A Local Reformulation of the Williams Cycle

Gereon Müller

Universität Leipzig

Workshop on Structure Building, Universität Konstanz, April 13-14, 2012

1. Introduction: Improper Movement

Generalization:
There is a correlation between the position targeted by a movement type (low vs. high) and the distance over which it can apply (short vs. long).

(1) *Scrambling vs. Long-Distance Scrambling in German*:

a. dass das Buch$_1$ keiner t$_1$ liest
that the book$_{acc}$ no-one-nom reads

b. *dass Karl das Buch$_1$ glaubt [$_{CP}$ dass keiner t$_1$ liest]
that Karl$_{nom}$ the book$_{acc}$ thinks that no-one$_{nom}$ reads

(2) *Pronoun fronting/object shift*:

a. dass es$_1$ Fritz t$_1$ gelesen hat
that it$_{acc}$ Fritz$_{nom}$ read has

b. *dass ich es$_1$ glaube [$_{CP}$ dass Fritz t$_1$ gelesen hat]
that I$_{nom}$ it$_{acc}$ think that Fritz$_{nom}$ read has

(3) *Raising vs. Super-Raising in English*:

a. Mary$_1$ seems [$_{TP}$ t$_1$ to like John]

b. *Mary$_1$ seems [$_{CP}$ t$'_1$ that t$_1$ likes John]

(4) *Clitic Climbing in Italian*:

a. Mario lo$_1$ vuole [$_{TP}$ leggere t$_1$]
Mario it wants to read

b. *Mario lo$_1$ odia [$_{CP}$ (t$'_1$) C [$_{TP}$ leggere t$_1$]]
Mario it hates to read

(5) *Extraposition* (Ross's (1967) Right Roof Constraint/Upward Boundedness Constraint):
*John always maintains [$_{CP}$ (t$'_1$) that [$_{DP}$ a review t$_1$] will appear shortly] whenever he is asked about it [$_{PP}$ of his new book]$_1$

Note:
The generalization is standardly accounted for by a conspiracy of two constraints.
(i) First, there is a locality constraint that permits extraction from a CP only via SpecC; e.g., the Phase Impenetrability Condition (PIC) in (6). This precludes skipping SpecC in (1-b), (2-b), (3-b), (4-b), and (5).
(ii) Second, there is a constraint on improper movement according to which movement to a TP-internal position may precede movement to SpecC, but not vice versa; this asymmetry can be taken to reflect the hierarchy of the target positions in the tree.

(6) *Phase Impenetrability Condition* (PIC; Chomsky (2000; 2001)):
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

2. Existing Analyses

2.1 Principle C

Refs.:
May (1979), Chomsky (1981)

Assumptions:
(i) Locally A-bar bound traces qualify as variables. (A trace whose immediate antecedent is in SpecC is locally A-bar bound.)
(ii) Variables obey Principle C of the Binding Theory: They must not be bound from an A-position.

Analysis:
Super-Raising is excluded by Principle C given that the initial trace qualifies as a variable: This trace in base position is A-bound from the matrix SpecT position. To extend this account to other phenomena, the respective movement types must be assumed to end up in A-positions, and the initial traces must also uniformly be locally A-bar bound; see Fanselow (1990) on scrambling in German.

2.2 Unambiguous Binding

Ref.:
Müller & Sternefeld (1993)

Claim:
A more general approach to improper movement is required because (a) scrambling in German is not A-movement (parasitic gaps, no licensing of reflexives and reciprocals, weak crossover (for some speakers), etc.); and (b) there are asymmetries between uncontroversial A-bar movement types as well, e.g., topicalization vs. wh-movement in German.

(7) *Topicalization vs. Wh-Movement from Wh-Islands in German* (Fanselow (1987)):

a. *Welches Radio$_1$ weißt du nicht [$_{CP}$ wie$_2$ C [$_{TP}$ man t$_1$ t$_2$ repariert]] ?
which radio know you not how one fixes

b. ?Radios$_1$ weiß ich nicht [$_{CP}$ wie$_2$ C [$_{TP}$ man t$_1$ t$_2$ repariert]]
radios know I not how one fixes

Assumptions:
(i) Different movement types are defined by targetting different landing sites: Scrambling targets SpecV, topicalization targets SpecTop, wh-movement targets SpecC, raising targets SpecT.
(ii) There is a constraint on uniform chains that makes use of these differences in landing sites, viz., the Principle of Unambiguous Binding (PUB).

(8) *Principle of Unambiguous Binding* (PUB):
A variable that is α-bound must be β-free in the domain of the head of its chain (where α and β refer to different types of positions).

Analysis:
In (1-b)–(4-b), (5) (and (7-a)), the derivations required by locality constraints violate the PUB because variables (i.e., locally A-bar bound traces) are ambiguously bound. (In (7), topicalization may use an embedded SpecTop escape hatch that is unavailable for wh-movement, because of the PUB.)

2.3 *The Williams Cycle*

Refs.:
Williams (1974; 2003), Sternefeld (1992), Grewendorf (2003; 2004), Abels (2008)

Assumptions (Williams):
(i) There are several (four or five) nested domains: S′ > S > Pred > VP (Williams (1974)), or embedded levels: FS, SS, CS, TS (Williams (2003)).
(ii) It follows from a Level Embedding Conjecture that operations that take place at one level cannot take place again at a higher, more comprehensive level, where other operations defining that level apply.

Origin of the name:
This was informally referred to as the 'Williams cycle' in Chomsky's 1974 MIT class lectures.

(9) *Generalized Ban on Improper Movement* (GBOIM, theorem; Williams (2003, 72)):
Given a Pollock/Cinque-style clausal structure X_1 > ... > X_n (where X_i takes X_{i+1}P as its complement), a movement operation that spans a matrix and an embedded clause cannot move an element from X_j in the embedded clause to X_i in the matrix, where i <j.

(10) *Prinzip der hierarchiekonformen Bewegung* (Sternefeld (1992)):
Bewegung von α nach β ist nur möglich, wenn der Typ(α) in der Hierarchie [Topik > SpecC > Adj > SpecI > SpecV] tiefer ist als der Typ(β).

(11) *Generalized Prohibition against Improper Movement* (GenPIM; Abels (2008)):
No constituent may undergo movement of type τ if it has been affected by movement of type σ, where $\tau < \sigma$ under UCOOL.

(12) *The Universal Constraint on Operational Ordering in Language* (UCOOL; Abels (2008)):
θ < scrambling < A-movement < wh < topicalization

(13) *Affectedness of constituents* (Abels (2008)):
A constituent α is affected by a movement operation iff
a. α is reflexively contained in the constituent created by movement, and
b. α is in a (reflexive) domination relation with the moved constituent.

Note:
The notion of affectedness is required because the GenPIM is supposed to restrict not only the interaction of movement operations applying to a single item, but also the interaction of movement operations applying to two different items that are base-generated in a dominance relation (surfing paths/freezing configurations and diving paths/remnant movement configurations). I will disregard these latter issues in what follows. (Ideally, the relevant data can be derived independently, without recourse to a theory of improper movement; in addition, it is not absolutely uncontroversial that the properties of the respective movement interactions are indeed identical.)

Analysis:
In all the approaches that rely on some version of the Williams Cycle, improper movement in (1-b)–(4-b), (5) (and elsewhere) can be accounted straightforwardly; in particular, movement from SpecC to SpecV, Specv, or SpecT is blocked.

2.4 *The Activity Condition*

Ref.:
Chomsky (2001) (and many others)

Assumption:
To be eligible for movement, an item must have an active feature sought by the movement-inducing head.

Analysis:
In English super-raising constructions, the moved DP has its ϕ- and case features checked in the lower TP; hence, it cannot be attracted by matrix T.

(14) *Super-Raising*:
*Mary$_1$ seems [$_{CP}$ t$'_1$ that t$_1$ likes John]

Problem (Nevins (2004)):
The Activity Condition is empirically problematic; in particular, it is at variance with the existence of non-nominative subjects in SpecT that have their ϕ- and case features checked independently (and earlier in the derivation).

2.5 *Feature Splitting*

Ref.:
Obata & Epstein (2011)

Assumptions:
(i) The PIC forces long-distance movement via SpecC.
(ii) Uninterpretable features (like case) are not permitted in the edge domain of a phase head (C) once the phase head's complement has undergone spell-out (based on Richards (2007), Chomsky (2008)).
(iii) In view of (ii), feature splitting must take place if a wh-subject is to undergo movement: The case/ϕ-features undergo movement to SpecT (under Agree with T, which has inherited the relevant probes from C); and the wh- (or Q-) feature undergoes a separate (but simultaneous) movement step to SpecC, as in (15).

(15) *Movement of a Wh-Subject*:
[$_{CP}$ Who$_{[wh]}$ C [$_{TP}$ who$_{[\phi],[case]}$ T [$_{vP}$ who$_{[wh],[\phi],[case]}$ left]]] ?

Analysis:
In cases like (16), matrix T does not find a matching goal: The copy in the lower SpecT position has undergone spell-out already, and the copy in the lower SpecC position does not have ϕ- and case features anymore.

(16) *Super-Raising of Wh-Phrases*:
*Who seems [$_{CP}$ who$_{[wh]}$ C [$_{TP}$ who$_{[\phi],[case]}$ [$_{T}$ will] [$_{vP}$ who$_{[wh],[\phi],[case]}$ leave]]] ?

Note:
This analysis should generalize to cases where the super-raised item is not a wh-phrase.

2.6 Problems With the Existing Analyses

2.6.1 Generality

The accounts developed in Chomsky (2001) and Obata & Epstein (2011) are confined to super-raising, and cannot be generalized to other cases of improper movement (like, e.g. long-distance scrambling in German) in any obvious way. In these other contexts, there is, by assumption, some head in the upper clause that attracts some item from the lower clause (i.e., that shares some feature with such an item) in a way that no other head (in the lower clause) does.

2.6.2 Locality

Except for, possibly, Chomsky (2001) and Obata & Epstein (2011), all analyses require scanning large amounts of syntactic structure. Therefore, they are incompatible with a strictly local derivational approach to structure-building that permits only a very small amount of accessible syntactic structure at any step of the derivation (given the PIC and Strict Cyclicity).

2.6.3 Promiscuity

Given the PIC and the assumption that CP, vP, and DP are phases, intermediate movement steps to Specv, SpecC, and SpecD are required for all movement types without necessarily giving rise to improper movement effects; i.e., these intermediate landing sites are highly promiscuous. Things get only worse if *all* intervening XPs must be crossed via intermediate movement steps to SpecX in the course of movement (Sportiche (1989), Takahashi (1994), Agbayani (1998); Chomsky (1995; 2005; 2008), Bošković (2002), Boeckx (2003), and Boeckx & Grohmann (2007), among many others). It seems that there are two straightforward options under such a view, and neither is unproblematic. First, only criterial positions (final landing sites of movement) count for improper movement. Then it is unclear how, e.g., long-distance scrambling via SpecC can be excluded. Second, all positions (including all non-criterial intermediate positions) count for improper movement. Then it is unclear how, e.g., long-distance wh-movement via matrix Specv can be permitted (given that long-distance scrambling targetting the same position needs to be ruled out).

Caveat:
Things are different if one assumes flavoured edge features (Abels (2012)). But these kinds of edge features cannot by themselves provide a comprehensive account of all relevant instances of improper movement (e.g., if two inherently feature-driven movement operations ending up in criterial positions are combined, flavoured edge features as such cannot rule out the combination as improper).

2.6.4 So?

Many cases of improper movement can in fact be derived differently. Concerning the phenomena tackled in Müller & Sternefeld (1993), this holds, e.g., for the asymmetry with topicalization from wh-islands vs. wh-movement from wh-islands in German (and other languages) (a maraudage effect under the Intermediate Step Corollary (see below) in Müller (2011)), and for the asymmetry with extraction from verb-second clauses in German (analyzed as a CED effect derivable from the PIC in Müller (2011)).
However, this does not hold for all cases, in particular, it does not hold for the core cases mentioned in the beginning (super-rasing, long-distance scrambling, etc.): No maraudage (no competing moved item), no CED islands involved (other items can be extracted in identical contexts).

Goal:
A local reformulation of the Williams Cycle that is compatible with a strictly derivational approach,

with extremely small accessible domains throughout (phases = phrases).

> *GenPIM [...] cannot be understood dirctly as a constraint on derivations (unless the standard assumption is given up that successive cyclic movement is launched before the target of movement is merged into the tree). I make no attempt to reformulate GenPIM in derivational terms here.* (Abels (2008))

3. Background: Edge Features and Successive-Cyclic Movement

Premise:
If edge feature insertion is free (or if phase heads simply have an edge property, cf. Chomsky (2008)), and if edge features are not flavoured (Abels (2012)), no restrictions can be imposed. So edge feature insertion is not free (Chomsky (2001)).

Assumptions (Müller (2011)):

- All phrases are phases.
- All syntactic operations are driven by designated features: Structure-building features ([•F•]) trigger internal and external Merge, and probe features ([∗F∗]) trigger Agree.
- Operation-inducing features are ordered; they show up on stacks, with a Last Resort condition demanding that only the topmost feature on a given stack can be discharged (and thereby deleted).
- Edge feature insertion effecting intermediate movement steps is restricted by the Edge Feature Condition (a modification of Chomsky's original proposal).

(17) *Edge Feature (EFC)*:
An edge feature [•X•] can be assigned to the head γ of a phase only if (a), (b), and (c) hold:
a. The phase headed by γ is otherwise not balanced.
b. γ has not yet discharged all its structure-building or probe features.
c. [•X•] ends up on top of γ's list of structure-building features.

Consequence:
Given these assumptions, MLC and CED follow from the PIC:
(i) MLC effects follow because the higher one of two items competing for movement to the domain of a movement-inducing head (i.e., the item that is merged later) ensures phase balance without edge feature insertion, which is therefore blocked; and subsequent movement of the lower item violates the PIC.
(ii) CED effects follow because edge feature insertion cannot take place for an item that is included in a last-merged specifier of a phase head, with the phase head qualifying as inert at this point; subsequent movement of such an item in a last-merged specifier then also violates the PIC.

A side effect:
Intermediate movement steps must take place before regular specifiers are merged; I refer to this as the Intermediate Step Corollary.

(18) *Intermediate Step Corollary*:
Intermediate movement steps to specifiers of X (as required by the PIC) must take place before a final specifier is merged in XP.

A problem that is acknowledged but not resolved in Müller (2011):
Edge feature insertion violates Inclusiveness.

A possible solution (Lahne (2009)):
Edge features do not exist as such; there is just an edge property (or a structure-building instruction: [• •]) that can be assigned to *some feature(s)* of a phase head, thereby creating an edge feature. Discharge (and deletion) of such derivative edge features then accounts for a generalization concerning the morphological form of intermediate reflexes of successive-cyclic movement.

(19) *Lahne's Generalization* (Lahne (2009, 60)):
Morphological reflexes of successive-cyclic movement involve elsewhere markers: If a language has different exponents in contexts with and without movement, then the marker that shows up in the movement context is less specific than the marker that shows up in the non-movement context.

Lahne's (2009) approach:
(i) Intermediate movement steps require an edge feature.
(ii) Edge features are not inserted as such; rather, an *edge feature property* [• •] is assigned (by edge property insertion rules) to a feature (or features) of the phase head.
(iii) The features that are thus affected trigger structure-building (i.e., intermediate movement) in syntax and undergo deletion.
(iv) The features that are deleted are not available for post-syntactic morphological realization anymore: *impoverishment* in the syntax.
(v) Consequently, a less specific, default exponent is chosen by spell-out.

Problem:
The feature may need to check something that is contradicted/not found on the moved item (e.g., an edge feature [•V•] is supposed to be able to attract a DP with a conflicting categorial feature; an edge feature [•voice:–ag•] may need to attract a wh-item with a conflicting feature value (viz., [voice:+ag]).

Conclusion:
(i) The idea to construct edge features from existing material on phase heads (rather than insert them out of nowhere) is on the right track.
(ii) The newly formed edge features are not exactly the features that show up on phase heads.
(iii) Additional assumptions are then required to capture Lahne's generalization.
(But there may well be a way out, assuming that categorial features percolate from X to XP; then (a defective version of) X may be used as an edge feature directly after all, without prior copying, and with XP retaining the categorial feature for subcategorization from outside.)

4. A Reformulation of the Williams Cycle

4.1 Assumptions

Basic assumptions:
(i) Edge features are defective copies of categorial features of phase heads.
(ii) Edge features successively value movement-related features of moved items, creating lists that record aspects of the derivational history of movement.
(iii) Such information is deleted when information of the same type is encountered.

(iv) When the final landing site is reached, f-seq must be respected on such lists.

Mechanics:

1. An edge feature is a defective copy of the categorial feature of a phase head accompanied by a structure-building instruction ([• •]):
   - [V] → [V] [•X_V•]
   - [v] → [v], [•X_v•]
   - [T] → [T], [•X_T•]
   - [C] → [C], [•X_C•]
   - etc.
2. As many edge features can be generated (by copying the categorial feature of the phase head) as are needed to effect intermediate movement steps of items, in accordance with the Edge Feature Condition.
3. The original content of the feature is lost (in the course of defective copying) to make the feature usable (i.e., there is no instruction anymore to merge an item with the exact same categorial feature as that of the phase head), but crucial aspects of the original information (viz., the categorial feature of the phase head) remain intact so as to make it possible to trace (recent) steps of the derivation: The categorial information is still there as part of the structure-building edge feature, but it does not by itself restrict the nature of the merge operation that the edge feature effects.
   (One can think of this in terms of a vaccination analogy: A virus, e.g., is rendered inactive by destroying its contents, but the protein shell is preserved.)
4. Movement-related features of moved items have lists as values:
   - [Σ:□] (scrambling)
   - [wh:□] (wh-movement)
   - [top:□] (topicaliztion)
   - [rel:□] (relativization)
   - [epp:□] (raising to SpecT, whatever ultimately underlies this feature)
5. Edge feature discharge involves valuation of the movement-related feature of the moved item by the (defective) categorial information on the phase head, so as to ensure complete matching of the two items. Categorial information is successively added on top of the list.
   - Merge(V:[•X_V•], DP:[wh:□]) → V DP:[wh:$\boxed{\text{V}}$]
   - Merge(v:[•X_v•], DP:[wh:$\boxed{\text{V}}$]) → v DP:[wh:$\boxed{\text{vV}}$]
   - Merge(T:[•X_T•], DP:[wh:$\boxed{\text{vV}}$]) → T DP:[wh:$\boxed{\text{TvV}}$]
   - Merge(C:[•X_C•], DP:[wh:$\boxed{\text{TvV}}$]) → C DP:[wh:$\boxed{\text{CTvV}}$]
6. When identical categorial information is added at the top, the original information is deleted at the bottom. (The system is derivational, and information gets lost during the derivation. Effectively, this instantiates a ban on recursion in feature value lists, possibly motivated by economy considerations.)

- Merge(C:[$\bullet X_C\bullet$], DP:[wh:$\boxed{\text{TvV}}$]) → C DP:[wh:$\boxed{\text{CTvV}}$]
- Merge(V:[$\bullet X_V\bullet$], DP:[wh:$\boxed{\text{CTvV}}$]) → V DP:[wh:$\boxed{\text{VCTv}\cancel{\text{V}}}$]
- Merge(v:[$\bullet X_v\bullet$], DP:[wh:$\boxed{\text{VCTv}\cancel{\text{V}}}$]) → v DP:[wh:$\boxed{\text{vVCT}\cancel{\text{vV}}}$]

7. When a moved item has reached its target position, it discharges the movement-related structure-building feature of the head; this feature must also carry the categorial information of the head it is associated with, e.g., [$\bullet wh_C\bullet$].

- Merge(v:[$\bullet X_v\bullet$], DP:[wh:$\boxed{\text{VCTv}\cancel{\text{V}}}$]) → v DP:[wh:$\boxed{\text{vVCT}\cancel{\text{vV}}}$]
- Merge(T:[$\bullet X_T\bullet$], DP:[wh:$\boxed{\text{vVCT}\cancel{\text{vV}}}$]) → T DP:[wh:$\boxed{\text{TvVC}\cancel{\text{TvV}}}$]
- Merge(C:[$\bullet wh_C\bullet$], DP:[wh:$\boxed{\text{TvVC}\cancel{\text{TV}}}$]) → C DP:[wh:$\boxed{\text{CTvVC}\cancel{\text{CTvV}}}$]

8. At this point, it is determined whether the information recording the intermediate landing sites conforms to the functional sequence (f-seq) independently established in syntactic structures (e.g., C≻T≻v≻V). This is the Williams Cycle.

- Merge(C:[$\bullet wh_C\bullet$], DP:[wh:$\boxed{\text{TvVC}\cancel{\text{TV}}}$]) → C DP:[wh:$\boxed{\text{CTvVC}\cancel{\text{CTvV}}}$]
- Check C DP:[wh:$\boxed{\text{CTvV}\cancel{\text{CTvV}}}$] → C DP:[wh:$^{ok}\boxed{\text{CTvV}}\ \cancel{\text{CTvV}}$]

(20) *Edge Feature Condition (EFC, revised)*:
An edge feature [$\bullet X_\gamma\bullet$] can be generated by copying the categorial feature of a head γ of a phase only if (a), (b), and (c) hold:
a. The phase headed by γ is otherwise not balanced.
b. γ has not yet discharged all its structure-building or probe features.
c. [$\bullet X\bullet$] ends up on top of γ's list of structure-building features.

(21) *Williams Cycle*:
Categorial information on a list of a movement-related feature α must conform to f-seq when α is checked by an inherent structure-building feature of a phase head (i.e., in criterial positions).

4.2 Simple Wh-Movement

(22) *Simple Wh-Movement in English*:
(I wonder) $[_{CP}\ \text{what}_2\ C\ [_{TP}\ \text{she}_1\ T\ [_{vP}\ t_1\ v\ [_{VP}\ \text{said}\ t_2\ ]]]]$

(23) *Derivation*:
a. $[_{V'}\ [_{V}\ \text{said}\ ]_{[\bullet X_V\bullet]}\ \text{what}_{[wh:\square]}\ ]$
b. $[_{VP}\ \text{what}_{[wh:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{said}\ ]]]$ (no Anti-Locality, pace Abels (2003)/Grohmann (2003))
c. $[_{v'}\ v_{[\bullet X_v\bullet]\succ[\bullet D\bullet]}\ [_{VP}\ \text{what}_{[wh:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]$
d. $[_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v_{[\bullet D\bullet]}\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]$
e. $[_{vP}\ \text{she}\ [_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]$ (Intermediate Step Corollary)
f. $[_{T'}\ T_{[\bullet X_T\bullet]\succ[\bullet epp\bullet]}\ [_{vP}\ \text{she}\ [_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]$
g. $[_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T_{[\bullet epp\bullet]}\ [_{vP}\ \text{she}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]$
h. $[_{TP}\ \text{she}\ [_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]$
i. $[_{C'}\ C_{[\bullet wh\bullet]}\ [_{TP}\ \text{she}\ [_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]]$
j. $[_{CP}\ \text{what}_{[wh:\boxed{CTvV}]}\ [_{C'}\ C\ [_{TP}\ \text{she}\ [_{T'}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]]]$ (^{ok}f-seq)

4.3 Long-Distance Wh-Movement

(24) *Long-Distance Wh-Movement in English*:
What_2 do you think $[_{CP}\ C\ [_{TP}\ \text{she}_1\ T\ [_{vP}\ t_1\ v\ [_{VP}\ \text{said}\ t_2\ ]]]]$?

(25) *Derivation*:
a. $[_{V'}\ [_{V}\ \text{said}\ ]_{[\bullet X_V\bullet]}\ \text{what}_{[wh:\square]}\ ]$
b. $[_{VP}\ \text{what}_{[wh:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]$
c. $[_{v'}\ v_{[\bullet X_v\bullet]\succ[\bullet D\bullet]}\ [_{VP}\ \text{what}_{[wh:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]$
d. $[_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v_{[\bullet D\bullet]}\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]$
e. $[_{vP}\ \text{she}\ [_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]$ (Intermediate Step Corollary)
f. $[_{T'}\ T_{[\bullet X_T\bullet]\succ[\bullet epp\bullet]}\ [_{vP}\ \text{she}\ [_{v'}\ \text{what}_{[wh:\boxed{vV}]}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]$
g. $[_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T_{[\bullet epp\bullet]}\ [_{vP}\ \text{she}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]$
h. $[_{TP}\ \text{she}\ [_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]$
i. $[_{C'}\ C_{[\bullet X_C\bullet]}\ [_{TP}\ \text{she}\ [_{T'}\ \text{what}_{[wh:\boxed{TvV}]}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]]$
j. $[_{CP}\ \text{what}_{[wh:\boxed{CTvV}]}\ [_{C'}\ C\ [_{TP}\ \text{she}\ [_{T'}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{said}\ ]\ ]]]]]]]]]$
k. $[_{V'}\ [_{V}\ \text{think}]_{[\bullet X_V\bullet]}\ [_{CP}\ \text{what}_{[wh:\boxed{CTvV}]}\ [_{C'}\ C\ [_{TP}\ \text{she said}\ ]]]]$
l. $[_{VP}\ \text{what}_{[wh:\boxed{VCTv\cancel{V}}]}\ [_{V'}\ [_{V}\ \text{think}]\ [_{CP}\ [_{C'}\ C\ [_{TP}\ \text{she said}\ ]]]]]$
m. $[_{v'}\ v_{[\bullet X_v\bullet]\succ[\bullet D\bullet]}\ [_{VP}\ \text{what}_{[wh:\boxed{VCTv\cancel{V}}]}\ [_{V'}\ [_{V}\ \text{think}]\ [_{CP}\ [_{C'}\ C\ [_{TP}\ \text{she said}\ ]]]]]]$
n. $[_{v'}\ \text{what}_{[wh:\boxed{vVCT\cancel{vV}}]}\ [_{v'}\ v_{[\bullet D\bullet]}\ [_{VP}\ [_{V'}\ [_{V}\ \text{think}]\ [_{CP}\ [_{C'}\ C\ [_{TP}\ \text{she said}\ ]]]]]]]$
o. $[_{vP}\ \text{you}\ [_{v'}\ \text{what}_{[wh:\boxed{vVCT\cancel{vV}}]}\ [_{v'}\ v\ [_{VP}\ [_{V'}\ [_{V}\ \text{think}]\ [_{CP}\ [_{C'}\ C\ [_{TP}\ \text{she said}\ ]]]]]]]]$
p. $[_{T'}\ T_{[\bullet X_T\bullet]\succ[\bullet epp\bullet]}\ [_{vP}\ \text{you}\ [_{v'}\ \text{what}_{[wh:\boxed{vVCT\cancel{vV}}]}\ [_{v'}\ v\ [_{VP}\ \text{think she said}\ ]]]]]$
q. $[_{T'}\ \text{what}_{[wh:\boxed{TvVC\cancel{TvV}}]}\ [_{T'}\ T_{[\bullet epp\bullet]}\ [_{vP}\ \text{you}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ \text{think she said}\ ]]]]]]$
r. $[_{TP}\ \text{you}\ [_{T'}\ \text{what}_{[wh:\boxed{TvVC\cancel{TvV}}]}\ [_{T'}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ \text{think she said}\ ]]]]]]]$
s. $[_{C'}\ C_{[\bullet wh\bullet]}\ [_{TP}\ \text{you}\ [_{T'}\ \text{what}_{[wh:\boxed{TvVC\cancel{TvV}}]}\ [_{T'}\ T\ [_{vP}\ [_{v'}\ [_{v'}\ v\ [_{VP}\ \text{think she said}\ ]]]]]]]]$
t. $[_{CP}\ \text{what}_{[wh:\boxed{CTvVC\cancel{TvV}}]}\ [_{C'}\ C\ [_{TP}\ \text{you think she said}\ ]]]$ (^{ok}f-seq)

4.4 Clause-Bound Scrambling

(26) *Clause-Bound Scrambling in German*:

dass das Buch_1	keiner	t_1	liest
that the book_{acc}	no-one-nom		reads

(27) *Derivation*:
a. $[_{V'}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\square]}\ [_{V}\ \text{liest}\ ]_{[\bullet X_V\bullet]}\ ]$
b. $[_{VP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]$
c. $[_{v'}\ [_{VP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]\ v_{[\bullet D\bullet]\succ[\bullet\Sigma\bullet]}\ ]$
d. $[_{v'}\ \text{keiner}\ [_{v'}\ [_{VP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{V}]}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]\ v_{[\bullet\Sigma\bullet]}\ ]]$
e. $[_{vP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{vV}]}\ [_{v'}\ \text{keiner}\ [_{v'}\ [_{VP}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]\ v\ ]]]$
f. $[_{TP}\ [_{vP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{vV}]}\ [_{v'}\ \text{keiner}\ [_{v'}\ [_{VP}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]\ v\ ]]]\ T\ ]$ (^{ok}f-seq)
g. $[_{CP}\ [_{C}\ \text{dass}\ ]\ [_{TP}\ [_{vP}\ [_{DP}\ \text{das Buch}\ ]_{[\Sigma:\boxed{vV}]}\ [_{v'}\ \text{keiner}\ [_{v'}\ [_{VP}\ [_{V'}\ [_{V}\ \text{liest}\ ]]]\ v\ ]]]\ T\ ]]$

4.5 The Ban on Long-Distance Scrambling

(28) *The Ban on Long-Distance Scrambling in German*:

*dass Karl das Buch glaubt dass keiner t_1 liest
that Karl$_{nom}$ the book$_{acc}$ thinks that no-one$_{nom}$ reads

(29) *Derivation*:

a. [$_{V'}$ [$_{DP}$ das Buch]$_{[\Sigma:\Box]}$ [$_{V}$ liest]$_{[\bullet X_V \bullet]}$]
b. [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{V}}]}$ [$_{V'}$ [$_{V}$ liest]]]
c. [$_{v'}$ [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{V}}]}$ [$_{V'}$ [$_{V}$ liest]]] v$_{[\bullet X_v \bullet] \succ [\bullet D \bullet]}$]
d. [$_{v'}$ [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{V}}]}$ [$_{V'}$ [$_{V}$ liest]]] v$_{[\bullet X_v \bullet] \succ [\bullet D \bullet]}$]
e. [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vV}}]}$ [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{V}$ liest]]] v$_{[\bullet D \bullet]}$]]
f. [$_{vP}$ keiner [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vV}}]}$ [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{V}$ liest]]] v]]]
g. [$_{T'}$ [$_{vP}$ keiner [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vV}}]}$ [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{V}$ liest]]] v]]] T$_{[\bullet X_T \bullet]}$]
h. [$_{TP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{TvV}}]}$ [$_{T'}$ [$_{vP}$ keiner [$_{v'}$ [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{V}$ liest]]] v]]] T]]
i. [$_{C'}$ [$_{C}$ dass]$_{[\bullet X_C \bullet]}$ [$_{TP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{TvV}}]}$ [$_{T'}$ [$_{vP}$ keiner liest v] T]]]
j. [$_{CP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{CTvV}}]}$ [$_{C'}$ [$_{C}$ dass] [$_{TP}$ [$_{T'}$ [$_{vP}$ keiner liest v] T]]]]
k. [$_{V'}$ [$_{CP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{CTvV}}]}$ [$_{C'}$ [$_{C}$ dass] [$_{TP}$ keiner liest v T] [$_{V}$ glaubt]$_{[\bullet X_V \bullet]}$]]]
l. [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{VCTv}\cancel{\text{V}}}]}$ [$_{V'}$ [$_{CP}$ [$_{C'}$ [$_{C}$ dass] [$_{TP}$ keiner liest v T] [$_{V}$ glaubt]]]]]
m. [$_{v'}$ [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{VCTv}\cancel{\text{V}}}]}$ [$_{V'}$ [$_{CP}$ dass keiner liest] [$_{V}$ glaubt]]] v$_{[\bullet D \bullet] \succ [\bullet \Sigma \bullet]}$]
n. [$_{v'}$ Karl [$_{v'}$ [$_{VP}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{VCTv}\cancel{\text{V}}}]}$ [$_{V'}$ [$_{CP}$ dass keiner liest] [$_{V}$ glaubt]]] v$_{[\bullet \Sigma \bullet]}$]]
o. [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vVCT}\cancel{\text{vV}}}]}$ [$_{v'}$ Karl [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{CP}$ dass keiner liest] [$_{V}$ glaubt]]] v]]]
p. [$_{vP}$ [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vVCT}\cancel{\text{vV}}}]}$ [$_{v'}$ Karl [$_{v'}$ [$_{VP}$ [$_{V'}$ [$_{V}$ glaubt]]] v] [$_{CP}$ dass keiner liest]]]] (extraposition, features ignored for perspicuity)
q. [$_{T'}$ [$_{vP}$ [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vVCT}\cancel{\text{vV}}}]}$ [$_{v'}$ Karl [$_{v'}$ glaubt dass keiner liest]]]] T$_{[\bullet epp \bullet]}$] (EPP is optional in German; Grewendorf (1989))
r. [$_{TP}$ Karl [$_{T'}$ [$_{vP}$ [$_{v'}$ [$_{DP}$ das Buch]$_{[\Sigma:\boxed{\text{vVCT}\cancel{\text{vV}}}]}$ [$_{v'}$ [$_{v'}$ glaubt dass keiner liest]]]] T]] (*f-seq → *Williams Cycle → crash)

Note:
The same consequence arises if long-distance scrambing targets SpecV rather than Specv (which might be an option yielding the same string given that EPP is optional in German). The only difference would be a fatal (f-seq-violating) value *$\boxed{\text{VCTv}\cancel{\text{V}}}$ of Σ on DP instead of *$\boxed{\text{vVCT}\cancel{\text{vV}}}$.

Note:
It is also predicted that a wh-phrase that undergoes long-distance movement cannot feed eventual wh-movement by intermediate, feature-driven long-distance scrambling to, say, Specv (as opposed to using Specv as an escape hatch provided by an edge feature). Assuming that it is the presence of a criterial (Σ) configuration that helps to avoid weak crossover configurations (for many speakers) and superiority configurations in German, the data show that true long-distance scrambling is not an option for long-distance wh-movement (Frey (1993), Büring & Hartmann (1994), Fanselow (1996), Heck & Müller (2000), Pesetsky (2000)).

(30) *Superiority Effects*:

a. (Ich weiß nicht) was$_2$ C wer$_1$ t_2 gesagt hat
I know not what$_{acc}$ who$_{nom}$ said has

b. *Wen$_2$ hat wer$_1$ geglaubt [$_{CP}$ dass der Fritz t_2 mag] ?
whom$_{acc}$ has who$_{nom}$ believed that the Fritz likes

(31) *Weak Crossover Effects*:

a. ?Wen$_1$ mag seine$_1$ Mutter t_1 ?
whom likes his mother

b. *Wen$_1$ hat seine$_1$ Mutter gesagt [$_{CP}$ dass wir t_1 einladen sollten] ?
whom has his mother said that we invite should

4.6 Super-Raising

(32) *Simple Super-Raising*:
*Mary$_1$ seems t_1' that t_1 likes John

Analysis:
The relevant movement-related feature on *Mary* is (by assumption) [epp]; matrix T (by assumption) bears the corresponding structure-building feature [•epp•]. Movement must take place via the embedded TP and CP domains, and via the matrix VP and vP domains. In the final matrix SpecT position where [•epp•] is discharged with the moved DP, [epp] on DP has the value $\boxed{\text{TvVC}\cancel{\text{T}}}$, which fatally violates f-seq (hence, the Williams Cycle) because C has not yet been removed.

(33) *Super-Raising Feeding Wh-Movement*:

a. *Who$_1$ t_1'' seems [$_{CP}$ t_1' C t_1 will leave] ?
b. *What$_1$ t_1'' seems [$_{CP}$ t_1' that it was said t_1] ?

Analysis:
The reasoning is identical. EPP-driven movement to matrix SpecT gives rise to a violation of the Williams Cycle which cannot subsequently be made undone by matrix wh-movement (there is no back-tracking or look-ahead).

4.7 Other Local Movement Types

Observation:
Other movement types that target positions in the TP, vP or VP areas can also not apply long-distance via CP, and for the same reason: When the (criterial) target position is reached, there will at least be an f-seq-violating C on the list of the movement-related feature on the moved item, and so a violation of the Williams Cycle will be unavoidable.

5. Extensions

5.1 DP-Internal PP Preposing

Note:
German has a movement type that involves PPs and targets SpecD (Lindauer (1995)).

(34) *DP-Internal Preposing of PPs in German*:

[$_{DP_2}$ [$_{PP_1}$ Über die Liebe] [$_{D'}$ das/ein Gerücht t_1]] kenne ich t_2
about the love the/a rumour know I

Assumption:
Suppose the f-seq is CTvVDNP. Then, DP-internal PP Preposing in CNPC configurations should additionally violate the Williams Cycle, in contrast to long-distance PP wh-movement. Long-distance PP scrambling is also predicted to violate the Williams Cycle in addition. (In (34), the list on the relevant feature, say [ω], of the moved PP in the final position respects the Williams Cycle: [ω: DN]. Predictions for P stranding may also arise under these assumptions.)

(35) *DP-Internal PP Preposing and the Williams Cycle* :

a.??[$_{PP_1}$ Über die Liebe] kenne ich [$_{DP}$ das/ein Gerücht [$_{CP}$ dass sie ein Buch t_1 geschrieben
about the love know I the/a rumour that she a book written
hat]]
has
Williams Cycle: ^{ok}[wh: CTvVDN~~CTvVDN~~]

b. *[$_{DP}$ PP$_1$ Über die Liebe] [$_{D'}$ das/ein Gerücht [$_{CP}$ dass sie ein Buch t_1 geschrieben hat]]]
about the love the/a rumour that she a book written has
kenne ich
know I
Williams Cycle: *[ω: DNCTvV~~DN~~]

c. *Es kennt [$_{PP_1}$ über die Liebe] keiner [$_{D'}$ das/ein Gerücht [$_{CP}$ dass sie ein Buch t_1
EXPL knows about the love no-one the/a rumour that she a book
geschrieben hat]]]
written has
Williams Cycle: *[Σ: vVDNCT~~vVDN~~]

5.2 Relativization

Observation (Bayer & Salzmann (2009); also Plank (1981)):
Many speakers of German do not permit long-distance relativization (in contrast to wh-movement or topicalization).

(36) *Impossible Long-Distance Relativization in German*:

a. *Das ist einer [$_{RelP}$ der$_1$ ich glaube [$_{CP}$ dass t_1 das schaffen wird]]
this is one who I believe that this manage will

b. *der Mann [$_{RelP}$ den$_1$ ich glaube [$_{CP}$ dass Maria t_1 liebt]]
the man who I believe that Maria loves

Analysis:
This follows if [C > Rel] according to f-seq. Then, the relative pronouns in (36) eventually bear a feature [rel] that violates the Williams Cycle: *[rel: RelTvVC~~T~~(~~vV~~)]

5.3 ECM Constructions

Observation (Abels (2008)):
A strict interpretation of the Williams Cycle (like the one adopted here) is problematic if ECM is analyzed as raising to object; the relevant movement-related feature on the raised object would then have a value vVT~~vV~~, which would violate the Williams Cycle. (Raising to object cannot possibly be assumed to reach the TP domain, as would be required.)
However, it is unclear whether there is raising to object.

(37) *Evidence for ECM Analyses* (Stowell (1991)):

a. John promised repeatedly to leave

b. John believed Mary repeatedly to have left (she left repeatedly)

c. *John believed Mary sincerely to have left

(38) *Conflicting Evidence for Raising to Object Analyses* (Lasnik (1999)):
?The DA proved the defendants$_1$ to be guilty during each other$_1$'s trials.

5.4 Languages that Permit Long-Distance Scrambling

Observation:
Languages like Japanese, Korean, and Russian have long-distance scrambling from uncontroversial CPs.

(39) *Long-Distance Scrambling in Russian* (Zemskaja (1973)):

a. Ty doktor$_1$ videl [$_{CP}$ kogda t_1 pod"ezžal] ?
you doctor$_{nom}$ saw when came

b. Vy pocylku$_1$ videli [$_{CP}$ kak zapakovali t_1] ?
you$_{plur}$ parcel$_{acc}$ saw how (they-)did-up

Analysis:
(i) Categorial features may well have a fine structure, in the sense that they consist of more primitive features (cf. Chomsky's [±N], [±V] yielding natural classes of V, N, A, P).
(ii) Suppose that [θ], [case], [A-bar] are more primitive features that capture natural classes of categorial features. E.g., [top], [rel], [wh] are all composed of [A-bar] plus other features; and the same may hold for the heads of landing sites of long-distance scrambling.
(iii) As a parametrized option, edge features may also be even more defective and impoverished copies of the original categorial feature, such that only a subfeature (that captures a natural class) is copied; e.g., [$\bullet X_{A-bar}\bullet$] instead of, say, [$\bullet X_{wh}\bullet$] and [$\bullet X_{\Sigma}\bullet$].
(iv) Discharge of [$\bullet X_{A-bar}\bullet$] in a scrambling position will then suffice to delete a lower [$\bullet X_{A-bar}\bullet$] based on earlier movement through SpecC, thereby satisfying the Williams Cycle.
(v) This option would also work for a raising to object analyses of ECM constructions.

References

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2008): Towards a Restrictive Theory of (Remnant) Movement, *Linguistic Variation Yearbook* 7, 53–120.

Abels, Klaus (2012): Syntax with a Gilded Edge; Or, Why Edge Features Come in Flavours. Talk, Workshop on Structure Building. Ms., University College London.

Agbayani, Brian (1998): Feature Attraction and Category Movement. PhD thesis, UC Irvine.

Bayer, Josef & Martin Salzmann (2009): That-Trace Effects and Resumption as the Result of Improper Movement. Ms., Universität Konstanz.

Boeckx, Cedric (2003): *Islands and Chains. Resumption as Stranding*. Benjamins, Amsterdam.

Boeckx, Cedric & Kleanthes K. Grohmann (2007): Putting Phases in Perspective, *Syntax* 10, 204–222.

Bošković, Željko (2002): A-Movement and the EPP, *Syntax* 5, 167–218.

Büring, Daniel & Katharina Hartmann (1994): The Dark Side of Wh-Movement, *Linguistische Berichte* 149, 56–74.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. In: R. Martin, D. Michaels & J. Uriagereka, eds., *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2005): Three Factors in Language Design, *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam (2008): On Phases. In: R. Freidin, C. Otero & M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.

Fanselow, Gisbert (1987): *Konfigurationalität*. Narr, Tübingen.

Fanselow, Gisbert (1990): Scrambling as NP-Movement. In: G. Grewendorf & W. Sternefeld, eds., *Scrambling and Barriers*. Benjamins, Amsterdam, pp. 113–140.

Fanselow, Gisbert (1996): The Proper Interpretation of the Minimal Link Condition. Ms. Universität Potsdam.

Frey, Werner (1993): *Syntaktische Bedingungen für die Interpretation*. Akademieverlag, Berlin.

Grewendorf, Günther (1989): *Ergativity in German*. Foris, Dordrecht.

Grewendorf, Günther (2003): Improper Remnant Movement, *Gengo Kenkyo: The Journal of the Linguistic Society of Japan* 123, 47–94.

Grewendorf, Günther (2004): Die linke Bewegung von Restkategorien. Ms., Universität Frankfurt/Main.

Grohmann, Kleanthes K. (2003): *Prolific Domains: On the Anti-Locality of Movement Dependencies*. John Benjamins, Amsterdam & Philadelphia.

Heck, Fabian & Gereon Müller (2000): Successive Cyclicity, Long-Distance Superiority, and Local Optimization. In: R. Billerey & B. D. Lillehaugen, eds., *Proceedings of WCCFL*. Vol. 19, Cascadilla Press, Somerville, MA, pp. 218–231.

Lahne, Antje (2009): Where There is Fire There is Smoke. Local Modelling of Successive-Cyclic Movement. PhD thesis, Universität Leipzig.

Lasnik, Howard (1999): *Minimalist Analysis*. Blackwell, Oxford.

Lindauer, Thomas (1995): *Genitivattribute: Eine morphosyntaktische Untersuchung zum deutschen DP/NP-System*. Niemeyer, Tübingen.

May, Robert (1979): Must Comp-To-Comp Movement Be Stipulated?, *Linguistic Inquiry* 10, 719–725.

Müller, Gereon (2011): *Constraints on Displacement. A Phase-Based Approach*. Vol. 7 of *Language Faculty and Beyond*, Benjamins, Amsterdam.

Müller, Gereon & Wolfgang Sternefeld (1993): Improper Movement and Unambiguous Binding, *Linguistic Inquiry* 24, 461–507.

Nevins, Andrew (2004): Derivations Without the Activity Condition. In: M. McGinnis & N. Richards, eds., *Proceedings of the EPP/Phase Workshop*. MITWPL, Cambridge, Mass.

Obata, Miki & Samuel David Epstein (2011): Feature-Splitting Internal Merge: Improper Movement, Intervention, and the A/A′ Distinction, *Syntax* 14, 122–147.

Pesetsky, David (2000): *Phrasal Movement and Its Kin*. MIT Press, Cambridge, Mass.

Richards, Marc (2007): Deriving the Edge: What's in a Phase?. Ms., Universität Leipzig.

Ross, John (1967): Constraints on Variables in Syntax. PhD thesis, MIT, Cambridge, Mass.

Sportiche, Dominique (1989): Le Mouvement Syntaxique: Contraintes et Paramètres, *Langages* 95, 35–80.

Sternefeld, Wolfgang (1992): Transformationstypologie und strukturelle Hierarchie. Ms., Universität Tübingen.

Stowell, Tim (1991): Small Clause Restructuring. In: R. Freidin, ed., *Principles and Parameters in Comparative Grammar*. MIT Press, Cambridge, Mass., pp. 182–218.

Takahashi, Daiko (1994): Minimality of Movement. PhD thesis, University of Connecticut.

Williams, Edwin (1974): Rule Ordering in Syntax. PhD thesis, MIT, Cambridge, Mass.

Williams, Edwin (2003): *Representation Theory*. MIT Press, Cambridge, Mass.

Zemskaja, E.A. (1973): *Russkaja Razgovornaja Reč'*. Nauka, Moskva.