

Handout 3: Locality theory in Minimalism (I)

1. The role of government (hence also of barriers, CED and ECP)

- Government played a crucial role in attaining a high level of descriptive adequacy and rather appealing level of explanatory adequacy, showing properties of a deductive structure.
- With the rise of the Minimalist Program and its *Inclusiveness Condition*, government became regarded as an illicit theoretical object. For Chomsky, Inclusiveness was more a desideratum of linguistic theory than a condition (Chomsky 1995: 228):
 - A “perfect language” should meet the condition of inclusiveness: any structure formed by the computation is constituted of elements already present in the lexical items selected for N; no new objects are added in the course of the computation apart from the rearrangement of lexical properties (in particular, no indices, bar levels in the sense of X-bar theory, etc.; see note 7).

It is obvious from its very formulation that the inclusiveness condition was in direct opposition to the formal mechanisms of government, which often included indices, gamma-marks (Lasnik and Saito 1984), and various star-notations. However, as Chomsky rightly notes in note 7, there is nothing in inclusiveness that precludes the relationships that were captured by the indices, gamma-marks, and stars (Chomsky 1995: ch3 note 7):

- ... Thus with sufficiently rich formal devices (say, set theory), counterparts to any object (nodes, bars, indices, etc.) can readily be constructed from features. There is no essential difference, then, between admitting new kinds of objects and allowing richer use of formal devices; we assume that these (basically equivalent) options are permitted only when forced by empirical properties of language.

So it seems that given sufficient empirical justification, the government *relationship* could survive inclusiveness, as long as the indices and gamma-marks were simply a notation for the relation.

- In chapter 3, however, Chomsky had already begun to attack the empirical coverage of the ECP. Chomsky argued that many of the major reasons of being for the ECP, such as Superiority effects, Superraising, the HMC, and [Spec, CP] islands, can be analyzed as a violation of some sort of ‘shortest move’ requirement, such as the Minimal Link Condition (Chomsky 1995: 181). In fact, in footnote 19 of chapter 3, he is quite clear that the ECP has no theoretical status within the minimalist program he is elaborating: “Note that the ECP will now reduce to descriptive taxonomy, of no theoretical significance...”

- Unfortunately, shortest move approaches to locality such as the Minimal Link Condition could not supplant the ECP in accounting for the distribution of island effects in *wh*-in-situ languages: how could shortest move be violated for *wh*-adjuncts but not for *wh*-arguments?

2. The empirical domain of the ECP/CED:

- Overt long extraction asymmetries (subject-object, adjunct-complement)
- Covert asymmetries:
 - *ne* personne
 - superiority (subject-object, adjunct-complement)
 - cf: some 'pure superiority' effects
- Covert asymmetries: *wh*-in-situ languages
- Extraction out of any island/barrier: subject or adjunct movement out of any island (in addition to violation of CED, etc.).
- Weak island violations (some)
- A note on *where*, *when*, *how* and *why* w.r.t. overt vs. covert
 - cf. Huang 1982; Tsai 1994 i.a.; Caponigro and Pearl 2009)
- Adjunct extraction and predicate clefts in Vata (Koopman 1984)
- *Who did t leave early? (Who left early? Who didn't leave, who would leave, why did he leave early?)
- And more.

3. Locality in MP

- Many researchers have since attempted to derive the descriptive generalizations of the CED / ECP from minimalist assumptions without reference to the notion of government, with some success.
- On CED effects:
 - Nunes and Uriagereka (2000)
 - Ishii (1997)
 - Stepanov (2007)
 - Truswell (2007)
 - Barrie (2007)
 - Narita (2009)
- On ECP effects:
 - Minimal Link Condition, Attract Shortest, etc.

4. Deriving the CED (I): Nunes and Uriagereka 2000

4.1. Introduction

- Some relevant data regarding CED:
 - (1) Standard extraction out of a subject or an adjunct:
 - a. *[_{CP} [which politician]_i [_{C0} did+Q [_{IP} [pictures of t_i] upset the voters]]]
 - b. *[_{CP} [which paper]_i [_{C0} did+Q [_{IP} you read *Don Quixote* [_{PP} before filing t_i]]]]
 - (2) $\sqrt{\text{Parasitic gap constructions}}$
 - a. [_{CP} [which politician]_i [_{C0} did+Q [_{IP} [pictures of pg_i] upset t_i]]]]
 - b. [_{CP} [which paper]_i [_{C0} did+Q [_{IP} you read t_i [_{PP} before filing pg_i]]]]
 - (3) *Parasitic gaps embedded within a CED island such as an adjunct clause
 - a. *[_{CP} [which politician]_i [_{C0} did+Q [_{IP} you criticize t_i [_{PP} before [pictures of pg_i] upset the voters]]]]
 - b. *[_{CP} [which book]_i [_{C0} did+Q [_{IP} you finally read t_i [_{PP} after leaving the bookstore [_{PP} without finding pg_i]]]]]]
- The major locality condition in MP (Chomsky 1995:311):
 - (4) *Minimal Link Condition*
K attracts α only if there is no β , β closer to K than α , such that K attracts β .
- Minimal Link Condition **well explains** the unacceptability of (5)
 - (5) a. *[I wonder [_{CP} wh_i [_{C0} Q [_{IP} who [_{VP} bought t_i]]]]]]
 - b. [_{CP} Q [_{CP} who [_{VP} bought what]]]]

the interrogative complementizer Q should have attracted the closest *wh* element *who*, instead of attracting the more distant *what*.

- However, Minimal Link Condition **says nothing** about the CED effect such as shown in (1)- (3). (For instance, there is no *wh*-element other than *which politician* that Q could have attracted in (1))

❖ This paper argues:

1. first, that CED effects arise when a syntactic object that is required at a given derivational step has become inaccessible to the computational system at a previous derivational stage;
 2. Second, that the contrasts between (1) and (2), on the one hand, and between (2) and (3), on the other, are due to their different derivational histories.
- These results arise as by-products of two independent lines of research on the role of Kayne's (1994) Linear Correspondence Axiom (LCA) in the minimalist framework.

- Uriagereka's (1999) Multiple Spell-out system, which derives the induction step of the LCA by eliminating the unmotivated stipulation that Spell-out must apply only once;
- And Nunes's (1995, 1998) version of the copy theory of movement, which permits instances of sideward movement (i.e., movement between two unconnected syntactic objects) if the LCA is satisfied.)

4.2. Basic CED Effects

- To show how the standard CED effects illustrated in (1) can be accounted for within Uriagereka's (1999) Multiple Spell-out theory.

Some fundamental ideas of Multiple Spell-out System (Uriagereka 1999):

- LCA: Recasting Kayne's (1994) LCA in terms of Chomsky's (1995) bare phrase structure and simplify its definition by eliminating the recursive step, as formulated in (6):

(6) *Linear Correspondence Axiom (LCA)*

A lexical item α precedes a lexical item β iff α asymmetrically c-commands β .

(7) *Linearize:*

The operation that maps a phrase structure into a linear order of terminals in accordance with the LCA in (6).

- Multiple Spell-out:

(8) Assuming that **failure to yield a total ordering** among lexical items leads to an illicit derivation, the system is forced to employ **multiple applications of Spell-out**, targeting **chunks** of structure that Linearize can operate with.

- The number of applications of Spell-out is determined by **economy**.

- Labeling

(9) Assuming that the label of a given structure provides the "address" for the appropriate plugging in, in both the phonological and the interpretive components.

⇒ The syntactic object $K = \{\gamma, \{\alpha, \beta\}\}$, with label γ and constituents α and β (see Chomsky 1995:chap. 4), Spell-out ships $\{\alpha, \beta\}$ to the phonological and interpretive components, leaving K only with its label.

⇒ Once the constituent parts of K are gone, the computational system treats it as a lexical item. (To use the notation $K = [\gamma \langle \alpha, \beta \rangle]$ to represent K after it has been spelled out)

4.3. Examining Extraction out of a Subject

- (10) *Which politician did pictures of upset the voters?
- Derivation of (10) under a multiple spell-out analysis:
 - Assembling the structure K and L in (11) by successive applications of Merge.
- (11) a. K = [vP upset the voters]
 b. L = [pictures of which politician]
- The computational system applies Spell-out to L, allowing its constituents to be linearized in the phonological component, and merges the spelled-out structure L' with K.
 - To merge of *did* and movement of L' to [Spec,TP]. (Copy Theory involved)
 - The interrogative complementizer Q merges with TP and *did* adjoins to it
 - There is no element that can check the strong *wh*-feature of Q.
 - The derivation therefore crashes.
- Crucially, the *wh*-element of either copy of L = [pictures <pictures, of, which, politician>] became unavailable to the computational system after L was spelled out.
 - Extraction out of a subject is prohibited because, at the relevant derivational point, there is literally **no** syntactic object within the subject that could be copied.

4.4. Examining Extraction out of an Adjunct

- (12) *Which paper did you read *Don Quixote* before filing?

⇒ Same story with the extraction out of a subject

4.5. Examining Extraction out of a derived Subject (Subject in Passives)

- (13) a. *who_i was [_α a picture of t_i]_k taken t_k by Bill
 b. who_i did Bill take [_α a picture of t_i]

- Some other assumptions:

(14) Chain uniformity

(15) Economy (the number of multiple spell out)

(16) Cyclicity is inviolable: the derivation proceeds in cyclic fashion.

- Derivation of (13):
 - The NP is spelled out before being copied in (17a) and its copy merges with the whole structure, as shown in (17b)

- the two copies of the NP can then form a licit chain and the derivation converges.
- (17) a. [TP were [VP taken [some <some, pictures, of, John>] by Bill]]
 b. [TP [some <some, pictures, of, John>] [T' were [VP taken [some <some, pictures, of, John>] by Bill]]]
- Given (16), a simplified structure can be presented as:
- (18) [CP was+Q [CP [α <a, picture, of, who>] [VP taken [α <a, picture, of, who>] by Bill]]]
- The strong *wh*-feature of Q must be checked and **neither** copy of *who* is accessible to the computational system.
 - The derivation therefore crashes.

5. Deriving the CED (II): Stepanov (2007)

5.1. N&U (2000)'s structure building account

- (2) A structure- building approach: If a phrase marker X was assembled in parallel with a phrase marker Y, and then X and Y were merged, whereupon Y projects, no extraction is ever possible from X.
- (3) X is assembled in parallel with Y iff there exists a derivational point at which X and Y coexist in the derivational space and are unconnected.
- Empirical challenge: cross-linguistic variation

Japanese (Kikuchi 1987, Takahashi 1994; see also Ross 1967, Kuno 1973)"

- (4) [Op [Mary-ga t yonda no]-ga akirakana yorimo John-wa takusan-no
 Mary-nom read that-nom is-obvious than John-top many-gen
 hon-o yonda.
 book-acc read
 (*)John read more books than [that Mary read _] is obvious.
- (5) *[Op Bill-ga [Mary-ga t yonda kara] odorcita yorimo]
 Bill-nom Mary-nom read because was-surprised than
 John-wa takusan-no hon-o yonda.
 John-top many-gen book-acc read
 *John read more books than Bill was surprised because Mary read.

Russian

(6) a. S kem by ty xotel c'toby govorit' t bylo by odno udovol'stvie?
 with whom subj you wanted that-subj to-speak were subj one pleasure
 Lit. 'With whom would you want that [to speak _] were sheer pleasure?'

b. C' to by ty xotel c'toby kupit' ne sostavljalo by nikakogo truda? [Su]
 what subj you wanted that-subj to-buy not constitute subj no labor
 Lit. 'What would you want that [to buy _] would not be an trouble?'

→ The inadequacy of the uniform approach

→ An eclectic approach is justified

6. The eclectic, non-unifying approach

6.1. Extraction from subjects

- Lasnik and Saito (1990), etc.
- CED effects are seen in those languages (or constructions in a language) where the subject has not been moved.
- A kind of *Freezing Principle* (going back to Culicover 1975)
- Maybe reduced to some version of chain uniformity

6.2. Extraction from adjuncts: LAH The Late Adjunction Hypothesis

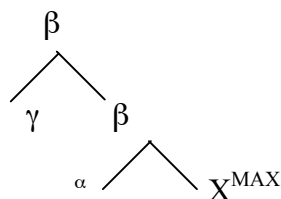
(7) ?* Who_i did Mary cry [after Peter hit t_i] ?

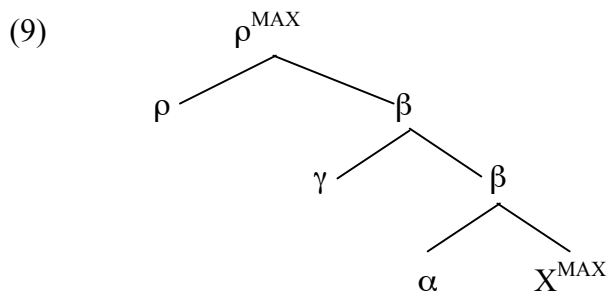
- Lebeaux (1988) makes this intuition precise by showing, on a solid empirical basis, that adjuncts can be Merged into the structure acyclically—that is, independently of the main cycle of the derivation.

More recent research has shown that the conclusion reached by Lebeaux must be strengthened to the extent that adjuncts not only can but must enter the structure late. Stepanov (2001b) adopts this viewpoint.

- Least Tampering condition: It is stated that given a choice of operation applying to a syntactic object labeled α , the computation must select one that does not change the set of c-command relations in that object.
- Substitution merge vs Adjunction merge. S.M is the main generator of cyclic portion of the derivation.

(8)





.....when all instances of substitution merge have applied, adjunctions may start.

Stepanov (2001b) states the criterion for determining structural argumenthood/adjuncthood in terms of uninterpretable features in the label of the element being Merged. In particular, if an element contains any uninterpretable feature (not necessarily relevant for Merger) then it has an additional property of entering the structure by substitution. Conversely, if an element does not contain any uninterpretable feature in its label, then it must enter the structure by adjunction and is therefore subject to LAH. The relevant uninterpretable features are taken to be *structural Case* and *the wh-feature*.

- Adjunct Condition effects with wh-movement arise because, at the time when the movement has to take place, the adjunct containing the moving element is unconnected to the phrase marker containing target of movement, due to the particular timing of structure building as encoded in the LAH.

7. Deriving the CED (III): Barrie (2007) based on cyclic linearization

Cyclic Linearization (Fox and Pesetsky 2006, Pesetsky and Fox 2003)

- Core idea: Linear order established at one stage cannot be altered at a later stage.
- Linearization domains: vP (VP) or CP
- Once a linearization domain is built, ordering statements are recorded.

(10) a. $[_{VP} A B C] \rightarrow \langle A, B \rangle, \langle A, C \rangle, \langle B, C \rangle$

- Subsequent operations may move elements, as long as previous ordering relations are not contradicted.

b. $[_{CP} A D [_{VP} t_A B C]]$ (ok)

c. $[_{CP} A D B [_{VP} t_A t_B C]]$ (ok)

d. $[_{CP} B D A [_{VP} t_A t_B C]]$ (crash): $\langle B, A \rangle$ contradicts $\langle A, B \rangle$

8. The CED a non-syntactic approach (Truswell 2007, 2008)

- Mainly on the Adjunct Condition