IGRA 01: Syntax III
Deriving Order
Gereon Müller (Universität Leipzig), July 7, 2014

1. Deriving Opacity Effects from the Order of Operations in Syntax

Observation: Both excitatory and inhibitory sequential interaction of operations can easily be captured in derivational models of grammar where all rules (more generally, building blocks) apply (or can apply) one after the other.

Challenge: How can the order of operations be determined?

Proposals:
- extrinsic ordering
- obligatoriness vs. optionality
- specificity
- anti-specificity
- strict cyclicity
- strata/levels
- rule vocabulary
- minimal search

1.1. Extrinsic Ordering


(1) **Extrinsic vs. intrinsic rule ordering** (Chomsky (1965)):
   a. In some cases, the order of rules is determined intrinsically, by the makeup of the rules.
   b. In some cases, the order of rules must be determined extrinsically, by stipulation; it follows from nothing.

Objection (Pullum (1979)):
The extrinsic/intrinsic distinction is not what is relevant. Suppose there is a universal principle determining that a rule with property A has to precede a rule with property B. Is this then an instance of extrinsic or intrinsic rule ordering?

(2) Pullum’s suggestion:
   a. *Parochial* rule ordering is language-specific (subject to variation) and not determined by general principles.
   b. *Universal* rule ordering follows from general properties of UG.

Massive extrinsic (or parochial) rule ordering is not really conceptually attractive; where possible, rule ordering should be made to follow from general principles.

(3) **Universally Determined Rule Application** (UDRA; Koutsoudas (1973), Pullum (1979)):
All restrictions on the application of rules are determined by universal principles.

1.2. Obligatoriness vs Optionality

(4) **Obligatory Precedence Principle** (Ringen (1972), Pullum (1979)):
Obligatory rules apply before optional rules.

Assumption (Pullum (1979)):
Reflexivization is obligatory, imperative subject deletion is optional.

(5) **Reflexivization** obligatory
   a. You kick you
   b. You kick yourself

(6) **Imperative subject deletion** optional
   a. You go home at once!
   b. Go home at once!

1.3. Specificity

Note:
Specificity has standardly been assumed for instances of *inhibitory simultaneous* interaction of building blocks (see, e.g., Elsewhere Condition/Subset Principle/Panimi’s Principle in morphology); but it can be extended to *sequential* interaction.

(7) **Specificity** (Sanders (1974), Pullum (1979)):
More specific rules apply before less specific rules.
(For any representation R, which meets the structural description of each of the two rules A and B. A takes applicational precedence over B with respect to R if and only if the structural description of A properly includes the structural description of B.)

(8) **Extraposition is more specific than *it*-deletion**
   a. Structural description (SD) of EXTRAP: X [ NP it ] S that/for X || X
   b. Structural description (SD) of ITDEL: X [ NP it ] S X || X

(9) **Extraposition precedes (and bleeds) *it*-deletion**
   a. *It that Purvis has been made Professor of Verbs is great
   b. It is great that Purvis has been made Professor of Verbs (extraposition)
   c. That Purvis has been made Professor of Verbs is great (i*-deletion)

Observation:
Extraposition is optional, *it*-deletion is obligatory. Wouldn’t the Obligatory Precedence Principle predict the reverse application?

Answer (Pullum (1979)):
Indeed, but the two principles are ranked: In the case of conflict, Specificity is respected, and
the Obligatory Precedence Principle is violated.

Note: Specificity as a means to predict the order of operations has more recently been re-discovered in minimalist syntax for operations like Move, Merge, and Agree; see van Koppen (2005), Lüne (2012), Georgi (2012).

Critical vs. intermediate movement steps (Georgi (2012)):

- There are two types of internal Merge: critical internal Merge and intermediate internal Merge.
- Internal Merge may bleed Agree with a subject (e.g., anti-agreement in Berber).
- Either both types of internal Merge bleed Agree, or none of them does, or critical internal Merge does and intermediate internal Merge does not; but the fourth possibility seems to be generally unavailable:
  - Critical internal Merge, intermediate internal Merge > Agree
  - Agree > critical internal Merge, intermediate internal Merge
  - Critical internal Merge > Agree > intermediate internal Merge
  - *Intermediate internal Merge > Agree > critical internal Merge

- Critical internal Merge (triggered by specific features) is inherently more specific than intermediate internal Merge (triggered by general edge features).
- This predicts a universal ordering of critical and intermediate Merge.

Observation:
Phenomena like anti-agreement with movement may hold for both critical movement and intermediate movement, or for none of them, or they may hold for critical movement, but not for intermediate movement steps.

(10) **Anti-agreement in Berber** (Ouhalla (1993)):

a. zi-n imdarn Mohand
   saw-3PL students Mohand
   ‘The students saw Mohand.’

b. man tamghart ay yzirin Mohand
   which woman COMP see.PART Mohand
   ‘Which woman saw Mohand?’

c. *man tamghart ay i-zaa Mohand
   which woman COMP 3sG.FEM-saw Mohand
   ‘Which woman saw Mohand?’

d. man tamghart ay bna-n qa t-zra Mohand
   which woman COMP S3M-3PL that 3sG-FEM-saw Mohand
   ‘Which woman did they say saw Mohand?’

(11) **Anti-agreement in Fiorentino** (Ouhalla (1993)):

a. Quante ragazze gli ha parlato con te?
   how-many girls CL.3sg have.3sg spoken to you
   ‘How many girls (it) has spoken to you?’

b. *Quante ragazze le hanno parlato con te?
   how-many girls cl.3pl have.3pl spoken to you
   ‘How many girls have spoken to you?’

c. Quante ragazze tu credi che gli hai telefonato?
   how-many girls you think that CL.3sg have.3sg phoned
   ‘How many girls do you think have phoned?’

d. *Quante ragazze tu credi che le hanno telefonato?
   how-many girls you think that CL.3PL have.3PL phoned
   ‘How many girls do you think have phoned?’

1.4. **Anti-Specificity**

Note: Chomsky (2000) adopts the opposite view.

(12) **Anti-Specificity** (Chomsky (2000; 2001; 2005; 2008)):
More general rules apply before less general rules.

Assumption: Move is defined in terms of more general Merge
(i) Merge
(ii) Move = Merge plus Agree plus Picked Piping
(iii) (13) follows from Anti-Specificity.

Chomsky (2000, 101): “Plainly Move is more complex than its subcomponents Merge and Agree, or even the combination of the two, since it involves the extra step of determining P(F) (generalized ‘pied-piping’). Good design conditions would lead us to expect that simpler operations are preferred to more complex ones so that Merge or Agree (or their combination) preempts Move.”

(13) **Merge before Move** (Chomsky (2000), Frampton & Gutman (1999)):
Suppose that the derivation has reached stage $\sum_n$, and $\Sigma_{n+1}$ is a legitimate instance of Merge, and $\Sigma_{n+1}$ is a legitimate instance of Move. Then, $\Sigma_{n+1}$ is to be preferred over $\Sigma_n$.

(14) **Expletive constructions in English**

a. There seems [TP t₁ to be [PP someone₂ in the room]]

b. *There seems [TP someone₂ to be [PP t₂ in the room]]

c. *[TP someone₂ [T to be [PP someone₂ in the room]]]

(15) **Optimization of derivational steps: T as input**

a. [T to be [PP someone₂ in the room]]

b. [TP there [T to be [PP someone₂ in the room]]]

(16) **Avoidance of the effect if no expletive is present in the numeration**

a. Someone₂ seems [TP t₂ to be t₂ in the room]
(17) A potential problem for Merge before Move, part (i):
   a. *John\textsubscript{1} expected [TP t\textsubscript{1} to be a proof\textsubscript{2} discovered]
   (Merge)
   b. John\textsubscript{1} expected [TP a proof\textsubscript{2} to be t\textsubscript{2} discovered]
   (Move)

(18) A potential problem for Merge before Move, part (ii):
   a. *C [TP Was a proof\textsubscript{2} discovered]
   (Merge)
   b. [CP C [TP A proof\textsubscript{2} was t\textsubscript{2} discovered]]
   (Move)

**Problem:**
Why is Move not blocked by Merge (of the external argument or the complementizer) after the generation of T' in (17) and (18)?

**Solution:**
Merge before Move in (14) demands that the preferred option be legitimate. Merge of the external argument John\textsubscript{1} in (17) violates the Theta-Criterion; and Merge of the complementizer in (18) violates subcategorization requirements of T.

(19) Another potential problem for Merge before Move, part (i):
   a. *H\textsubscript{1} seems [CP (that) t\textsubscript{1} was told John [CP that Bill left]]
   (Merge)
   b. H\textsubscript{1} seems [CP (that) John\textsubscript{2} was told [CP that Bill left]]
   (Move)

(20) Another potential problem for Merge before Move, part (ii):
   a. *H\textsubscript{1}'s fun [CP t\textsubscript{1} to [\textit{PRO} PRO\textsubscript{2} go to the beach]]
   (Merge)
   b. H\textsubscript{1}'s fun [CP PRO\textsubscript{2} to \textit{PRO} t\textsubscript{2} go to the beach]]
   (Move)
   c. *H\textsubscript{1} was decided [CP t\textsubscript{1} to [\textit{PRO} PRO\textsubscript{2} be executed at dawn]]
   (Merge)
   d. H\textsubscript{1} was decided [CP PRO\textsubscript{2} to [\textit{PRO} t\textsubscript{2} be executed at dawn]]
   (Move)

**Problem:**
Why doesn't the first example in (19) and in (20) block the second one because of Merge before Move? Here, the expletive must be part of the numeral.

Chomsky's (2000) solution via lexical subarrays:

Suppose ... that at each stage of the derivation a subset LA\textsubscript{i} is extracted, placed in active memory (the 'work space'), and submitted to the procedure L. When LA\textsubscript{i} is exhausted, the computation may proceed if possible. Or it may return to LA and extract LA\textsubscript{j}, proceeding as before. The process continues until it terminates. Operative complexity in some natural sense is reduced, with each derivation accessing only part of the LA. If the subarray in active memory does not contain EXPR, then Move can take place in the corresponding stage; if it does, Merge of EXPR preempts Move. The next step is to determine the subarrays LA\textsubscript{i} that can be selected for active memory. LA\textsubscript{i} should determine a natural syntactic object ... the counterpart to a proposition. ... LA\textsubscript{i} can then be selected straightforwardly: LA\textsubscript{i} contains an occurrence of C or of v ... exactly one occurrence if it is restricted as narrowly as possible.

Chomsky (2000)

Thus, each LA\textsubscript{i} corresponds to a phase.

**Consequence:**
In there-constructions (where Merge before Move effects obtain), the expletive and the DP always have to be in the same LA\textsubscript{i}. In it-constructions (where there are no Merge before Move effects), the expletive and the DP do not have to be in the same LA\textsubscript{i}; they can be in different subarrays.

**Note:**
There is more evidence for Merge before Move; see Frampton & Gutmann (1999), Horstein (2001; 2009), Castillo, Drury & Grohmann (2009), Boekx, Horstein & Nunes (2010), Drummond (2011), Weisser (2013), Wiktö (2013), among others.

1.4.1. Control into Adjuncts

An argument for Merge before Move from object control (Horstein (2001; 2009), Boekx, Horstein & Nunes (2010)):

Merge before Move, together with the Movement Theory of Control (MTC) and the idea of sideward movement, predicts that objects cannot control into adjuncts, whereas subjects can.

(21) No object control into adjuncts:
John\textsubscript{1} saw Mary\textsubscript{2} before PRO\textsubscript{1,2} leaving the party

**Horstein's analysis of the impossible derivation:**
(i) At the relevant point in the derivation, there are two workspaces: [before PRO\textsubscript{1,2} leaving the party] is in the first one, saw is in the second one: John\textsubscript{1} is still in the numeration.
(ii) For object control, Mary\textsubscript{2} would have to sideward-move out of the adjunct and attach to the main verb saw; given Merge before Move, the preferred option will be Merge of John\textsubscript{1} from the numeration, followed by movement of Mary\textsubscript{2} to matrix subject position, yielding subject control.

(22) Object control into complements (minimality):
John\textsubscript{1} persuaded Mary\textsubscript{2} [PRO\textsubscript{1,2} to leave]

1.4.2. Left-Subordinating and Constructions

Another argument for Merge before Move: Weisser (2013) on extraction from second conjuncts in the English left-subordinating and-constructions.

(23) Left-subordinating and-constructions: conditional interpretation (Culicover & Jackendoff (2007)):
(You drink) one more can of beer and I'm leaving

**Observation:**
The construction permits asymmetric extraction from only one conjunct; either the left one (which is irrelevant in the present context) or the right one.

(24) Extraction in apparent violation of the Coordinate Structure Constraint
a. "This is the loot $O_{p1}$ that | you just identify $t_1$ | and | we arrest the thief on the spot |

b. "This is the thief $O_{p2}$ that | you just identify the loot | and | we arrest $t_2$ on the spot |

c. *This is the thief $O_{p3}$ that | you have identified the loot | and | we have arrested $t_2$ on the spot |

Weisser's analysis: (i) The construction involves two TP's, $T_{p1}$ and $T_{p2}$.
(ii) Initially, $T_{p1}$ is a part of $T_{p2}$.
(iii) $T_{p2}$ is first merged with $\&$ and.
(iv) $T_{p1}$ undergoes movement out of $T_{p2}$ to SpecP.

(25) Structure of English left-subordinating 'and'-constructions:

$k_p T_{p1} [\& \& T_p t_1 \& \ldots \& ]$

(26) Coordinate Structure Constraint (based on Ross (1967)): In a coordinate structure $k_p A [\& \& B ]$, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

Why (34-b) is possible: Movement from $T_{p2}$ (first to an intermediate phase edge, viz. SpecP) precedes movement of $T_{p1}$ to SpecP; such an extraction happens at a stage of the derivation when there is no coordinate structure yet in the sense of (26).

This is thus an instance of counter-bleeding: Raising of $T_{p1}$ would bleed [further] extraction from $T_{p2}$ (via creation of a CSC island), but it doesn’t because it comes too late.

Question: Why is this option not available for regular coordination, given that there should also be a stage of the derivation where the second conjunct has merged with $\&$, and the first conjunct is not yet present?

Answer: For regular coordination, such a derivation is blocked by Merge before Move because the first conjunct is generated by Merge.

Final question: There is a lot of evidence for Merge before Move; is there evidence for Move before Merge in the literature? Perhaps there is (Shima (2000), II, e.g.); but there is certainly much less around, and it can arguably all be accounted for in some other way.

1.5. (Strict) Cyclicity

Observation (McCawley (1984; 1998)): Cyclic rule application predicts orders among operations; and the smaller the cyclic domains are, the more orders are predicted.

(27) Cyclic Principle:

When two operations can be carried out, where one applies to the cyclic domain $D_2$ and the other applies to the cyclic domain $D_{n-1}$ included in $D_2$, then the latter is applied first.

Side remark:

In what sense do the Cyclic Principle and the Strict Cycle Condition (see (28)) differ? And is the latter needed at all? It has sometimes been claimed that the Strict Cycle Condition is not needed if the Cyclic Principle is adopted; see Jacobson & Neubauer (1974); Pullum (1979). However, this is not the case: A rule applying in a higher cycle may change the context for the application of a rule in a lower cycle, such that, e.g., the second rule can now apply (as part of the higher cycle). This interaction is excluded by the Strict Cycle Condition, but not by the Cyclic Principle (under the standard reading at least).

(28) Strict Cycle Condition (Chomsky (1973), Perlmutter & Soames (1979)): Within the current cyclic $D_2$, a syntactic operation may not target a position that is included within another cyclic domain $D_{n-1}$ that is included in $D_2$.

Question: What does the Cyclic Principle predict for the order of Merge and Agree in (29), where $v$ has a subcategorization feature for external Merge of the external argument ([*D*]; and a probe feature for case assignment [*c:inf*] that will trigger Agree?

(29) Stage $\Sigma$: $\Sigma$

Options:

- Assumption: Each phrase is a cyclic node.
  - Then the Cyclic Principle does not predict any order of Merge and Agree here.
- Assumption: Each projection is a cyclic node.
  - Then the Cyclic Principle may or may not predict an order Agree before Merge, depending on the exact interpretation of "operation to the cyclic domain":
    - Agree applies in $v'$ Merge applies in $v$: Agree before Merge.
    - Agree applies in $v'$ Merge applies in $v'$ (both features are on $v'$): no order predicted.

1.6. Strata/Levels

Assumption: If two rules (more generally, building blocks) are assigned to different strata, or levels, which are in a fixed order, then the two rules will also have to apply in a fixed order.

A version of this proposal: Some rules apply in core areas of a grammatical component; other rules apply at interfaces (before or after the core component). Pullum (1979): Post-cyclic (post-syntactic, phonology-oriented) operations will always be fed and bled by cyclic (genuinely syntactic) operations. Pre-cyclic (pre-syntactic, lexicon-oriented) operations will never be fed or bled by syntactic operations.

More recent work: Post-syntactic operations always come too late to feed syntactic operations. Such counterfeeding effects have been argued for by Watanabe (2012) (on default agreement with numerals in Slavic); Embick (2000) (on deponent verbs in Latin), and Sauerland & Elbourne (2002) (on blocked scope inversion with scrambling in Japanese and German).

Note: More on interfaces later in the course.

1.7. Rule Vocabulary
Assumption (Arregi & Nevins (2012, ch. 6)): Assuming a Distributed Morphology approach, there are various operations that apply post-syntactically (after all regular syntactic operations) but before phonological realization: copying, fission, dissimilation, impoverishment, metathesis. Here the order is relevant, and it follows from how close to syntax or close to phonology, a given post-syntactic operation is: Rules where concepts like hierarchy play a role apply before rules that mention phonological features.

1.7.1. Minimal Search
Suggestion (Chomsky (2013), and Chomsky’s lecture 2 in his MIT class this spring (30:00-33:00)); If anything, Move should be simpler than Merge “since it requires vastly less search” because external Merge “must access the workspace of already generated objects and the lexicon”.

Norbert Hornstein in his blog entry from June 19, 2014 on this issue:

He notes that the ‘simplest’ application of Merge is one where you pick an expression X that is within another expression Y and combine X and Y. Thus I [internal]-Merge is the simplest application instance of Merge. The cognoscenti will recognize that this is not how Chomsky elaborated things before. In earlier versions, taking two things neither of which was contained in the other and merging them (viz. E-merge) was taken to be simpler. Not now, however. Chomsky does not go into why he changes his mind, but he hints that the issue is related to search. It is easier to find a term within a term than to find two terms in a workspace (especially one that contains a lexicon).
(I heard Chomsky analogize this to finding something in your pocket vs finding it on your desk, the former being clearly simpler. This clearly says something about Chomsky’s pockets versus his desk. But substitute purses or school bags for pockets and the analogy, at least in my case, strains. This said, I like this analogy better than Chomsky’s old rotary analogy in his motivation of numerations.)
So, the simplest operation is I-merge, E-merge being only slightly more complex, and so also available.

2. The No Tampering Condition
(30) The Order of Syntactic Operations

(31) Hypothesis:
(i) The order of Merge and Agree is determined parachorially.
(ii) All other orders are fixed; left precedes right.

Hypothesis: All the fixed orders can be shown to follow from a third-factor principle of efficient computation, viz., a version of the No Tampering Condition (NTC, Chomsky (2007; 2008; 2013)) that incorporates Pullum’s (1992) assumptions about the origins of the Cyclic Principle.

Observation: According to the original NTC, Merge of two syntactic objects leaves these objects unchanged. Arguably, this should be generalized, but then it cannot be categorical anymore (the more liberal Strict Cycle Condition, in contrast, is categorical): Operations like feature valuation...

---

2 See facultyoflanguage.blogspot.co.uk
by Agree, and generation of copies by Move (or adding an additional mother if multidominance is adopted) do change syntactic objects (see Branigan [2013], among others).

(32) Pulham’s evolutionary motivation:
   a. “Complex structures in language are assembled from well-formed parts which may be modified in the process of being concatenated [...] but retain much of their structural integrity.” (p. 227)
   b. “The only way to make a complex object that exhibits stability in the face of disruptions and accidents is to give it a hierarchical structure.” (p. 230)

(33) No Tampering Condition (NTC; new version):
   Minimize changes to existing structures.
   (The more deeply embedded the affected area is, the more the structure as a whole is changed.)

- **Merge vs. Agree:**
  The NTC must not discriminate between two operations that are radically different, like Merge and Agree (structure-building vs. modification of structures).

- **External Merge vs. Internal Merge:**
  External Merge adds an item at the top of the current structure; internal Merge requires access to a lower part of the existing structure (even though access is typically quite local, given the PIC).

- **Critical Internal Merge vs. Intermediate Internal Merge:**
  Critical movement steps typically (though not always, as in raising followed by wh-movement) imply that the moved item stays in place for the rest of the derivation. In contrast, intermediate movement steps, by definition, will lead to a disruption of existing structure on the next cycle.

- **Lower Intermediate Internal Merge vs. Higher Intermediate Internal Merge:**
  Intermediate movement steps for lower features will lead to fixed structures (i.e., criterial positions) earlier than intermediate movement steps for higher features. (This does not hold if the lower feature finds its criterial position in an even higher clause; but these derivations are typically ruled out as involving improper movement, e.g., by the Williams Cycle.)

- **Müller-Takano Generalization:**
  Other things being equal, moving the more inclusive category affects less structure. (Cf. number of c-command reversals. Possible extension: Minimality effects in general.)

- **External Agree vs. Internal Agree:**
  All Agree operations are dispreferred by the NTC, but Agree with a specifier affects less structure than Agree with, or into, the complement, so the former is preferred to the latter.

**Overall conclusion:**
(i) If there are elementary operations like Merge, Move, and Agree, they will interact; interaction has empirical consequences.
(ii) Such interaction leads to opacity effects (counters-bleeding, countere-feeding), which are ubiquitous in grammar, and which thus support a derivational approach to syntax.
(iii) Either all logically possible orders of operations are available (parochial ordering as a parameter), or there are restrictions. Assuming the latter, the parameter space can be significantly reduced.
(iv) A plausible third-factor principle that restricts possible orders of operations is the No Tampering Condition.

**References**
Branigan, Phil (2013); Cyclicity and the Approach the Probe Principle. Ms. Memorial University of Newfoundland.