*_i

Why are these sentences bad? (3) might lead us to think that both occurrences of John in the main clause are good antecedents for himself. But notice that in (3), John is understood to be the agent of taking better care, that is, John is linked to PRO in its own way. On the other hand, in (4) you are understood to be the agent of taking better care, that is, John is not linked to PRO. To summarize, main clause John is linked to subordinate clause himself only if John provides the reference of the PRO subject of the subordinate clause.

This leaves us with the last question:

The interpretation of PRO: #1 control

In the above sentences, PRO is like an anaphor in that it fully depends on another DP for its interpretation. But PRO is not like an anaphor in that its antecedent, although it c-commands PRO, is not within the same clause. (It cannot be, since PRO is the most prominent DP within its own clause.) PRO is also not quite like an anaphor in that it does not always seem to have an antecedent:

- (5) [PRO to go to the beach] is fun.
 It is fun [PRO to go to the beach].

But this latter is not quite true. Because it is understood that the one who goes to the beach is the same as the experiencer of the fun. You might think this simply follows from the nature of the situation, but observe:

- (6) [PRO to vanish] would be devastating.
 It would be devastating [PRO to vanish].

It is pragmatically quite possible for it to be devastating for person X that person Y vanishes, compare: For John to vanish would be devastating to Bill. But the sentences in (6) cannot mean that. We conclude that when PRO does not appear to have an antecedent, its antecedent is in fact an "implicit argument" (typically: experiencer) of the main clause predicate.

Because the interpretation of PRO does not work in entirely the same way as that of anaphors, we give the mechanism a separate label: **control**. The DP that serves as the overt or implicit antecedent of PRO is called its controller. The lexical entry of control verbs specifies which argument will control the subject of the infinitive:

Mary promised John [PRO to leave]	subject control verb
Mary persuaded John [PRO to leave]	object control verb

The interpretation of PRO: #2 "arbitrary"

When PRO does not have an overt controller, it does not necessarily have an implicit controller. It may have a so-called arbitrary interpretation (PROarb):

I don't know [what PRO to do].

In this example, arb=generic. When we turn to pro, we'll see that arb can be existential too.

Structures: IP or CP?

Verbs that **exceptionally case mark** the complement subject must be taking IP (not CP) complements, or else they could not govern that subject:

ECM:	V'	<u>want</u> governs <u>him</u>
	V	IP
	want	DP I'
		him to be here

Now consider **raising**. Conforming to Burzio's generalization is unfortunately not enough to ensure that a predicate allows raising:

- (a) It is probable that John has a fever.
- (b) * It is probable John to have a fever.
- (c) * John is probable to have a fever.

We may try to make a distinction between the raising predicate likely and the non-raising predicate probable in such a way that the trace of John in the specifier of the infinitival IP is legitimate (licensed) with likely but not with probable. The proposal in the 1980's uses the following general assumption:

Trace must be governed by a lexical head.

We can then distinguish the two kinds of predicates by saying that likely can have an IP complement, but probable must have a CP complement (possibly with an empty C). The CP level will prevent probably from governing the specifier of IP:

Raising:

A'		<u>likely</u> governs <u>trace</u>
A	IP	
likely	DP	I'
	trace	to be here

Non-raising:

A'		<u>probably</u> does not govern <u>trace</u>
A	CP	
probable	C'	
	C	IP
	DP	I'
	* trace	to be here

Finally, consider **control**. Recall the claim (which we did not examine in detail) that PRO must be a subject, it cannot occur in the complement position. This can be ensured if we stipulate the following:

PRO must not be governed by a lexical category.

Itt nem pro járnak biciklin. (generic tense, generic pro)
 here not go-3pl bicycle-on
 ‘People here do not use bicycles’

pro Kopogtak - vajon Mari az? (episodic tense, existential pro)
 knocked-3pl perhaps Mari that
 ‘Someone knocked, is it perhaps Mari?’

It is interesting to note that crosslinguistically, all “arb” expressions are [+human]: PROarb, proarb, one, Mann, etc. This makes it easy to diagnose an empty category as one of these guys.

Case

We now see that either PRO or trace can occur in the subject of infinitives. But pro cannot. The final form of the Case Filter is this:

An overt DP and pro must receive abstract Case.

Laure Lim presentation: Acquisition and pro-drop (based on Hyams chapter in Fromkin)

Assignment:

Blurb about LI, NLLT, TLR, Language, Syntax, Ling. Anal., Lingua, Journal of Linguistics.

Syntax I, 11/8/00
Szabolcsi

9 Wh-movement: General notions

Reading: Roberts, Ch. 4, skipping 4.4. Barriers and 4.6 Locality.
van Riemsdijk – Williams: Subjacency (Introduction to..., 1985).

We have considered two kinds of movement: head movement and XP movement, both structure preserving. Within the XP movement class, we looked at DP movement, motivated by the need for DP to get case. The last phenomenon we are looking at is another major type of XP movement, called wh-movement.

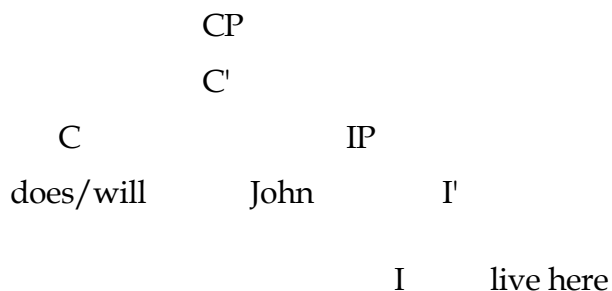
- (1) (a) John read this book
(b) Which book did John read?
(c)*John read which book? [only as an "echo-question"]
- (2) (a) John lives in London
(b) Where does John live?
(c) *John lives where? [only as an "echo-question"]

As examples (1) and (2) show, which book and where must occur in a sentence-initial position. This cannot be the product of phrase structure rules (because they are context-free), it must be the result of movement. But this movement is not motivated by the need to get case, because which book originates in a position where it gets accusative, and where does not need case at all. It appears that this movement is driven by nothing else than the fact that these phrases are question words (wh-phrases).

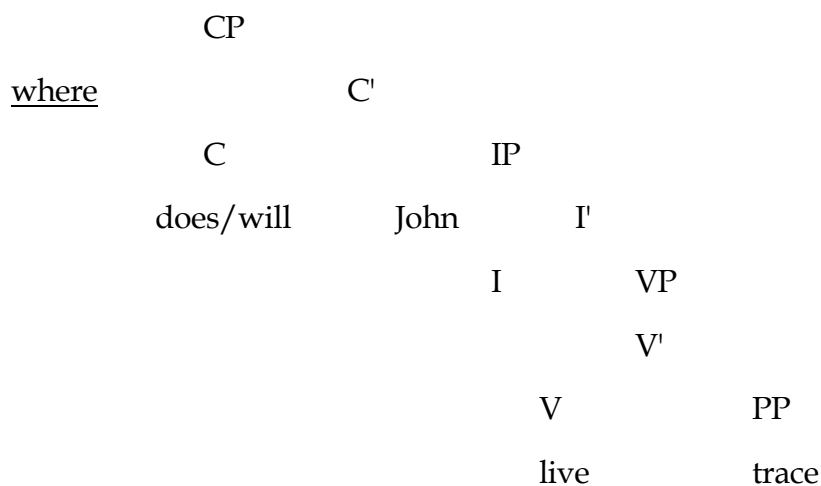
What position do question words move to? Focus on the underlined elements:

- (3) Where does John live? Does John live here?
Where will John live? Will John live here?
Where was John seen? Was John seen here?

We analyzed yes/no questions with the auxiliary in C:



In both which book and where-questions, the auxiliary precedes the subject exactly as it does in yes/no questions, suggesting that it is in C. The wh-phrase precedes the auxiliary, suggesting that the wh-phrase is in the specifier of CP:



Wh-movement and subcategorization

Do question words always move to the specifier of CP within their own clause?

- (4) (a)* John wonders [that I live here].
 (b) John wonders [whether I live here].
 (c) John wonders [where I live].
 (d)* Where does John wonder [(that) I live]?
- (5) (a) John believes that I live here.
 (b)* John believes whether I live here.
 (c) * John believes where I live.
 (d) Where does John believe (that) I live?
- (6) (a) John wonders where Sue think that Bill said that I live.
 (b) Where does John believe that Sue think that Bill said that I live?

These examples show that how far a question word moves depends on what kind of verb governs the CP containing that question word. A verb like wonder requires that its complement be a question (witness: wonder can have a whether-clause but not that-clause complement).

wonder, V, [___ CP_{+interrogative}]

If CP is the complement of wonder, the question word must stop in the specifier of this CP, see (4). A [+interrogative] CP must contain either if/whether in its head or a wh-

phrase in its specifier. (This leaves main clause yes/no questions unaccounted for: they have neither. Maybe they have an empty if-like element.)

A verb like believe, think, or say does not take questions as complements (witness: they can have a that-clause, but not a whether-clause, as a complement).

believe, V [__ CP_{+declarative}]

If CP is the complement of believe, etc., the question word must move out. A [+declarative] CP must not contain if/whether in its head or a wh-phrase in its specifier. The wh-phrase keeps moving until it finds either a verb like wonder, which takes a question complement, or a main clause, which may always be a question. See (5).

So-called “partial wh-movement” (German, Hindi, Hungarian, etc.) is interesting, among many other things, because it seems to contradict this regularity:

Mit gondolsz, hogy mikor aludt el a gyerek?
 what think-you that when fell-asleep the child
 ‘When do you think the child fell asleep?’

See Dayal 1996 (Natural Language Semantics) and Horvath 1997 (Natural Language and Linguistic Theory) for arguments against assuming that mikor ‘when’ moves covertly to replace “expletive” mit ‘what’.

The examples in (6) show that such movements may span several clauses.

Wh-movement and cross-linguistic variation. The wh-criterion

The claim that wh-phrases must occur in the specifier of CP is too simplistic. Consider:

E: John read this book in London
 Where did John which book?
 ? Which book did John read where?
 John read which book where [at most as an "echo-question"]

F: Qui as-tu vu?
 whom have-you seen
 Tu as vu qui?
 you have seen whom

R: Kto kogo videl?
 who whom saw

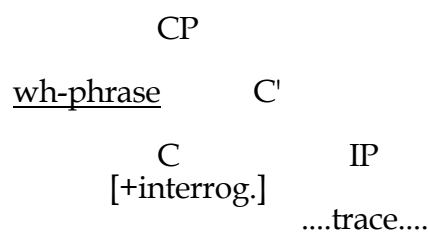
In English, when a clause contains more than one question word, only one need to (and

indeed, only one can) move to the specifier of CP. In French, a question word may be left in its original position even if it is the only question word of the sentence (and in Chinese, it must always be left there). In Russian, normally all question words move to the front.

Such cross-linguistic variation raises the issue why question words move at all. If we take French (or Chinese) to be the "ideal" case, the fact that some question words move seems entirely superfluous. If however Russian is the "ideal," all non-movement examples seem fatally defective.

Wh-Criterion (Rizzi 1991):

A [wh] head must have a [wh] XP in its specifier at some level;
a [wh] XP must appear in the spec of a [wh] head at some level.



Whether Russian or Chinese is regarded as "ideal" depends on whether the grammar includes, additionally, economy principles like Procrastinate. – See further Watanabe 1992, Lisa Cheng's dissertation, recent work by Z. Boskovic, etc.

What moves and where: Successive cyclicity, the strict cycle, the ECP, and superiority

Successive cyclicity. Wh-movement seems to be able to span several clauses in one swoop (thus is sometimes called "unbounded"). One kind of evidence that it really does not do so comes from anaphora interpretation:

- a. Which book about **herself** does **Madonna** like ___?
- b. Which book about **herself** does **Madonna** think [___ that we will publish ___]?

Sentence (a) requires that we interpret [...book about herself] in the initial ___ position, because that is where it is c-commanded from close by Madonna. This is taken care of by reconstruction or the copy theory of movement. But (b) requires something even funnier. Its acceptability is understood only if we assume that [...book about herself] has a chain link in a position that is roughly in the Spec of the embedded CP, where it is c-commanded by Madonna from close. (Barss dubbed this chain binding.)

Another kind of evidence comes from Spanish, where the intermediate clauses exhibit subject/aux inversion just like the clause where the wh-phrase lands. This suggests that it transits through Spec-CP in the intermediate clauses (Torrego, LI 1984), causing I-to-C:

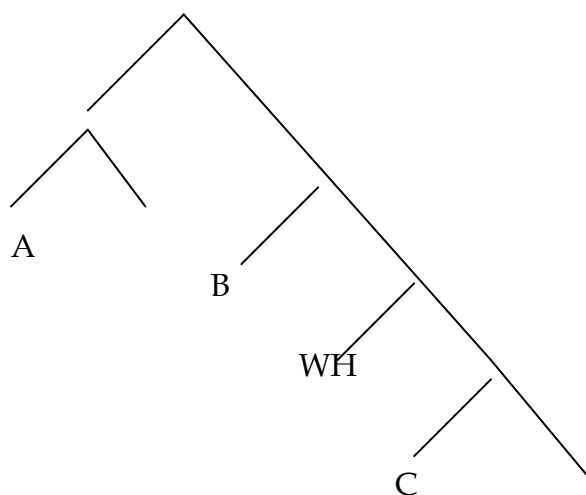
- a. Qué pensaba **Juan** que le había dicho **Pedro** que había publicado **la revista**?
 'What did J think that P had told him that the journal had published?'
 b. *Qué pensaba **Juan** que **Pedro** le había dicho que **la revista** había publicado?

Yet another kind of evidence comes from the observation that when the Spec-CP of an intermediate clause is filled (by another wh-phrase), wh-extraction is degraded:

?What film do you wonder [when to see ___]?

Wh-movement is assumed to move **successive cyclically** (i.e. from Spec-CP to Spec-CP or in even smaller steps), and thus turns out to be similar to head movement and DP-movement in always proceeding to the closest available landing site.

The strict cycle. Another constraint is that wh-movement (as all other transformations) observe the **Strict Cycle** condition. This has various formulations. One is this: once you left a domain (say CP) you cannot go back and move something so that it lands inside that domain. In Minimalist terms, where trees are built bottom-up by Merge and Move, this principle is captured by Chomsky's **Extension Condition**: each step of the derivation must extend the tree. Suppose you already built a tree like this and want to move WH:



The Extension condition says that neither A, nor B, nor C are possible landing sites for WH (or anything else in its place, for that matter): these movements will not extend the tree. The only way to move WH is to build a new landing site on top of what we have and move right into it.

The Strict Cycle was initially used to account for certain wh-island phenomena (the case of moving to position B, where B is not topmost), but it is much more generally useful. The fact that WH cannot move to position C corresponds to the fact that there is no lowering movement. The fact that WH cannot move to A corresponds to the fact that even upward

movement is out if the target does not c-command the launching site.

One advantage of obtaining the c-command condition on movement in this way is that it allows **remnant movement**. In remnant movement, all movements are to c-commanding positions, although a trace may end up not c-commanded by its binder. This is now perfectly acceptable if the resulting structure is interpretable. It may be interpretable (i) if the trace is treated as a free variable in need of binding, but interpretation is aided by reconstruction/copies of the moved element, or (ii) if the trace is treated as a locally lambda-bound variable, in which case remnant movement is not an interpretive problem.

The Empty Category Principle. Both these conditions pertain to where things move. But what can move at all? We have already assumed that traces need to be governed by a lexical head, to account for the distribution of DP-trace versus PRO in the infinitival subject position. This predicts the following:

- (a) Complements (of V, P, ...) can move.
- (b) Spec-IP cannot move when a CP-layer is present.
- (c) Adjuncts cannot move.

In reality, subjects and adjuncts can move, although sometimes with some difficulty. Subjects in languages like English suffer from the “overt C–trace” effect in the source clause:

- Which contestant do you believe (*that) __ won?
- Which contestant (*did) __ win?

Once a subject gets out of its own clause, though, it moves as easily as a direct object:

- ?Which contestant do you wonder whether Mary believes __ is winning?

On the other hand, adjuncts do not mind complementizers but are more sensitive to islands (e.g. wh-islands) than subjects or objects (=arguments) are:

- How do you think that I told Mary that she was behaving __?
- *How did you ask whether I told Mary that she was behaving __?

Thus, lexical government is sufficient but not necessary.

Empty Category Principle: A trace must be properly governed, viz.,

- (i) governed by a lexical head, or
- (ii) antecedent governed, i.e. not separated from a co-indexed c-commander by any “barrier”.

The idea is that if the next link in the movement chain is sufficiently close to a subject/adjunct trace, it antecedent-governs the trace and this is enough. In the first step, both subjects and adjuncts must use the Spec-CP of their own clause as an escape hatch. The story then gets complicated because, as the above description demonstrates, subjects and adjuncts are sensitive to quite different things. Lasnik and Saito's On the nature of proper government (1984) and Chomsky's Barriers (1986) try to unify these and the result is in no way beautiful. More promising is Rizzi's Relativized Minimality (1991), where the two are essentially separated. The "C-t" effects for subject extraction are dealt with by "the conjunctive ECP", whereas the difficulties with adjunct extraction fall together with the intervention effects causing various different kinds of Weak Island violations.

Superiority. There are constraints on what is movable at all (in an absolute sense), but there is a further curious condition. When either of two things might move, the higher one must move. This condition is called **Superiority**:

initial: you told **who** to read **what**
 * What/i did you tell who [to read ___/i] ?
 Who/i did you tell ___ /i[what to read]?

Pied piping

When a phrase moves to check a feature, it sometimes drags the constituent that contains it along. Pied piping may be obligatory (a) or optional (b):

(a) Whose brother did you visit? compare: *Whose did you visit brother?
 (b) To whom are you talking? compare: Who are you talking to?

The best behaved and most typical case is when the larger phrase inherits the relevant feature via spec-head agreement, roughly: D agrees with who and is thereby [+wh]; the whole DP whose brother has the features of its head and is thereby [+wh]; the whole DP moves to Spec-CP.

	DP	
who	D'	
	D	NP
	's	brother

Features of the complement are rarely inherited by the whole phrase and need a different mechanism. Thus to whom is exceptional. Feature inheritance from a complement is freer in relatives than in interrogatives:

These are the documents the size of the lettering on the cover of which the government prescribes.

*The size of the lettering on the cover of which documents does the government prescribe?

Pied piping has a great significance in current theories where not only wh-phrases but many other phrases are assumed to be involved in pied piping. Essentially, the (as yet not fully known) rules of pied piping determine what size constituents move.

Wh-movement = a family of movements

What we called DP-movement was, descriptively, a family of movements: passive, raising, and the treatment of unaccusatives all involved a movement motivated by the Case Filter. Is wh-movement instigated by a single operation, the one that fronts question words, or is it also a family of operations?

The so-called island constraints on wh-movement can be used as a diagnostic. We may assume that if an operation is similar to question word fronting in that its motivation is meaning-related, not case-related, and it is subject to the same constraints, then it is an instance of the same general mechanism question word fronting is. (We come back to the island constraints below.)

One obvious candidate is relative clause formation (the fronting of the relative pronoun) in English.

a man <u>who</u> likes Bill	a man who I think likes Bill
a man <u>whom</u> Bill likes	a man who I think Bill likes
a man <u>to whom</u> Bill sent a letter	a man to whom I think Bill ...
a place <u>where</u> I want to live	a place where I think I ...
a reason <u>why</u> I cannot live there	a reason why I think I ...

We demonstrate that relativization is subject to the Coordinate Structure, Complex DP, and wh-island constraints:

- CS
- Bill likes this guy and my mom likes Bill.
 - * The guy who Bill likes ___ and my mom likes Bill called.
 - The guy who Bill likes ___ and Peter hates ___ called.
- CDP
- You heard the rumor that they fired this guy.
 - * I met the guy who you heard the rumor that they fired ___.
- WH
- You read the book which I got from this guy.
 - * I met the guy who you read the book which I got from ___.

Assignment

(i) Read Roberts' 4.3 ECP, up to 4.3.3 (connectedness and parasitic gaps) and Ch.2 (The conjunctive ECP) of Rizzi's Relativized Minimality.

(ii) Using the data below, propose a structure for Hungarian noun phrases. Discuss the parallelisms, if any, with clausal structure. Develop an analysis for possessor extraction, making use of the ECP. The core data are on this page; the rest are supporting data.

(iii) Evaluate Rizzi's ECP in comparison with other theories (as discussed in Rizzi and in Roberts) in the light of Hungarian possessor extraction.

(1)

a hat fekete kalap-ot

the six black hat-acc

'the six black hats (accusative)'

(FYI; don't analyze numerals and adjectives)

(2)a.

a te kalap-ja-i-d

the you hat-poss-pl-2sg

'your hats'

(2)b.

(a) Mari kalap-ja-i

the Mari hat-poss-pl(-3sg)

'Mari's hats'

(3)

Mari-nak a kalap-ja-i

Mari-dat the hat-poss-pl(-3sg)

'Mari's hats'

(4)a.

* Mari volt fekete a kalap-ja

Mari was black the hat-poss(-3sg)

'It was Mari whose hat was black'

(cf. 2b)

(focus; similarly for topic)

(4)b.

Mari-nak volt fekete a kalap-ja

Mari-dat was black the hat-poss(-3sg)

'It was Mari whose hat was black'

(cf. 3)

(focus; similarly for topic)

(5)

a te valamennyi titk-od

the you each secret-poss.2sg

'your every secret'

- (6)
 (a 'the') Mari { minden 'every'
 e, eme, ezen 'this'
 ama, azon 'that'
 melyik 'which'
 keves 'few'
 sok 'many'
 egy(ik) 'one'
 valamennyi 'each'
 bármelyik 'either'
 semelyik 'neither'
 * a(z) 'the'
 * az a(z) 'that the'
 * ez a(z) 'this the' }
- húg-a 'sister-infl'
 szav-a 'word-infl'
 kalap-ja 'hat-infl'
 kutyá-ja 'dog-infl'
 ...

- (7)
 Kati **ki-nek a kalap-já-t** látta?
 Kati who-dat the hat-poss.3sg-acc saw
 'Whose hat did Kati see?'

- (8)a.
 Kati (nem) csak **Mari-nak a kalap-já-t** látta
 Kati not only Mari-dat the hat-poss.3sg-acc saw
 'Kati saw (not) only Mari's hat (but...)'

- (8)b.
 Kati a sál-at és **Mari-nak a kalap-já-t** látta
 Kati the shawl-acc and Mari-dat the hat-poss.3sg-acc saw
 'Kati saw the shawl and Mari's hat'

- (9)a.
 * ki kalap-ja
 who(-nom) hat-poss.3sg

- (9)b.
 ki-nek a kalap-ja
 who-dat the hat-poss.3sg
 'whose (interrog.) hat'

and similarly for aki 'who, relative,' melyik 'which one,' mindenki 'everyone,' senki 'no one,' etc.

- (10)
 Kati **kinek** látta a kalap-já-t?
 Kati who-dat saw the hat-poss.3sg-acc
 'Whose hat did Kati see?'

(11)a.
 az én kalap-om az én kalap-ja-i-m
 the I hat-poss.1sg the I hat-poss-pl-1sg
 'my hat' 'my hats'

(11)b.
 *én kalap-om *én kalap-ja-i-m
 I hat-poss.1sg I hat-poss-pl-1sg

(12)a.
 *Az én isz-om
 the I(-nom) drink-1sg

(12)b.
 Én isz-om
 I(-nom) drink-1sg
 'I drink'

(13)a.
 a János kalap-ja
 the Janos hat-poss.3sg
 'Janos's hat'

(13) =Upper Tisza dialect

(13)b.
 * János kalap-ja
 Janos hat-poss.3sg

(13)c.
 * A János isz-ik
 the Janos drink-3sg

(13)d.
 János isz-ik
 Janos(-nom) drink-3sg
 'John drinks'

If you think you crucially need further specific pieces of data, email me.

Syntax I, 11/28
Szabolcsi

10 Strong and Weak Islands (excerpts from my SynCom case)

The fundamental work on islands is Ross 1967/1986. Ross distinguished between transformations that move a constituent and leave behind a pro-form (i.e., resumptive pronoun) and ones that "chop" a constituent and move it without leaving anything behind. An example of the former is left dislocation (that is, on Ross's analysis, left dislocation is not base generated). An example of chopping is topicalization. Of the two, only chopping operations are constrained by islands. Islands will be enclosed in angled brackets:

- (1) This kid, I must call <the teacher who punished him>.
(2) * This kid, I must call <the teacher who punished _>.

Diagnosing islands has become more complex since Ross, because it is assumed that there exist invisible pro-forms (empty resumptive pronouns), in addition to overt ones. I will use "gap" as a pretheoretical term that covers both traces and empty pronouns.

Islands come in two main varieties: "weak" and "strong". Note that the weak/strong qualification does not correspond to degrees of ungrammaticality.

- (3) Strong islands:
No extraction is allowed (constructions with an appropriate resumptive pronoun may be allowed).
- (4) Weak islands (WI):
Some phrases can extract, others cannot extract.

Up until the late Eighties, nothing much beyond wh-islands had been thought to be a weak (selective) island. Beginning with Relativized Minimality, however, an ever growing range of weak islands have been recognized. Thus, theories of weak islands have mushroomed, each coming with a significant set of new data and important new connections to other domains.

2. Absolute versus selective islands

Although the notion of strong versus weak islands as defined in (3) and (4) seems like a simple descriptive one, it is not. Consider the following examples involving wh-islands and adjunct islands:

- (5) a.* About which topic did John ask <who was talking _>?
b.*? Which topic did John ask <who was talking about _>?
c.* How did John ask <who behaved _>?

- (6) a.* About which topic did you leave <because Mary talked _>?
 b.* Which topic did you leave <because Mary talked about _>?
 c.* How did you leave <because Mary behaved _>?

These two sets might suggest that both wh-clauses and adjunct clauses are strong islands: the extraction of three different types of phrases was attempted and failed. On the other hand, it is easy to construct other examples in which some, but not other, phrases can extract. The percentage sign % indicates that there is variation among the speakers of English:

- (7) a.% About which topic did John ask <whether to talk _>?
 b. Which topic did John ask <whether to talk about _>?
 c.* How did John ask <whether to behave _>?
- (8) a.* About which topic did you leave <without talking _>?
 b. Which topic did you leave <without talking about _>?
 c.* How did you leave <without behaving _>?

The difference between the two sets involves Tense in the first place. While tensed clauses are themselves not islands, cf.

- (9) Which topic do you think that I talked about _?

we see that the presence of Tense may considerably strengthen other islands. This fact must be systematically controlled for in the evaluation of various islands. (Chomsky 1986 briefly touches upon the role of Tense. Cinque 1990 uses it as an important test, and Manzini 1992 proposes a theory of its behavior.)

Thus, the specific constructions in (5)-(6) may preclude the presence of any gaps, but this in itself does not mean that wh-complements and adjuncts per se are totally impenetrable.

Shall we conclude now that wh-clauses and adjuncts per se are merely weak islands? Note a fine distinction: while adverb extraction is blocked in both, the infinitival whether-clause allows the extraction of both DPs and PPs (at least for some speakers), while the gerundival adjunct clause allows only DP-extraction.

The interpretation of this situation is a theoretical, rather than descriptive, matter. Making the implicit assumption that the PP versus DP facts are highly consistent in each island type, Cinque 1990 argues that this difference warrants radically different analyses, because a DP gap may be an empty resumptive pronoun but a PP gap cannot be. Wh-islands are indeed weak (selective); adjunct-islands on the other hand are strong (absolute), and the DP-gap that they may contain is an empty resumptive pronoun.

Cinque's distinction will serve as a basic organizing principle in this chapter, because it is typically taken for granted in the weak island literature: theories of weak

islands do not try to account for the islands that Cinque classifies as strong.

- (10) Cinque's diagnostic of strong versus weak islands:
 Among the domains that do not allow all standard extractions, those that allow a PP-gap are weak islands, and those that can at best contain a DP-gap are strong islands (and their DP-gap is an empty pronoun).

This is not the only possible way to cut the cake. For instance, Postal 1997, 1998, who uses the terms locked (absolute) versus unlocked (selective) island, does not attribute this much significance to PP versus DP. Following Ross, he assumes that extraction from unlocked islands always involves empty resumptive pronouns and subsumes Cinque's weak islands under this rubric.

A note on the data. Question formation, relativization, and topicalization are all instances of *wh*-movement. Because the bulk of the literature to be surveyed focuses on examples involving question formation, this chapter follows the same practice, unless the discussion of a particular work requires us to do otherwise.

3. Strong islands

3.1 Classical strong islands

While recent literature has identified a host of weak islands, the classical inventory, with the exception of certain *wh*-islands, consists of strong ones. What follows is a list, sometimes interrupted by comments.

Complex DP (with relative clause):

- (11) * Which kid must you call <the teacher who punished _>?
 (12) * This kid, you must call <the teacher who punished _>.
 (13) * What size shoes did you call <the man who wears _>?
 (14) * How did you call <the man who behaved _>?

Complex DP (with complement clause):

- (15) * Which man did you hear <the rumor that my dog bit _>?
 (16) * This man, I heard <the rumor that my dog bit _>.
 (17) * What size shoes did you hear <the rumor that I wear _>?
 (18) * How did you hear <the rumor that I behaved _>?

But just as it was possible to arrange for a DP-gap in adjuncts by removing Tense, the same is possible with Complex DPs if the definiteness of the "head" is removed. (Pollard--Sag (1994, vol.2) assume, for this reason, that there is no need for a Complex DP constraint.) PP-gaps remain practically excluded.

- (19) a.? What the police arrested <everyone who saw _> was this video. (Postal

- 1998)
- b.* I know in which building the police arrested <everyone who lives _>.
- (20) a.? Which man did they consider <rumors that Bob would betray _>?
(Rothstein 1988)
- b.* About which man did they consider <rumors that Bob would talk _>?

In distinction to Tense, Definiteness may create an island by itself (Fiengo and Higginbotham 1981, Manzini 1992):

Definites:

- (21) * Which man did you discover <Mary's poem about _>?
(22) ?? Which man did you discover <the poem about _>?
(23) Which man did you discover a poem about _?

Subjects:

- (24) * Which man did <his visiting _> shock you?
(25) * Which book do you believe <the first chapter of _> to be full of lies?
(26) * Which man does <everyone who knows _> admire his sincerity?

Adjuncts (see the discussion of Tense above):

- (27) * Which topic did you leave <because Mary talked about _>?
(28) a. Which topic did you leave <without talking about _>?
b.* About which topic did you leave <without talking _>?

Authier 1991 argued that adjuncts can contain CP gaps besides DP gaps; for an alternative analysis, see Postal 1994:

- (29) We suggest _ to our employees, <without actually requiring _ of them>, that they wear a tie.

Coordinate structures, unless extraction is across-the-board:

- (30) * Which man did you invite <Mary and _>?
(31) * Which man did you invite <Mary and a friend of _>?
(32) Which man did you invite <a friend of _ or a brother of _>?

However, DP-gaps are acceptable under various circumstances (see Postal 1998 for detailed discussion):

- (33) This is the beer that I <bought _ , loaded _ into the car, went home, and then fell asleep>. (Jacobson 1996)

The status of tensed constituent *wh*-complements seems to vary cross-linguistically. Some speakers of English reject extractions from them entirely, whereas others find them tolerable. Likewise, they are rejected by speakers of Dutch, whereas they are acceptable in Scandinavian languages (see Engdahl and Ejerhed 1982) and in Hungarian. (Hungarian has object pro-drop, but only for singulars, thus we know that the gap in (c) is not a dropped pronoun.) This variation is not well-understood (though see Bayer 1984, 1996:6.6.4).

Tensed constituent wh-complements (in some languages/dialects)

- (34) a.% Which man did John ask <who invited _>
 b.* Welke man heb jij je afgevraagd <wie _ gezien heeft>? (Dutch)
 which man have you self asked who seen has
 c. Mely fiúkat találgattad, hogy <ki látta _>? (Hungarian)
 which boys were-guessing-you that who saw

Left branches:

- (35) * Which (man's) did you see <_ picture>?

But comparable examples are perfect in other languages:

- (36) Combien as-tu lu <_ de livres>? (French)
 how-many have-you read of books
 (37) Cuius legis <_ librum>? (Latin)
 whose read-you book
 (38) Kinek olvastad <_ a könyvét>? (Hungarian)
 who-dative read-you the book+agr

Whether all the counterexamples are well-understood, the Left Branch condition seems like a less plausible generalization than the others. See Corver (1990).

Finally, some islands constrain left-extraction but not Right Node Raising, indicating that the latter may not be an extraction (Bouma 1987, Postal 1998):

- (39) I patted <the dog that bit _> and hit <the dog that adored _> the man who stole my beloved cat.

3.2 Types of explanation

The standard explanation of the islandhood of Complex DPs, Subjects, Wh-

complements, and Left Branches is in terms of Subjacency: they are out because movement crosses more than one bounding node (barrier).

Adjunct islands are standardly explained by Huang's 1982 Condition on Extraction Domains, hence by the Empty Category Principle: an extraction domain needs to be properly governed. Likewise, Pesetsky 1982 subsumes the Coordinate Structure Constraint under the path containment version of the ECP.

Manzini 1992 is the only attempt to unify the effect of Definiteness and Tense with other locality phenomena: D and T block dependencies based on Case-addresses, which, in her theory, DPs otherwise rely on to escape from islands.

3.3 Ways of salvaging strong island violations

As has been mentioned, a strong island violation can sometimes be salvaged by resorting to the resumptive pronoun strategy (whether the pronoun is overt or empty).

Second, the gap in the island may be parasitic on a well-behaved trace of movement, e.g.:

- (40) Which papers did you file _ <without reading _>?
 (41) He is a man who <everyone who knows _> admires _.

See Engdahl 1983, Kayne 1983, Chomsky 1982, 1986, Cinque 1990, among others.

Third, the island itself may be pied piped, overtly or at Logical Form:

- (42) a.* Whose did you visit <_ brother's sister>?
 b. <Whose brother's sister> did you visit?

Pied piping at LF has been proposed for apparent violations of Complex DPs in wh-in-situ languages in Pesetsky 1987, and for apparent violations of a single adjunct island in Cinque 1990, for instance.

3.4 More on Subjacency

Subjacency is classically understood as a condition on movement and requires that movement not cross more than one bounding node. Bounding nodes were originally defined as a list: NP and S (=DP and IP) in English (Chomsky 1973, 1977), NP and S' (=DP and CP) in Italian (Rizzi 1978).

Chomsky 1986 redefines bounding nodes as barriers. An XP is a blocking category for α iff it is not theta-marked by a sister lexical head and dominates α . β is a barrier for α iff (i) β is a blocking category for α but not IP, or (ii) β is the first XP that dominates a blocking category for α . Chain-formation (which takes the place of movement) requires 1-subjacency: no more than one barrier may be crossed. (For an excellent introduction to "Barriers" as well as other theories of locality, see Roberts 1997.)

Adopting Rizzi's 1990 Relativized Minimality (the pertinent aspects of which will

be discussed when we turn to Weak Islands), Cinque 1990 proposes two important changes in the understanding of strong islands. (i) Strong islands constrain binding chains, not necessarily movement, and (ii) Not only government but also binding chains require 0-subjacency (no barrier may be crossed).

3.4.1 Cinque 1990

We take a closer look at some aspects of Cinque's proposal, because they are fundamental in defining the division of labor between theories of strong and weak islands, as assumed in most of the literature.

Following Obenauer, Cinque observes that both parasitic gaps and gaps inside strong islands are restricted to the category DP and takes this to indicate that these gaps are not variables, but A-bar bound empty pronominals: *pro*. This *pro* is unmoved in syntax but must move at LF (like some kind of abstract *wh*-phrase). But instead of moving on its own, it pied pipes the minimal island it is contained in. This accounts for the fact that, if all goes well, one strong island can be evaded.

Now recall the contrast between tensed and gerundival adjunct islands:

- (43) * Which topic did you leave <because Mary talked about _>?
 (44) Which topic did you leave <without talking about _>?

Bona fide cases of pied piping are also constrained by Tense (Nanni and Stillings 1978):

- (45) * They bought a car that their son might drive which was a surprise to them.
 `they bought a car and the fact that their son might drive it was a surprise to them'
 (46) The elegant parties, to be admitted to one of which was a privilege, had usually been held at Delmonico's.

Cinque stipulates that Tensed Inflection weakly blocks the upward percolation of features. Thus, even one island cannot be pied piped by *pro* if it is tensed.

Cinque predicts that at most one island can be evaded: strong islands cannot be compounded:

- (47)* the book that we left Russia <without being arrested <after distributing _>>

The reason is that the movement of the pied piped domain continues to be sensitive to islands (can at best be parasitic on overt movement). Postal (1998) notes, though, that an adjunct island can be compounded with a complex DP:

- (48) It was Lucille that Mike went home <without criticizing <anyone who defended _>>.
 (49) It was Lucille that Mike criticized <everyone who went home <without

defending _>>.

To summarize, the observation that certain islands only contain DP gaps leads Cinque to a theory according to which these islands are strong, and the DP gap is not a trace of movement but A-bar bound *pro*. This contrasts with Weak Islands, which may also contain PP gaps. The latter must be traces since, according to Cinque, human languages generally lack resumptive pronouns of category PP. The behavior of weak islands is explained by considerations that do not apply to strong islands, in Cinque's work and in the many other proposals this chapter will review. Thus, the DP versus PP distinction carries a great burden in deciding which data are to be accounted for by each theory.

3.4.2 Some intriguing similarities

While the division of labor so determined has proven very useful in the literature, it may be worth pointing out some intriguing similarities between the phrases that may or may not escape from the two types of islands.

First, there is dialectal variation among speakers of English regarding the acceptability of PP-extraction out of *wh*-islands (one of the weak islands). Unfortunately, no systematic empirical study of this variation exists, to my knowledge, wherefore it is difficult to assess its significance for Cinque's theory.

Second, consider the following contrast:

Wh-phrase associated with *pro* in a strong island:

- (50) Which politician did you go to England <after meeting _>?
 (51) * How much water did you make the pasta <after boiling _>?

Wh-phrase associated with a variable in a weak island:

- (52) a. Which politician did John ask <whether to worry _>?
 b. About which politician did John ask <whether to worry _>?
 (53) a.* How much gravy did John ask <whether to cook _>?
 b.* With how much gravy did John ask <whether to cook _>?

In both cases, the *wh*-phrase needs to be referential in some sense. But the reasons are different. In the case of strong islands, the *wh*-phrase needs to be of the kind that resumptive pronouns tend to associate with; in the case of weak islands, it has to carry a referential index and corefer with its variable (see in 7.2). Cinque 1990 characterizes the suitable *wh*-phrases rather similarly in the two cases, but neither this book nor any subsequent work known to me addresses the question whether the two requirements are exactly identical. Cinque (pers. comm., 1998) has kindly suggested the following contrast

between a weak and a strong island,

- (54) ? Quanti pazienti volevi sapere <se ogni dottore avesse potuto visitare _>?
 `How many patients did you want to know whether each doctor had been able to examine?'
- (55) * Quanti pazienti vuoi incontrare <ogni dottore che abbia visitato _>?
 `How many patients do you want to meet each doctor who examined?'

but the exact generalization remains an open question.

4. Weak Islands: A preview

The historical starting point is the assumption, made in Huang 1982, Lasnik and Saito 1984, 1992, Chomsky 1986, that the paradigmatic (if not the only) case of weak (selective) islands is wh-islands, and the expressions whose extraction is sensitive to WIs are adjuncts, as opposed to arguments.

Since "Barriers," a number of new theories of weak islands have appeared in quick succession. What makes this process especially interesting is the fact that practically each theory comes with a significant new set of data. The survey below will reflect this spirit. The critical data come in two dimensions:

- (A) What extractions are sensitive to WIs?
 (B) What induces a WI?

Much of the literature can be conveniently surveyed along the dimensions in (A) and (B). I will thus begin by drawing ever-widening circles of data, first for (A), then for (B). The third part of the discussion will review what theories account for what data sets.

In both these sections and in section 7, where we turn to theories, we will see that the proposals are less and less syntactic. Some of them include pragmatic or semantic factors in the description of what expressions are WI-sensitive or WI-inducers, although they formulate the explanation in syntactic terms. Others even derive the explanation from semantics.

The data to be discussed are as follows:

Ad (A) What extractions are sensitive to WIs:

- (A1) Arguments versus adjuncts
 (A2) Referential versus non-referential
 (A3) Re-evaluating the role of D-linking
 (A4) Individual versus non-individual, and how many-phrases
 (A5) Functional readings and event-related readings
 (A6) Individuals: is it being type e that matters?

- (A7) Split constructions
- (A8) Negative polarity item (NPI) licensing
- (A9) Cross-sentential anaphora
- (A10) Elements of anti-pronominal contexts

Ad (B) What induces a WI:

- (B1) Wh-islands
- (B2) Negatives and other affective operators
- (B3) Response stance and non-stance, versus volunteered stance predicates
- (B4) Extraposition islands
- (B5) VP-adverbs
- (B6) Scope islands

Section 7 presents **the theories** that account for these data and generalizations. To anticipate, the division of labor is as follows. When a theory accounts for only some of the data falling under some generalization, the generalization is marked with a percentage sign (%):

- ECP and Subjacency for (A1)--(B1)
- Relativized Minimality for (A2, A7)--(B1, B2, B3%, B4, B5%)
- Monotonicity for (A3, A4, A7%)--(B1, B2, B3%, B5%)
- Scope Theory, algebraic version for (A3, A4, A5, A6, A7%)--(B1, B2, B3, B5, B6)
- Scope Theory, dynamic semantic version (A4, A5, A7, A8, A9)--(B1, B2, B3, B5, B6)

6.1 (B1) Wh-islands

As was mentioned in 3.1, infinitival, modal, or subjunctive whether-complements are the weakest islands, as illustrated by English and Dutch below. Tensed constituent-wh questions are often strong islands. The cross-linguistic and cross-dialectal variation is not well-understood.

- (141) a. Which man are you wondering <whether to invite _>?
 b.* How are you wondering <whether to behave _>?

6.2 (B2) Negatives and other affective operators

That negatives block extraction was observed as early as in Williams 1974, together with the fact that only unstressed negatives do (the latter remains unaccounted for in the literature):

- (142) * I asked how John <didn't behave _>.
 (143) I asked how John did NOT behave _.

The island-inducing effect of negatives was analyzed systematically in Ross 1984, who called them Inner Islands:

- (144) What did <no imitation pearls touch _>?
 (145) * What did <no imitation pearls cost _>?

Rizzi 1990 observes that the generalization extends to all items that are affective operators (i.e. NPI triggers) in the sense of Klima 1964.

- (146) * How did <few men think that you behaved _>?
 (147) * How did <only John think that you behaved _>?
 (148) * How did <you deny that you behaved _>?

In terms of Ladusaw 1980, affective operators are downward entailing (on the whole, or have such a component, as is the case with only). Compare:

- (149) No one / Few men / Only John had ever been here.
 (150) I deny that you have ever been here.

6.3 (B3) Response stance (e.g., negative) and non-stance (e.g., factive), in contrast to volunteered stance predicates

That factive contexts are WIs is well-known from Cinque 1990. This is just a special case, though, as Hegarty 1992 points out. The full picture is presented in Cattell 1978. Cattell distinguishes response stance, non-stance, and volunteered stance predicates. Response stance includes, alongside with deny, items like accept, confirm, verify, admit, and others. Non-stance includes a big class of factives: realize, know, regret, remember, surprise, notice, etc. Finally, volunteered stance includes think, believe, suspect, allege, assume, claim, etc. Cattell's generalization is that the first two classes generally block adjunct extraction:

- (161) * How did you accept that he behaved _?
 also: deny, confirm, verify, admit, ...
 (162) * How did you realize that he behaved _?
 also: know, regret, remember, surprise, notice, ...

- (163) How did you think that he behaved _?
 also: believe, suspect, allege, assume, claim, ...

6.4 (B4) Extraposition islands

Again, this is one of the cases discussed in Cinque 1990. Whether extraposition islands constitute a separate case is not immediately evident, since they are overwhelmingly also factive islands, e.g:

- (164)* How <is it a scandal that he behaved _>?

But here is at least one extraposition island quoted by Cinque that is not factive:

- (165)* How <is it time to behave _>?

On the other hand, the following extrapositions involve non-factive, non-negative predicates and do not appear to create islands (P. Postal (p.c.)).

- (166) How much wine is it possible/legal/fun/dangerous to drink _ at a party?

This seems to indicate that extraposition per se is not the culprit. It is unclear to me exactly what makes it is time to... an island. Further relevant examples can be drawn from Heavy NP Shift.

6.5 (B5) VP-adverbs

Obenauer 1984/85 examines the following pattern:

- (167) J'ai beaucoup conduit [_ de camions].
 I have a lot driven of trucks
 'I drove many trucks'

Although the phrase beaucoup de camions would be grammatical, Obenauer argues in careful detail that VP-initial beaucoup is not an extracted determiner but a VP-adverb. The quantified interpretation 'many trucks' is obtained through quantification of V in terms of 'many times'. He calls this phenomenon QAD (quantification at a distance). The presence of the adverb nevertheless blocks combien-extraction:

- (168) * Combien as-tu <beaucoup consulté [_ de livres]>?

The analysis according to which a VP-initial quantifier renders split constructions ungrammatical without being derivationally related to the empty determiner slot of the DP is corroborated by the fact that iterative adverbs also block wat voor split (de Swart 1992):

- (169) * Wat heb je < twee keer [_ voor boeken] gelezen >?
 what have you two times for books read
 `What kind of books did you read twice?'

It may be interesting to point out that these adverbs block how-extraction as well:

- (170) *How did you < behave a lot _ >?
 *How did you < behave twice _ >?

6.6 (B6) Scope islands

At least one important subset of the islands above, namely, *wh*-islands, negative (affective) islands, and adverbial islands, can easily be described as intervention islands: the WI is induced by the intervention of a certain element between the final locus of the extracted phrase and its extraction site.

The intervention of other operators seems harmless, e.g., every boy may intervene between how and its extraction site, without making the sentence ungrammatical.

- (171) How did every boy behave _?

As we shall see, Rizzi's 1990 generalization is that bad interveners are A-bar specifiers. Expressions in A-specifier or adjoined position do not matter. Universals are adjoined to some projection by QR, hence they cannot matter.

It turns out that the grammaticality of (171) type extractions is misleading. Every boy is an operator that may scope either below or above a *wh*-phrase; it may even be scopally independent from it. Let us illustrate the three possibilities with an example where the *wh*-phrase is immune to WIs:

- (172) Which book did every boy read _?
 (i) every > wh
 `For every boy, which book did he read?'
 (ii) wh > every
 `Which book is such that every boy read it?'
 (iii) independent scope (uniformity presupposition)

'Taking for granted that every boy read the same book, what was this book?'

Reading (i) is often called a pair-list reading: it is answered by a list of pairs: "Bill read Magic Mountain, Jim read The Russia House, ...". Readings (ii) and (iii) both ask for a single book that was read by every boy, but differ as to the possibility of what else each boy may have read. For instance, if Bill read Jurassic Park and Tom Jones, Jim read Jurassic Park and Airframe, and so on, reading (ii) is felicitous and the answer is "Jurassic Park." Reading (iii) is not felicitous in the same situation: it presupposes that each boy read just one book, moreover, the same one, and merely asks to identify the book. The question is, are these three readings equally possible when every boy interacts with a WI-sensitive expression, rather than a which-phrase.

Kiss 1993 and de Swart 1992 make the fundamental observation that universals are harmless only when they (i) scope above or (iii) independently of the sensitive wh-phrase. When they scope below it (ii), they induce a WI.

Kiss states the generalization for non-specific extractees, de Swarts for split constructions:

(173) How did every boy behave _?

(i) every > wh

'For every boy, how did he behave?'

(ii) wh > every

* 'What was the common element in the boys' non-uniform behavior?'

(iii) independent

'Taking for granted that every boy behaved the same way, what was it like?'

(174) Combien ont-ils tous lu _ de livres?

(i) 'For each of them, how many books did he read'

(ii)* 'For what n, they all read n books'

In sum, the fundamental observation is that it is not so much the nature of the syntactic position (A-bar specifier versus other) but the scope of the intervener that matters: an operator that scopes between the WI-sensitive extractee and its extraction site blocks the extraction.

Why does the mere presence of the previously reviewed operators induce a WI? Because, for reasons of their own, they are unable to scope above (or independently of) the extracted phrase. For instance, no one or twice never induce pair-list questions at all:

- (175) a. Which book did no one read?
 * `for no one, tell me which book he/she read'
- b. Which book did you read twice?
 * `for two occasions, tell me which book you read then'

Following Kiss's and de Swart's insight, this point is demonstrated systematically in Szabolcsi and Zwarts 1993.

7. Theories

7.1 ECP and Subjacency for (A1)--(B1)

The historical starting point is the assumption, made in Huang 1982, Lasnik and Saito 1984, 1992, Chomsky 1986, that the paradigmatic (if not the only) case of weak (selective) islands is wh-islands (A1), and the expressions whose extraction is sensitive to WIs are adjuncts, as opposed to arguments (B1).

- (185) Which problem are you wondering how to [t' [phrase t]]?
 (186) ? Which student did he wonder whether to [t' [consider [t intelligent]]]?
- (187) How have you [t'''[decided [t'' to [t' [phrase the problem t]]]]]?
- (188) * How are you [t'' wondering [which problem to t' [phrase t]]]?
 `what is the manner such that you are wondering which problem to phrase in that manner'

In all four examples, the lowest trace t is properly governed by t', an intermediate trace adjoined to the closest VP. The distinction between arguments and adjuncts is drawn in the following way.

An empty category that is in argument position is indelibly marked as to whether it is properly governed at S-structure. Thus in (185)-(186), t is marked as properly governed. t' itself is not properly governed when it is separated from its antecedent by a wh-island (since government requires 0-subjacency). However, intermediate traces can delete before Logical Form, because they are not needed for interpretation. Thus in (185)-(186), the only trace that must appear at LF is t, which is legitimate.

On the other hand, non-argumental (adjunct) empty categories are marked for proper government as late as at LF, whence intermediate traces need to be retained. In (187), the antecedent government chain <how, t''', t'', t', t> is unbroken by any subjacency violation. But in (188), t'' does not govern t', since they are separated by a wh-island. t' cannot be deleted, because then t would not be governed. Thus, on the whole, the extraction of an adjunct out of a wh-island inescapably involves an empty category that is not properly governed.

As Cinque (1990) puts it, this theory forces adjuncts to move strictly successive cyclically, whereas arguments may, in effect, undergo long movement.

We see that this theory of weak islands is purely syntactic. The extraction of a phrase is sensitive to *wh*-islands simply because it originates in a non-argument position. Likewise, a moved *wh*-phrase creates an island purely with its bulk -- by occupying the specifier position of CP and thus preventing the non-argument extraction from proceeding through that position.

7.2 Relativized Minimality for (A2, A3)--(B1, B2, B3%, B4, B5%)

The fact that a fronted *wh*-phrase blocks certain extractions can be attributed to the fact that, by occupying an escape hatch, it gives rise to a subjacency violation. But this explanation does not extend to negation or VP-adverbs, for instance: they do not occupy any escape hatch whatsoever. Relativized Minimality (Rizzi 1990, Cinque 1990) is a particularly exciting theory because it pulls together a host of new islands and seeks to provide a unified account. Specifically, it seeks to account, besides *wh*-islands (B1), also for negative and affective islands (B2), factive islands, a subset of those discussed in (B3), extraposition islands (B4), and the islands created by VP-adverbs (B5). The expressions that are sensitive to WIs are defined as non-referential, as in (A2), which includes split constructions, as in (A3).

Rizzi 1990 builds primarily on the theoretical analysis of Quantification at a Distance (QAD) in Obenauer 1984/85. Obenauer's insight is that a local relation between an operator and its variable is blocked by the intervention of any third party that may be derivationally totally unrelated to them but is sufficiently similar to the operator. In the same spirit, Relativized Minimality is a representational theory of "like" intervention. This is the basic intuition as to what expressions create WIs. (For recent versions of Relativized Minimality, see the Minimal Link Condition in Chomsky 1995 and its revision in Manzini 1998.)

The assumption that WI-sensitive operators need to be connected to their variable by a series of local relations, while WI-immune operators may be linked to their variable long-distance is retained and is recast in the following way. Rizzi assumes that there are two (non-exclusive) ways to connect an operator and its variable:

- (i) binding and
- (ii) a chain of government relations.

Binding is an arbitrarily long-distance relation that requires the identity of referential indices; consequently, only operators that are referential can use it. Non-referential operators need to rely on a government chain. Rizzi makes intervention relevant specifically for government. Consequently, "like" interveners block the government relation between a non-referential operator and its variable. This is the basic intuition as to what expressions are sensitive to WIs. Rizzi 1990 defines referential operators as those that bear an event-participant thematic role like Agent, Patient, etc. Cinque 1990 adds the requirement that the operator must also be Discourse-linked. Non-referential *wh*-phrases are those that bear a role like Reason, Manner, Measure, etc. or are not D-linked.

The role of intervention is technically explicated as follows. "Barriers" incorporates a notion of minimality: a governor cannot govern into the domain of another, closer

governor. Minimality is defined rigidly: only an intervening head governor counts, and it blocks both head and antecedent government. Rizzi proposes to relativize minimality to the kind of relation that obtains between governor and governee:

(189) Relativized Minimality:

$X \alpha$ -governs Y only if there is no Z such that

(i) Z is a typical potential α -governor for Y , and

(ii) Z c-commands Y and does not c-command X .

Rizzi finds it useful to distinguish four kinds of values for α : head government, antecedent government in an A-chain, antecedent-government in an A-bar chain, and antecedent government in head-chain.

Since we are concerned with chains headed by a wh-phrase in A-bar specifier position, all and only A-bar specifiers are relevant interveners.

Rizzi analyzes whether, who, beaucoup, not, no one, only John, and deny as A-bar specifiers, at S-structure or at LF. In contrast, he points out that everyone and two men acquire their scope by adjunction, according to May 1985, so they are predicted not to block non-referential extraction.

Extraposition islands and factive islands (and the other islands noted by Cattell, not mentioned by Rizzi or Cinque) are not covered by these considerations. To account for these, Cinque adds that the clause from which extraction occurs must be a sister of a theta-marking [+V] head.

While the general idea is very insightful, there are some analytical problems with this execution. One is the movement of deny, a head, to an A-bar specifier position at LF, and the assumption that the complement of regret is not a sister to the verb. These have an alternative solution within Relativized Minimality, however: the adoption of Progovac's 1988 and Melvold's 1991 proposals to place empty operators in the [SPEC,CP] of the complements of deny and regret, which then serve as standard interveners.

More important perhaps is the problem posed by the cross-linguistic variation in the syntax of negation, attributed to the fact that the negative particle may be a head, a specifier, or an adjunct. This would suggest that the island-creating effect of negation varies accordingly, but it does not: we are not aware of any language in which negation does not create a weak island. Rizzi 1992 proposes to solve this problem by assuming an empty A-bar specifier when NEG is a head, and an empty head when NEG is a specifier. This proposal may well be correct, but it makes the original distinction of Relativized Minimality vacuous: how do we tell apart the effects of a head versus a specifier?

But there is a more important problem: the class of WI-inducers cannot be identified as intervening A-bar specifiers. This is highlighted by briefly considering another proposal that is closely related to Relativized Minimality.

7.4 From Relativized Minimality to a Scope Theory

7.4.1 A major problem with Relativized Minimality

Recall a crucial background assumption of Relativized Minimality. The theory of LF that it relies on is that of May 1985. According to this theory, structure (usually) does not disambiguate scope. (191), for instance, is assigned a single structure in which how is higher than everyone, but they govern each other, whence they can be interpreted in either scope order or even independently. The adoption of this theory for the purposes of Relativized Minimality results in the assumption that it does not matter which reading of the sentence we are considering; all we have to know is that everyone is in an adjoined position, whence its intervention between how and its trace must be harmless. (192) is also assigned a single structure, but no one occupies an A-bar specifier position in it, whence it must block how-extraction.

- (191) How did everyone behave?
 (192) * How did no one behave?

As seen in 6.6, results by É.Kiss 1993 and de Swart 1992 indicate that this assumption is wrong. For instance, sentences like (193) are potentially ambiguous, but they are ungrammatical on the reading where the universal takes narrower scope than the wh-phrase.

- (193) Combien ont-ils tous lu de livres?
 how many have-they all read of books
 (i) 'For each of them, tell me what number of books he read'
 (ii) * 'For what number, they all read that number of books'

7.4.2 Scopal intervention

É.Kiss 1993 and de Swart 1992 propose that any operator that scopes between the sensitive expression and its trace blocks the relation. É.Kiss identifies the class of sensitive expressions as non-specifics in the sense of Enc 1991, including nominals and adverbials. (See also Frampton 1991.) De Swart identifies sensitive expressions as those whose operator part is split away from the restrictive clause:

The Scope Generalization:

- (194) Specificity Filter: If Op_i is an operator which has scope over Op_j and binds a variable in the scope of Op_j , then Op_i must be specific. (É.Kiss)
 (195) A quantifier Q_1 can only separate a quantifier Q_2 from its restrictive clause if Q_1 has wide scope over Q_2 (or is scopally independent from Q_2). (de Swart)

These generalizations put the WI phenomenon in an entirely new light. Just as Relativized Minimality was based on the observation that the range of WI-inducers is much wider than Subjacency can account for, the Scope Generalization expresses the

observation that both the range and the nature of WI-inducers is different from what Relativized Minimality (in its original form or in its monotonicity reincarnation) can take care of. We come back to the possible theoretical explanation of the generalization in 7.5.

What is the coverage of these proposals? As regards island sensitive expressions, Kiss's proposal that non-specific operators are island-sensitive essentially covers the classical data (A3, A4, A7) and event-related readings (A5), but it is not easy to see how functional wh-phrases (A5) and collectives (A6) can be said to be generally non-specific, since their determiner may be which; also, why the trigger--NPI relation (A8) might fall under the generalization is not obvious. De Swart's generalization is stated with reference to splitting: here the question is what data, beyond the classical ones (A7) might be analyzed as involving splitting. Many pieces of data involving anti-pronominal contexts (A10) fall outside both proposals.

As regards WI-inducers, it is easy to see that, with the obvious exception of purely syntactic WI-inducers like it is time to... extraposition and some of the constructions discussed by Postal, the known WI-inducers are scope-bearing operators and thus fall under the Kiss/de Swart proposal. The problem may be, rather, that there are expressions that some well-established theory or other classifies as scope-bearing operators and they nevertheless do not induce WIs. Such are indefinite DPs and intensional verbs like want:

(196) How did a boy behave _?

(197) How do you want me to behave _?

Clearly, there are two ways out. One is to adopt analyses according to which indefinites and intensional operators are not scopal. Another is to draw some principled demarcation line between scopal expressions, predicting some of them to be innocuous.

7.4.3 Prospects for a revision of Relativized Minimality

We have seen that in order to capture the Scope Generalization, LF ought to disambiguate scope and, according to May 1985, it does not. Although this feature played a major role in diverting attention from some critical data, it is not very difficult to fix. It is not too difficult to imagine a scope theory which produces disambiguated LFs.

A second, more serious aspect of the problem is that, if the spirit of Relativized Minimality is to be maintained, QR cannot be an adjunction rule. Instead, it should move the quantifier into an A-bar specifier position. This in turn requires that quantifiers all have designated landing sites. This requires much more radical departure from May 1985. A theory that at least partially fulfils this expectation has been proposed recently, namely, by Beghelli and Stowell 1997.

According to this theory, quantifier scope is a by-product of feature checking. Definites and specific indefinites check features in positions called RefP and ShareP; distributive universals check features in DistP. But Beghelli and Stowell argue, with equal force, that modified numeral (counting) QPs do not have comparable features and never move beyond their Case position. This accounts for the fact that they only marginally take

inverse scope. This is a major stumbling block in the way of a revision of Relativized Minimality, because as we saw above, more/fewer than five boys is as much of a WI-inducer as narrow scope every boy. But then it is difficult to see how the positions of, say, universals and modified numerals can be brought under a single heading.

Likewise, if one goes beyond quantifiers, it is not obvious how all of Cattell's WI-creating predicates can be placed in a position (or can be associated with an operator in a position) that falls under the same generalization.

In other words, the Scope Generalization implies a radical widening of the set of WI-inducers, and it is not at all clear how all these might fit a single syntactic recipe.

The syntactic recipe does not necessarily have to make reference to "A-bar specifier". But the beauty of Relativized Minimality was that it established a very tight connection between the properties that make extractees WI-sensitive and the properties that make interveners WI-inducers. The revision should recapture this explanatory connection.

In fact, it is also not obvious how all WI-sensitive expressions will fit a single syntactic recipe. Note that, alongside non-individual *wh*-phrases, WI-sensitive expressions include the amount and the event-related readings of numerical QPs, functionally interpreted which-phrases, definite dependents of "one time only" predicates, and negative polarity items.

A summary of the scope theories can be found in the SynCom case and in den Dikken-Szabolcsi, Islands, GLOT International 1999.

References

Syntax I, 12/6
Szabolcsi

11 Theories with or without movement, theories with or without empty categories
(a comparison of GB, HPSG, and categorial grammar)

Reading: Szabolcsi, Combinatory grammar and projection from the lexicon, Part I. In Sag – Szabolcsi, eds., *Lexical Matters*, CSLI, 1992.