Derivations/Representations

1 Long-Distance Dependencies: HPSG

(1) Argument Realization Principle (revised) [p. 432]



[The subtracted list C may be null, in which case the value for GAP is null as well.]

(2) **The GAP Principle**

A local subtree Φ satisfies the GAP Principle with respect to a headed rule ρ iff Φ satisfiesL



(3) Head-Filler Rule

$$[phrase] \rightarrow [1][GAP \langle \rangle] H HEAD \begin{bmatrix} verb \\ FORM fin \end{bmatrix}$$
$$VAL \begin{bmatrix} SPR & \langle \rangle \\ COMPS \langle \rangle \end{bmatrix}$$
$$STOP - GAP \langle [1] \rangle$$
$$GAP \langle [1] \rangle$$

Properties of interest:

- non-derivational/non-directional
- crash-proof: follow the rules and you win

2 Long-Distance Dependencies: Radical Minimalism (Brody)

- (4) Existence of copies
 [cf. the percolation of the GAP list up the tree]
 Structure may contain identical elements (copies).
- (5) **Command condition** [cf. the Gap Principle + the Head filler Rule] $C = (\alpha_1, \dots, \alpha_n)$ is a *chain* if and only if a. all members of C are identical b. α_i c-commands α_{i+1}

(6) Main Thematic Condition (MTC)

All non-root positions of chains are non-thematic. [cf. the fact that the only gap-producing device in HPSG is the Argument Realizational Principle, which removes elements from the COMPS list]

Deriving the MTC:

Call the set of chain-root positions in a structure its *D*-set. (p. 13). Then: "Projectional requirements can only involve positions that belong to the D-set."

(7) **Full Interpretation (FI)**

LF can contain only elements that receive an interpretation.

(8) **Pronunciation principle [cf. the right side of the Head Filler Rule]** In an "overt movement chain" $C = (\alpha_1, \dots, \alpha_n)$, pronounce only α_1 .

μμμ

David Pesetsky, course materials for 24.960 Syntactic Models, Spring 2006. MIT OpenCourseWare (http://ocw.mit.edu/). Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

3 Some History

"Trace theory" and c-command

Step 1: Mid-1970's excitement:

- Movement is free -- up, down, sideways -- fine!
- But: movement leaves a trace.
- The trace is a kind of "anaphor", and must be c-commanded by an "antecedent".
- The c-command condition is a *surface filter* on derivations.

Consequence: Downward movement is ok, if some later process obliterates the trace.

- **Examples:** Expletive *there* and *it* obliterate traces of downward movement of an indefinite and a CP.
- **Logic:** Crucial that c-command is *not* a condition on movement, but a condition on representations to which movement (and other operations) have applied.

Koster (1978) Locality Principles in Syntax

Generalized a number of locality conditions on anaphora with conditions on movement, concluding that "cyclic movements, construal rules, and bounded deletions can all be considered instances of the coindexing rule."

Step II: 1980s retrenchment

Trace is not *literally* an anaphor:

Rizzi: *wh*-trace does not observe the locality conditions of reflexives Freidin & Lasnik: *wh*-trace obeys BT(C), not BT(A)

So if the c-command condition is independent of the definition of movement, it is a condition of some other sort -- e.g. a condition on an object called *chain*.

Chomsky (1981) Lectures on Government and Binding

(9) Movement (LGB era, reinterpreted and embellished)

- In the base component, each constituent (each NP) is entered in the Registry of Chains[™] as a member of a unique (therefore singleton) chain.]
- ii. Copy the constituent α_i as α_i and place α_i in a c-commanding position.
- iii. Coindex α_i and α_j
- iv. Delete all material dominated by the mother node of α_i (a construction called a "trace", an instance of an "empty category").
- v. [In the Registry of ChainsTM add α_j to the chain that contains α_i .]

(10) Representational definition of chain from *LGB* [p.333]

- $C = (\alpha_1, \dots, \alpha_n)$ is an [A-]*chain* if and only if
- (i) α_1 is an NP.
- (ii) α_i , <u>locally</u> A-BINDS α_{i+1} .
- (iii) for i > 1,
 - (a) α_i is a non-pronominal empty category, or

(b) α_i is A-free.¹

(iv) C is maximal, i.e. is not a proper subsequence of a chain meeting (i-iii).

Do chains exist?

Answer:

After movement takes place, do we need to know that a particular instance of an XP α is the "trace" of another occurrence of XP β ?

Subquestion 1A:	Can something happen to α that affects β ?
Answer:	Feature-checking (feature-deletion) on α seems to affect β as well, in that the uninterpretable feature on β that allowed the movement to take place/motivated it, is no longer a problem.
Subquestion 1B:	Can something happen to β that affects α ?

(11) **No bleeding of BT(C) without ACD, because trace material is not deleted:** *I [sent him; [every letter that John; expected I would write t]

Deletion patterns in ACD and reconstruction.

(12) But ACD requires deletion of the trace-internal material, which in turn bleeds BT(C): I [sent him; [every letter that John; expected I would [$_{VP} \Delta$]] —>

[every letter that John_i expected I would [VP send him_i t]] I [sent him_i t].

• The negotiation of deletion/non-deletion in the trace requires knowledge of the interpretation of the trace's antecedent.

¹ Case b was a hack for *there* constructions, where *there* was taken to "BIND", but not "bind", its associate. Note that there are no A-bar chains in this formulation.

Subquestion 1C: Does later syntax care about the space between α and β ?

- Data from Rizzi (1986) "On Chain Formation"
- (13) a. Paolo **gli** affiderà Gianni. Paolo to-him will-entrust Gianni
 - b. Paolo **si** affiderà Gianni. Paolo to-himself will-entrust Gianni
- (14) Gianni_i si_i impone [di PRO_i fare il suo dovere] Gianni himself compells to do his duty
- (15) a. Gianni **gli** è stato affidato. Gianni to-himself was entrusted
 - b. *Gianni **si** è-stato affidato. Gianni to-himself was entrusted
- (16) a. Il ladro e il poliziotto gli sono caduti addosso
 - b. ?*Il ladro e il poliziotto si sono caduti addosso.
- (17) a. Gianni non **gli** sembra fare il suo dovere.
 - b. *Gianni non **si** sembra fare il suo dovere. Gianni neg to-himself seems to-do the his duty
- **Rizzi:** Since chains have a property independent of Move, and are otherwise redundant with Move, we have an argument for a representational theory of chain formation and against Move.
- **Brody:** A theory that has both chain formation and Move is a theory with a redundancy.
- Of course, you could investigate other treatments of Rizzi's condition.

4 Chomsky (1995, chapter 4, 223-224) clarifies the issues

"A related question is whether. CHL is derivational or representational: does it involve successive operations leading to (π, λ) (if it converges), or does it operate in one of any number of other ways — say, selecting two such representations and then computing to determine whether they are properly paired, selecting one and deriving the other, and so on? These questions are not only imprecise but also rather subtle; typically, it is possible to. recode one approach in terms of others. But these questions too are ultimately empirical, turning basically on explanatory adequacy. Thus, filters were motivated by the fact that simple output conditions made it possible to limit considerably the variety and

complexity of transformational rules, advancing the effort to reduce these to just Move α (or Affect α , in the sense of Lasnik and Saito 1984) and thus to move toward explanatory adequacy. Vergnaud's theory of abstract Case, which placed a central part of the theory of filters on more solid and plausible grounds, was a substantial further contribution. Similarly, Rizzi's proposals about chain formation were justified in terms of explaining facts about Romance reflexives and other matters.

"My own judgment is that a derivational approach is nonetheless correct, and the particular version of a minimalist program I am considering assigns it even greater prominence, though a residue of filters persists in the concept of morphologically driven Last Resort movement, which has its roots in Vergnaud's Case theory. There are certain properties of language, which appear to be fundamental, that suggest this conclusion. Viewed derivationally, computation typically involves simple steps expressible in terms of natural relations and properties, with the context that makes them natural "wiped out" by later operations, hence not visible in the representations to which the derivation converges. Thus, in syntax, crucial relations are typically local, but a sequence of operations may yield a representation in which the locality is obscured. Head movement, for example, is narrowly "local," but several such operations may leave a head separated from its trace by an intervening head. This happens, for example, when N incorporates to V, leaving the trace t_N and the [$_V V - N$] complex then raises to I, leaving the trace t_V : the chain (N, t_N) at the output level violates the locality property, and further operations (say, XP-fronting) may obscure it even more radically, but locality is observed by each individual step.

"In segmental phonology, such phenomena are pervasive. Thus, the rules deriving the alternants *decide-decisive-decision* from an invariant lexical entry are straightforward and natural at each step, but the relevant contexts do not appear at all in the output; given only output conditions, it is hard to see why *decision* should not rhyme with *Poseidon* on the simplest assumptions about lexical representations, output conditions, and matching of input-output pairings. Similarly, intervocalic spirantization and vowel reduction are natural and simple processes that derive, say, Hebrew *ganvu* 'they stole' from underlying *g-n-B*, but the context

for spirantization is gone after reduction applies; the underlying form might even all but disappear in the output, as in *hitu* 'they extended', in which only the It I remains from the underlying root /ntC/ (C a "weak" consonant).

"It is generally possible to formulate the desired result in terms of outputs. In the head movement case, for example, one can appeal to the (plausible) assumption that the trace is a copy, so the intermediate V-trace includes within it a record of the local N –>V raising. But surely this is the wrong move. The relevant chains at LF are (N, t_N) and (V, t_V), and in these the locality relation satisfied by successive raising has been lost. Similar artifice could be used in the phonological examples, again improperly, it appears. These seem to be fundamental properties of language, which should be captured, not obscured by coding tricks, which are always available. A fully derivational approach both captures them straightforwardly and suggests that they should be pervasive, as seems to be the case.

"I will continue to assume that the computational system CHL is strictly derivational and that the only output conditions are the bare output conditions determined "from the outside," at the interface." (Excerpt courtesy of MIT Press. Used with permission.)

David Pesetsky, course materials for 24.960 Syntactic Models, Spring 2006. MIT OpenCourseWare (http://ocw.mit.edu/). Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

5 Is strict derivationalism motivated? Brody's (1995) reworking of phenomena otherwise explained in terms of derivation

Chapter 2

Claim in the literature:

Subjacency-type conditions are conditions on *movement* -- not on LF chains. <u>Argument 1</u>: Intermediate traces get erased. (For example: intermediate traces of nominative subject movement do not show a *that*-trace/ECP effect because they are erased in the course of the derivation (Lasnik & Saito 1984.)

<u>Argument 2:</u> It matters where in the derivation the movement takes place (overt vs. covert movement; Huang 1983) and whether the chain was formed by movement or not (parasitic gaps).

But we can equally well stipulate that **primary chains** obey subjacency, while **parasitic chains** are immune. That is:

- a. one chain terminating at a given head is always "primary";
- b. designate a covert chain as "parasitic" in preference to designating a overt chain as parasitic. (This is not in Brody, but he needs something ike this.)

Chapter 4

Positional principles stated in terms of movement have natural chain restatements:

(18) **The** *wh***-criterion**

a. A +WH C must have the head of a chain that contains a wh-phrase in its spec position.

b. A wh-phrase must be in a chain whose head is in the spec of a +WH C.

- Earliness/procrastinate principles can be stated as a a question of which position in the chain is occupied by the lexical category -- a "matter of morphology" (p. 104).
- (19) Transparency (cf. Earliness) The contentive category in the chain must be in the highest position licensed by morphology.
- Procrastinate is an unlikely part of grammar because it makes LF and PF maximally dissimilar.
- Suggestion: there is no QR (i.e. the only chains that violate transparency are secondary chains). ACD arises from vehicle change of a full DP to a variable.

Chapter 5

(20) <u>A reconstruction paradox</u>

Mary wondered [which claim that pictures of herself disturbed me/*Bill_{*i*}] he_{*i*} made.

Examples like (20) teach us that there's no reconstruction operation.

- BT(A) is satisfied via any one of the chain members (in (20), the upper one).
- BT(C) must be satisfied in all the chain members (so the lower one triggers the BT(C) effect). But...

"Late merge" representationalized:

Adjunct does not have to be projected as part of each member of the chain, though (by the Projection Principle/Full Interpretation) the complement does.

[Derivational alternative:

Mark the reflexive +BT before reconstruction, and mark *Bill* -BT after reconstruction ... or before *wh*-movement? Binding Theory requires that all DPs be +BT.]

6 What *might* be an argument for "derivations": gravity arguments

Are there genuine bottom->top (or top->bottom) asymmetries in syntax?

• Chomsky, Chapter 4: *there* story

Suppose we know that Spec,TP (finite) of a finite-clause or finite-clause-plusdefective-TP will contain *there*, and we know that a low *theta*-position will contain a DP associate. What should the intermediate Specs contain? Answer: trace of *there*.

- (21) a. There seems [______ to be an argument for derivations in the garden].b. *There seems [an argument for derivations to be _____ in the garden].
- (22) **Representational counterpart of Chomsky's story with similar intuition:** A structure containing chains $C_{i}=(\alpha_{1}...\alpha_{n})$ and $C_{j}=(\beta_{1}...\beta_{n})$, where α_{n} ccommands β_{1} , is well-formed only if an otherwise identical structure formed with the same lexical items in which $C_{i}=(\alpha_{1}...\alpha_{n}, \beta_{1}...\beta_{j})$ and $C_{j}=(\beta_{j+1}...\beta_{j})$ is ill-formed. [That is, maximize chain lengths top-to-bottom.]

(23) **Descriptive alternative:**

An expletive-argument chain is well-formed iff the argument occupies its theta-position

• An imaginary *there* story.

Imagine a language called *Moveglish* (spoken in *Movegl*, naturally) in which all instances of *there* pooled at the bottom of the tree. In Moveglish, (24a) is bad because the instance of *there* found in the matrix clause could have been found in the lower clause:

(24) Moveglish data

- a. [*]There was [a claim that a man was arrested] reported in the newspaper.
- b. [A claim that there was a man arrested] was reported in the newspaper.
- c. A claim that a man was arrested was reported in the newspaper.

Moveglish is unexpected under a representational theory, and is no more expected that an alternative language in which instances of *there* pool at the *top* of the tree, such that (24a) is good and (24b) is bad.

Why don't we speak Moveglish...

7 What *might* be an argument for ''derivations'': opacity arguments from Remnant Movement

(25) a. **DP scrambling, then infinitival clause topicalization**

 $[t_j \text{Zu lesen}]_i$ hat keiner [das Buch]_j [___i versucht]. to read has nobody the book tried

b. PP scrambling, then AP topicalization

[Stolz t_j] ist der Fritz gestern [auf sein Kind]_j nicht _____i gewesen. proud is the Fritz yesterday of his child not been

c. DP scrambling, then VP topicalization

 $[t_j \text{ Gelesen}]_i$ hat $[\text{das Buch}]_j$ keiner. read has the book nobody

(26) An algorithm for checking c-command in chains

<u>Given</u>: D= (structure K, a candidate list of chains for K)
(i) Search for a pair of chain members (α, β) such that α c-commands β.
(ii) If search is successful, delete α and repeat (i). Otherwise terminate.
If all chains are now singleton, congratulations! You had good chains.

(27) The notion of "head of a chain" (for "overt" movement)

Given: D= (structure K, a candidate list of chains for K)

(i) Search for a pair of chain members (α, β) such that α c-commands β .

David Pesetsky, course materials for 24.960 Syntactic Models, Spring 2006. MIT OpenCourseWare (http://ocw.mit.edu/). Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

(ii) Copy any occurences of [+silent] in β to corresponding positions in α .

- (ii) Mark β [+silent].
- To terminate, the procedure should be bottom-to-top.

Incomplete constituent fronting: remnant movement analyses

- (28) a. **DP scrambling from VP, then infinitival I+V topicalization** $[t_j \operatorname{Zu} \operatorname{lesen}]_i$ hat keiner [das Buch]_j [_____i versucht]. to read has nobody the book tried
 - b. **DP scrambling from VP, then infinitival VP topicalization** $[t_j \text{ verkaufen}]_i \text{ wird er das Pferd } ___i \cdot \text{ sell-INF}$ will he the horse
 - c. DP scrambling from VP, then participial VP topicalization
 [t_j Gelesen]_i hat [das Buch]_j keiner _____i.
 read-PRT has the book nobody
 - d. **PP scrambling from DP, then DP topicalization** [Ein Buch t_j]_{*i*} hat Hans [über Syntax]_{*j*} _____*i* ausgeliehen. a book has Hans about syntax borrowed

a

e. **PP scrambling from AP, then AP topicalization** [Stolz t_j]_{*i*} ist der Fritz gestern [auf sein Kind]_{*j*} nicht __*i* gewesen. proud is the Fritz yesterday of his child not been

e'. **PP scrambling from AP, then AP** *wh*-movement [Wie stolz t_j] ist der Fritz gestern [auf sein Kind]_j nicht ____i gewesen? how proud

An argument for derivation:

- The c-command property of movement follows from viewing Move as a case of Merge and imposing an Extension condition (or Featural Cyclicity; see Richards).
- What might a pure representational theory of remnant movement do?

(29) An algorithm for checking c-command in (structure, chains) pairs

<u>Given</u>: D= (structure U, a candidate list of chains for U)

- (i) Search for a pair of chain members (α, β) such that α c-commands β .
- (ii) If search is successful, delete $\boldsymbol{\alpha}$ and repeat (i). Otherwise terminate.

If all chains are now singleton, congratulations! You had good chains. If not, sorry! Play again...

Where's the argument, you ask? The argument is: what on earth is (29)?

How do we pronounce instances of (overt) remnant movement?

(30) A derivational timing theory of chain pronunciation [overt movement] The head of a chain is the most recently merged member of the chain. Pronounce the head and don't pronounce the other members of a chain.

Can we do better? Fox & Pesetsky, etc....? Maybe.

(31) A representational theory of chain pronunciation.

Given: D= (structure U, a list of chains for U)

- (i) Search for a pair of chain members (α, β) such that α c-commands β .
- (ii) Copy any occurences of [+silent] in β to corresponding positions in α .
- (ii) Mark β [+silent].

Keep going until (ii) and (iii) cannot reapply.

An efficient application of (31) will apply bottom-to-top, essentially duplicating the derivation in a derivational theory.

• But: What if we consider a derivational counterpart to a "re-merge" theory of movement... [Carson Schütze, p.c. last time I taught this class]

8 Yes, but does Remnant Movement exist?

The "reanalysis" alternative to remnant movement:

Some homegrown terminology:

R = the fronted incomplete constituent (*r*emnant) **K** = the missing piece of R (*k*ompletion)

- K does not exit R by movement or its equivalent (e.g. presence in R's GAPS list).
- Instead, a process of reanalysis makes K and an unsaturated R co-arguments in a higher domain.

What both proposals have in common:

• The phenomenon of K exiting R should *ceteris paribus* be observable independent of the fronting of R.

What might distinguish the proposals:

Does the phenomenon of K exiting R look like other instances of movement?

9 Arguments for remnant movement from Müller and others vs. counter-arguments from Dekuthy & Meurers

$\mathbf{R} = \mathbf{DP}$

- Subject condition
- Müller (pp. 10ff): K may not exit a subject R

(32)a. K exits subject R...

*Worüber_{*j*} hat [ein Buch t_j] Karl beeindruckt? about what has a book-NOM Karl impressed

- b. **...and R fronts** *[Ein Buch t_j]_{*i*} hat Karl [über Syntax]_{*j*} <u>___</u>*i* beeindruckt.
- D-M, citing Haider (1993, 173): sometimes K from subject R is ok!
- (33)a. [Über Strauß]_j hat [ein Witz t_j]_i die Runde gemacht. about Strauss has a joke the round made 'A joke about Strauss went round.'
 - b, [Ein Witz t_i]_i hat [über Strauß]_i ____i die Runde gemacht.

[An unaccusativity/stage-level (Diesing 1990) effect?]

Specificity effect

- Müller: "yes"
- D-M, citing Pafel (1993): "no"

Possessive subject acts as an intervener

- Müller: "yes"
- D-M, citing Fanselow (1991): "no"

Freezing effect (Wexler & Culicover 1980)

- Müller: when R has moved, K may not exit
- (34) *Worüber_j hat [ein Buch t_j] keiner gelesen? about what has a book-ACC nobody read
- D-M, citing Fanselow (1991, 189): something else is going on
- (35)a.Worüber_j kann [einen Südkurier-Artikel t_j] selbst Peter nicht am Strand verfassen? about what can a Südkurier article even Peter not at the beach create

b.Worüber_j kann [einen SK-Artikel t_j] jeder Schwachkopf am Strand verfassen? every idiot

$\mathbf{R} = \mathbf{V}\mathbf{P}$

"Coherent verb" = restructuring verb [selects +LEX complement] "Incoherent verb" = non-restructuring verb [selects -LEX complement]

Scrambling of the argument of an embedded infinitival V is possible only in a coherent (restructuring) context:

(36)a. Scrambling: coherent context

daß das Pferd _j keiner	[zu verkaufen t	j] versucl	ht hat.
that the horse nobody	to buy	tried	has

b. Scrambling: incoherent context *...daß das Pferd_j keiner [zu verkaufen t_j] abgelehnt hat. disapproved

This contrast persists with fronting of R = VP

- (37) a. Scrambling + R fronting: coherent context .[Zu verkaufen t_j]_i versuchte er das Pferd_j ____i to buy tried he the horse
 - b. Scrambling + R fronting: incoherent context .*[Zu verkaufen t_j]_i empfahl er ihr das Pferd_j ____i to buy advised he her the horse
- Remnant movement advocates: so there!
- D-M: Not so fast...

(38)a. Scrambling [movement into the *Mittelfeld*] of a VP complement to an obligatorily coherent verb is impossible

*Er wird [das Pferd verkaufen]_j noch heute $__j$ wollen. He will [the horse sell] still today want-to

b. Topicalization of such a VP is ok

[Das Pferd verkaufen]_{*j*} wird er noch heute $__{j}$ wollen.

"Reanalysis-like approaches such as the argument-raising proposal developed [below] can account for the above data. The distinction between coherent and incoherent verbs is captured by specifying coherent verbs as obligatorily raising all complements of their verbal argument so that they become the arguments of a head cluster whereas the complement of an incoherently selecting verb is required to be a complete VP. As a result, in the case of a coherent verb there is no full VP that could be scrambled to obtain examples like (38a). In a sentence with an incoherently selected infinitive such a VP exists, and it can thus [scramble]. Finally, the requirement that coherently selected verbal complements combine in a verbal cluster does not extend to non-local dependencies, which makes it possible to license (38b) but exclude (38a)...."

Proposal (sketch)

Coherent verb:	selects [LEX +] complement
Incoherent verb:	selects [LEX -] complement
	[p. 26]

(39) **Basic lexical argument raising principle**

word]	SYN	$\operatorname{SUBJ}\left(1\right)$
SYN[HEAD VFO	$verb \qquad] \rightarrow$		COMPS raised $3 \oplus 2$
	VFORM base_]	ARG	$G-ST\left\langle 1 \mid 2 \right\rangle \wedge \left(3 \circ indep \right)$

The function *raised* applied to the raising source (tag 3) returns the unsaturated members of the raising source's COMPS list. It is defined for verbs and adjectives only if they are [LEX +], and is also defined for all nouns.

The conjunction in ARG-ST is an additional restriction on top of the requirement that ARG-ST be a list whose first member is the subject and whose other members appear on the COMPs list. The additional restriction requires everything except the raising source ("3") to be "independent" — i.,e. to have null COMPS).

The open circle is the shuffle operation.

- A gap may be [LEX -] and its filler [LEX +], thus allowing partial constituent topicalization (p. 36, top). [Only [LEX -] may topicalize (p. 35, para 2 line 5)]
- Scrambling of an embedded VP under an obligatorily coherent verb is impossible because the VP is [LEX -] and the verb requires a [LEX +] complement. Scrambling in the Mittelfeld is just an ordering fact about complements. [I think...]
- (40) Müller-Takano generalization: Also, if R scrambles, K may not have exited R via scrambling

a. DP scrambling, then infinitival clause scrambling

*...daß $[t_j \text{ zu lesen}]_i$ keiner [das Buch]_j [____i versucht hat]. that to read nobody the book tried has

cf. ok:

...dass [das Buch zu lesen]_i keiner $__i$ versucht hat.

b. PP scrambling, then AP scrambling

*...daß [t_j stolz] der Fritz gestern [auf sein Kind]_j nicht ____i gewesen ist. that proud the Fritz yesterday of his child not been is

c. DP scrambling, then VP topicalization

*daß [t _j gelesen] _i	[das Buch] _j	keiner \underline{i} hat.	
read	has the book	nobody	has