On Deriving MLC Effects

Gereon Müller

Institut für Linguistik
Universität Leipzig

SoSe 2006

www.uni-leipzig.de/∼muellerg

Claim

Claim:
A derivational approach to syntax should minimize search space, its representational residue (Brody (2001)). Constraints that minimize search space should therefore be strengthened (Chomsky's (2000; 2001) PIC); constraints that presuppose search space should be abandoned (MLC).

Empirical domain:
MLC effects (superiority and superiority-like effects in German and English) can be derived from a strengthened version of the PIC (Chomsky (2000; 2001)) that holds for phrases rather than phases.

SSC and PIC

Note:
In Chomsky's (2000; 2001) system, the SCC and the PIC impose strong restrictions on active parts of derivations. The SCC restricts possible positions for the probe, and the PIC restricts the probe's search space, i.e., possible positions for the goal.

(1) Strict Cycle Condition (SCC) (Chomsky (1973)):
Within the current XP α, a syntactic operation may not target a position that is included within another XP β that is dominated by α.

(2) Phase Impenetrability Condition$_1$ (PIC$_1$) (Chomsky (2000, 108; 2001, 13)):
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.

Edge and Phase

(3) Edge (Chomsky (2001, 13)):
The edge of a head X is the residue outside of X′; it comprises specifiers of X (and adjuncts to XP).

(4) Phase:
The propositional categories CP and vP are phases; other XPs (except perhaps for DP) are not.

Consequence:
Suppose that ZP, XP, and UP are phases in (5). Then, in (5-a), operations can have a probe only in YP (SCC), and look for a goal only in YP or in the residue or head of XP (PIC_1). In the subsequent step (5-b), the probe must be in ZP, and the search space for a goal grows as indicated.

Search Space, PIC_1

(5) Search space under PIC_1:

a. $\overbrace{\overbrace{[_{YP}\ldots Y}^{SCC}\ [_{\underline{XP}}\ldots[_{X'}X}^{PIC_1}\ [_{WP}\ldots W\ [_{\underline{UP}}\ldots U\ldots\ ]]]]]$

b. $\overbrace{\overbrace{[_{\underline{ZP}}\ldots Z}^{SCC}\ [_{YP}\ldots Y\ [_{\underline{XP}}\ldots[_{X'}X}^{PIC_1}\ [_{WP}\ldots W\ [_{\underline{UP}}\ldots U\ldots\ ]]]]]$

Problem:
The PIC_1 does not allow an operation involving Y and an element of WP. Suppose that YP = TP, XP = vP, and WP = VP. The PIC_1 then precludes an operation involving T and NP in VP, which is arguably necessary for cases of long-distance agreement with nominative objects. Solution: The PIC_1 is weakened: A phase is evaluated with respect to the PIC_2 only at the next phase level.

Search Space, PIC_2

(6) Phase Impenetrability Condition$_2$ (PIC_2) (Chomsky (2001, 14)):
The domain of a head X of a phase XP is not accessible to operations at ZP (the next phase); only X and its edge are accessible to such operations.

Consequence:
The search space is enlarged: Operations in YP can now look for a goal in YP, in XP, in WP, or in the residue or head of UP.

(7) Search space under PIC_2:

a. $\overbrace{[_{YP}\ldots Y}^{SCC}\ \underbrace{[_{\underline{XP}}\ldots[_{X'}X\ [_{WP}\ldots W\ [_{\underline{UP}}\ldots U}_{PIC_2}\ldots\ ]]]]]$

b. $\overbrace{[_{\underline{ZP}}\ldots Z}^{SCC}\ \underbrace{[_{YP}\ldots Y\ [_{\underline{XP}}\ldots[_{X'}X}_{PIC_2}\ [_{WP}\ldots W\ [_{\underline{UP}}\ldots U\ldots\ ]]]]]$

Successive-Cyclic Wh-Movement

Wh-movement:
Movement is an agreement relation that is accompanied by an EPP feature on the probe. Checking is deletion under matching. Both PIC_1 and PIC_2 require successive-cyclic wh-movement to proceed via Specv and SpecC. What triggers intermediate movement steps?

(8) Optional EPP Feature Condition (Chomsky (2000, 109; 2001, 34)):
The head X of phase XP may be assigned an EPP- (oder P-) feature (after the phase XP is otherwise complete), but only if that has an effect on outcome.

(9) Derivation of wh-questions:
(I wonder) what John read

a. $[_{\mathrm{VP}}\ \mathrm{read}_3\ \mathrm{what}_1\ ]$
b. $[_{\mathrm{vP}}\ \mathrm{what}_1\ \mathrm{John}_2\ \mathrm{read}_3\ [_{\mathrm{VP}}\ \mathrm{t}_3\ \mathrm{t}_1\ ]]$ (EPP on v)
c. $[_{\mathrm{TP}}\ \mathrm{John}_2\ \mathrm{T}\ [_{\mathrm{vP}}\ \mathrm{what}_1\ \mathrm{t}_2\ \mathrm{read}_3\ [_{\mathrm{VP}}\ \mathrm{t}_3\ \mathrm{t}_1\ ]]]$ (EPP on T)
d. $[_{\mathrm{CP}}\ \mathrm{what}_1\ \mathrm{C}\ [_{\mathrm{TP}}\ \mathrm{John}_2\ \mathrm{T}\ [_{\mathrm{vP}}\ \mathrm{t}'_1\ \mathrm{t}_2\ \mathrm{read}\ [_{\mathrm{VP}}\ \mathrm{t}_3\ \mathrm{t}_1\ ]]]]$ ([wh], EPP on C)

MLC

(10) Minimal Link Condition (Chomsky (2000; 123; 2001, 27)):
If β and γ both match a probe α and β asymmetrically c-commands γ, a syntactic operation cannot involve α and γ.

Problem:
Subject raising to SpecT should be blocked by the MLC if object movement to Specv has occurred: what_1 is closer to T in (9-c) than t_2.

Solution:
Since an equidistance solution (Chomsky (1995)) is to be avoided, the MLC in its strict form must be fulfilled in (9) after all. Idea: After wh-movement, the subject NP is the closest goal for T. Execution of this idea seems to imply giving up the SCC (movement in TP would have to follow movement in CP). Chomsky's solution: The MLC is not evaluated at each step of the derivation; it is only evaluated at the phase level. In (9-d), no overt NP separates the subject trace and T, and the MLC is respected.

Conceptual Problems

Background assumption:
An attractive feature of incremental derivational approaches to syntax is that complexity can be reduced, compared to representational approaches.

1. Lack of look-ahead: At each step of the derivation, subsequent operations and their effects need not (and cannot) be considered.
2. Cyclicity: The SCC prohibits going back to earlier parts of the derivation.
3. Phase Impenetrability: The $\text{PIC}_{1,2}$ significantly reduces the search space. In effect, all syntactic material in the domain that the $\text{PIC}_{1,2}$ renders opaque can (and must) be forgotten for the remainder of the derivation ("spell-out").

Observation:

- PIC_1 and, in particular, PIC_2 could reduce search space even more radically.
- MLC presupposes search space.

Problem no. 1

Three problems with $PIC_{1,2}$ and MLC:
(i) Weak/Strong Representationality (Brody (2001)):
A representational approach can be strictly non-derivational. However, a derivational theory must always be representational to some extent. It is weakly representational if "derivational stages are transparent (i.e., representations), in the sense that material already assembled can be accessed." It is strongly representational if it "is weakly representational and there are constraints on the representations." On this view, Chomsky's (2000; 2001) approach is strongly derivational (see phase evaluation of PIC_2 and MLC). Ideally, though, a strictly derivational theory should not even be weakly representational.

Problem no. 2

(ii) PIC/MLC Redundancy:
Chomsky (2001, 47, fn. 52) notes: "The effect on the MLC is limited under the PIC, which bars 'deep search' by the probe." The MLC can only become relevant in the relatively small portions of structure permitted by PIC_1 and PIC_2; it thus loses much of its original empirical coverage. In line with problem (i), this could be taken to suggest that strictly derivational approaches should dispense with the MLC in toto since this constraint presupposes an articulated representation (the search space for the probe). Arguably, in a derivational approach, minimality effects should not be covered by a constraint that accesses a significant amount of syntactic structure, i.e., a representation (MLC); rather, they should emerge as epiphenomena of constraints that reduce the search space (PIC).

Problem no. 3

(iii) SCC/$PIC_{1,2}$ Asymmetry:
The SCC and the $PIC_{1,2}$ have complementary tasks and look like two sides af the same coin. Therefore, it is possibly suspicious that the two constraints talk about domains of such a different size: From the point of view of symmetry, the local domain of the SCC should be the phase (not the phrase); or the local domain of the PIC should be the phrase (not the phase).

Problem no. 3

(iii) SCC/$PIC_{1,2}$ Asymmetry:
The SCC and the $PIC_{1,2}$ have complementary tasks and look like two sides af the same coin. Therefore, it is possibly suspicious that the two constraints talk about domains of such a different size: From the point of view of symmetry, the local domain of the SCC should be the phase (not the phrase); or the local domain of the PIC should be the phrase (not the phase).

Goal:
A derivational approach that evades problems (i), (ii) and (iii) by having the following properties:
(i) The material that can be accessed at any given step of the derivation is a small bundle of categories that can hardly be called a representation anymore. Hence, the approach is not even weakly representational.
(ii) The MLC is dispensed with in favour of a strengthened version of the PIC.
(iii) The new version of the PIC has the same kind of local domain as the SCC: the phrase.

Phrase Impenetrability

Movement:
Movement is triggered by [*F*] features on the probe; these features can directly trigger movement of [F]-bearing items (no recourse to generalized EPP features). (See Stechow & Sternefeld (1981), Sternefeld (2000).) The condition that triggers movement is FC; the condition that requires movement to be feature-driven is LR.

(11) Feature Condition (FC):
A feature [*F*] must be checked by movement.

(12) Last Resort (LR):
Movement must result in deletion of [*F*].

PIC_3

The SCC is as before, but PIC_3 now is restricted to phrases. This denies a special role of CP and vP for the purposes of movement theory, contra Chomsky (2000; 2001), Fox (1999), Nissenbaum (1999), Barbiers (2002), and others. (But it is of course compatible with the all the evidence suggesting that SpecC and Specv are used by successive-cyclic movement.)

(13) Strict Cycle Condition (SCC):
Within the current XP α, a syntactic operation may not target a position that is included within another XP β that is dominated by α.

(14) Phrase Impenetrability Condition$_3$ (PIC_3):
The domain of a head X of a phrase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

Search Space, PIC_3

(15) Search space under PIC_3:

a. $\overbrace{\overbrace{[_{\underline{YP}}\ldots Y}^{SCC}\ [_{\underline{XP}}\ldots[_{X'}X}^{PIC_3}\ [_{\underline{WP}}\ldots W\ [_{\underline{UP}}\ldots U\ldots\ ]]]]]$

b. $\overbrace{\overbrace{[_{\underline{ZP}}\ldots Z}^{SCC}\ [_{\underline{YP}}\ldots Y}^{PIC_3}\ [_{\underline{XP}}\ldots\ [_{X'}X\ [_{\underline{WP}}\ldots W\ [_{\underline{UP}}\ldots U\ldots\ ]]]]]]$

Phrase Balance

Finally, a trigger is needed for successive-cyclic movement to phrase edges. In the present system, the role of the Optional EPP Feature Condition is played by Phrase Balance, a strengthened version of Phase Balance in Heck & Müller (2000).

(16) Phrase Balance (PB):
Every XP has to be balanced: For every feature [*F*] in the numeration there must be a distinct potentially available feature [F].

(17) Potential availabiliy:
A feature [F] is potentially available at the XP level if (i) or (ii) holds:
a. [F] is on X or on an edge element of X.
b. [F] is part of the workspace of the derivation.
(The workspace of a derivation D comprises the numeration and material in trees that have been created earlier and have not yet been used in D.)

Last Resort

Note:
Phrase Balance can trigger movement without feature checking, but Last Resort prohibits such movement. Assumption (Heck & Müller (2000)): LR is minimally violable if this is the only way to fulfill the inviolable conditions FC, SCC, PIC_3, and PB.

(This can be encoded in an optimality-theoretic way by a ranking FC, SCC, PIC_3, PB $\gg$ C $\gg$ LR; an additional condition C is needed to ensure that the higher-ranked constraints are in fact never violable in a well-formed output. C punishes the candidate that derives absolute ungrammaticality/ineffability, e.g., an empty output (null parse), or an unfaithful output that removes an offending property and leads to neutralization of different input specifications.)

Alternative: The Inclusiveness Condition (Chomsky (2001)) is violable (Last Resort is not): Features triggering intermediate movement steps can be inserted if this is forced by Phrase Balance.

Locality of Feature Transfer

Consequence:
Wh-movement must proceed via every XP on its way to its ultimate target position (the $C_{[*wh*]}$ node that attracts it). The resulting system is thus close to analyses in Sportiche (1989; 1998) Takahashi (1994), and Agbayani (1998), among others. It also bears a certain similarity to GPSG analyses that rely on SLASH feature percolation (Gazdar (1981), Gazdar, Klein, Pullum & Sag (1985)), and to Koster's (2000) analysis based on feature percolation in gap phrases.

(18) Derivation of wh-questions:
(I wonder) what John read

a. $[_{\mathrm{VP}}$ what$_3$ read$_3$ t$_1$ $]$

→ workspace: $\{C_{[*wh*]}, \text{John}, T_{[*D*]}, v\}$

b. $[_{\mathrm{vP}}$ what$_1$ John$_2$ v+read$_3$ $[_{\mathrm{VP}}$ t$'_1$ t$_3$ ~~t$_1$]]~~

→ workspace: $\{C_{[*wh*]}, T_{[*D*]}\}$

c. $[_{\mathrm{TP}}$ what$_1$ John$_2$ T $[_{\mathrm{vP}}$ t$''_1$ t$_2$ v+read$_3$ ~~[$_{\mathrm{VP}}$ t$'_1$ t$_3$ t$_1$]]]~~

→ workspace: $\{C_{[*wh*]}\}$

d. $[_{\mathrm{CP}}$ what$_1$ C $[_{\mathrm{TP}}$ t$'''_1$ John$_2$ T ~~[$_{\mathrm{vP}}$ t$''_1$ t$_2$ v+read$_3$ [$_{\mathrm{VP}}$ t$'_1$ t$_3$ t$_1$]]]]~~

→ workspace: $\{-\}$

Note:
Material that is crossed out is not accessible anymore for further operations in the derivation.

Superiority Effects in English

(19) Subject and object:

a. (I wonder) who_1 bought $what_2$

b. *(I wonder) $what_2$ who_1 bought t_2

(20) Object and object:

a. Who_1 did you persuade t_1 [$_{CP}$ to read $what_2$] ?

b. *$What_2$ did you persuade who_1 [$_{CP}$ to read t_2] ?

Analysis:
(i) Movement of NP_2 to SpecV (Specv, SpecT, ...) cannot be triggered by the Feature Condition ([*F*]): English does not have object shift or scrambling.
(ii) Movement of NP_2 to SpecV (Specv, SpecT, ...) cannot be triggered by Phrase Balance: VP is balanced because there is another wh-phrase in the workspace, viz., NP_1; vP and TP in (19) are balanced because NP_1 occupies the respective specifiers.

Superiority Derivation

(21) Derivation of (19):

a. $[_{\mathrm{VP}}$ bought$_3$ what$_2$ $]$

→ workspace: {$C_{[*wh*]}$, who$_1$, $T_{[*D*]}$, v}

b. $[_{\mathrm{vP}}$ who$_1$ v+bought$_3$ $[_{\mathrm{VP}}$ t$_3$ ~~what$_2$]]~~

→ workspace: {$C_{[*wh*]}$, $T_{[*D*]}$,}

c. $[_{\mathrm{TP}}$ who$_1$ T $[_{\mathrm{vP}}$ t$_1$ v+bought$_3$ ~~$[_{\mathrm{VP}}$ t$_3$ what$_2$]]]~~

→ workspace: {$C_{[*wh*]}$}

d. $[_{\mathrm{CP}}$ who$_1$ C $[_{\mathrm{TP}}$ t$_1''$ T ~~$[_{\mathrm{vP}}$ t$_1$ v+bought$_3$ $[_{\mathrm{VP}}$ t$_3$ what$_2$]]]]~~

→ workspace: {–}

Double Object Constructions: NP-PP

Observation 1 (Chomsky (1973, 246), Fiengo (1980, 123)):
If there are two wh-objects in double object constructions with a prepositional object, either object can move; but preposition stranding becomes impossible.

(22) Double object constructions with a prepositional object:

a. What$_1$ did you give t$_1$ to whom$_2$?
b. To whom$_3$ did you give what$_1$ t$_3$?
c. *Who$_2$ did you give what$_1$ [$_{\mathrm{PP}_3}$ to t$_2$] ?

Double Object Constructions: NP-NP

Observation 2 (Barss & Lasnik (1986, 349)):
If there are two wh-objects in double object constructions with dative shift, only the shifted object can (marginally) move. (The marginality is due to a general weak ban on A-bar movement of dative-shifted objects in English and thus independent of superiority.)

(23) Double object constructions with dative shift:
a. (?)Who$_2$ did you give t$_2$ what$_1$?
b. *What$_1$ did you give who$_2$ t$_1$?

Assumptions

Assumption 1 (modification of Larson (1988)):
The direct (THEME) object is in a complement position of V throughout; the indirect (GOAL) object is in SpecV if it has undergone dative shift, and in a right-peripheral V′ sister position if it is prepositional.

(24) Base generation of double object constructions:

a. $[_{\mathrm{VP}}\ [_{\mathrm{V'}}\ [_{\mathrm{V'}}\ \mathrm{V}\ \mathrm{NP}_1\ ]\ [_{\mathrm{PP}_3}\ \mathrm{P}\ \mathrm{NP}_2\ ]]]$

b. $[_{\mathrm{VP}}\ \mathrm{NP}_2\ [_{\mathrm{V'}}\ \mathrm{V}\ \mathrm{NP}_1\ ]]$

Assumption 2:
Pied piping involves optional feature percolation, which can be viewed as actual displacement of a feature (this assumption is probably wrong, but it may suffice for present purposes; see Heck (2004)).

Analysis

Analysis of (22):
(i) Suppose that [wh] percolation has taken place, and PP bears [wh]. The two objects are merged in VP-internal non-edge positions. Phase Balance forces movement of one wh-phrase to SpecV. It does not matter which one, but whichever wh-phrase is moved first forces the other wh-phrase to stay in situ, to avoid an unforced violation of Last Resort. The wh-phrase in SpecV is then passed on through further cycles of the derivation, until CP is reached and [*wh*] on C is checked.
(ii) Suppose that [wh] percolation has not taken place. Then, PP cannot move. However, NP_2 in PP cannot move either: To leave PP, it must move to SpecP. This operation is not available because there is no [*F*] that might trigger it, and because PB is independently satisfied (with another wh-phrase in the workspace).

Analysis of (23):
In dative shift constructions, NP_2 is in SpecV for independent reasons (either because it is base-generated there, or because there is an [*F*] that triggers dative shift to that position). Hence, PB can be fulfilled without a LR violation, and any such violation incurred by movement of the lower wh-phrase is fatal.

A Prediction

If two wh-phrases are further embedded in objects, preposition stranding is predicted to be blocked throughout by PB. (Note: Examples like (25-d) are classified as ill formed by Jackendoff (1990, 433), and as well formed by Fiengo (1980, 124).)

(25) Two embedded wh-phrases:

a. *Who$_2$ did you give [$_{\mathrm{NP}}$ pictures of t$_2$] [$_{\mathrm{PP}}$ to whom$_1$] ?
b. *Who$_1$ did you give [$_{\mathrm{NP}}$ pictures of whom$_2$] [$_{\mathrm{PP}}$ to t$_1$] ?
c. *Who$_2$ did you talk [$_{\mathrm{PP}}$ to t$_2$] [$_{\mathrm{PP}}$ about whom$_1$] ?
d. *Who$_1$ did you talk [$_{\mathrm{PP}}$ to whom$_2$] [$_{\mathrm{PP}}$ about t$_1$] ?

(25-ac) (classified as acceptable by Jackendoff) should involve an additional violation of the Clause Nonfinal Incomplete Constituent Constraint; see Kuno (1973), Lasnik & Saito (1992, 91).

(26) Movement from clause non-final constituents:

a. Who$_2$ did you give [$_{\mathrm{NP}}$ pictures of Mary] [$_{\mathrm{PP}}$ to t$_2$] ?
b.?*Who$_1$ did you give [$_{\mathrm{NP}}$ pictures of t$_1$] [$_{\mathrm{PP}}$ to John] ?

Lack of Superiority Effects in German

Observation (Haider (1983; 1993; 2000), Grewendorf (1988), Bayer (1990)):
German does not exhibit superiority effects with clause-mates.

(27) Lack of superiority effects with clause-mates:

a. (Ich weiß nicht) wer_1 C t_1 was_2 gesagt hat
I know not who_{nom} $what_{acc}$ said has

b. (Ich weiß nicht) was_2 C wer_1 t_2 gesagt hat
I know not $what_{acc}$ who_{nom} said has

Lack of Superiority Effects in German: Infinitives

Observation (Fanselow (1991), Kim & Sternefeld (1997), Haider (2000)):
German does not exhibit superiority effects with control infinitives.

(28) Lack of superiority effects in control infinitives:

a. (Ich weiß nicht) wen_1 er t_1 überzeugt hat [was_2 zu kaufen]
I know not whom_{acc} he convinced has what_{acc} to buy

b. (Ich weiß nicht) was_2 er wen_1 überzeugt hat [t_2 zu kaufen]
I know not what_{acc} he whom_{acc} convinced has to buy

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).
4. Superiority violations are only apparent in German; D-linking is involved (Fanselow (1991), Wiltschko (1997), Grohmann (1998)).

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).
4. Superiority violations are only apparent in German; D-linking is involved (Fanselow (1991), Wiltschko (1997), Grohmann (1998)).
5. Superiority results from a feature-based version of relativized minimality: a wh-element in situ must not c-command the trace of a wh-element whose overt licensing features are non-distinct (Haider (2000), based on Fanselow (1991)). German can avoid superiority effects because Case is a licensing feature and overt Case forms are often distinct; this is not the case in English.

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).
4. Superiority violations are only apparent in German; D-linking is involved (Fanselow (1991), Wiltschko (1997), Grohmann (1998)).
5. Superiority results from a feature-based version of relativized minimality: a wh-element in situ must not c-command the trace of a wh-element whose overt licensing features are non-distinct (Haider (2000), based on Fanselow (1991)). German can avoid superiority effects because Case is a licensing feature and overt Case forms are often distinct; this is not the case in English.
6. Superiority results from the MLC. German has variable base generation of arguments (Kim & Sternefeld (1997), based on Fanselow (2001)).

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).
4. Superiority violations are only apparent in German; D-linking is involved (Fanselow (1991), Wiltschko (1997), Grohmann (1998)).
5. Superiority results from a feature-based version of relativized minimality: a wh-element in situ must not c-command the trace of a wh-element whose overt licensing features are non-distinct (Haider (2000), based on Fanselow (1991)). German can avoid superiority effects because Case is a licensing feature and overt Case forms are often distinct; this is not the case in English.
6. Superiority results from the MLC. German has variable base generation of arguments (Kim & Sternefeld (1997), based on Fanselow (2001)).
7. Superiority results from the MLC. The English C attracts all wh-phrases, but only the first one that is attracted by C is PF-realized (Pesetsky (2000)). German has a different C that requires only one wh-phrase in SpecC and attracts the wh-features of the others. If feature attraction precedes phrase attraction, superiority violations occur.

Previous Analyses

1. Superiority results from a lack of government (ECP). All arguments are governed in German (Haider (1983; 1993)).
2. Superiority results from a lack of government (ECP). C in V/2 languages can govern the subject position (Noonan (1988)).
3. Superiority results from an IP barrier that restricts LF movement of the in-situ wh-phrase (ECP). This barrier does not exist in German (Müller (1995)).
4. Superiority violations are only apparent in German; D-linking is involved (Fanselow (1991), Wiltschko (1997), Grohmann (1998)).
5. Superiority results from a feature-based version of relativized minimality: a wh-element in situ must not c-command the trace of a wh-element whose overt licensing features are non-distinct (Haider (2000), based on Fanselow (1991)). German can avoid superiority effects because Case is a licensing feature and overt Case forms are often distinct; this is not the case in English.
6. Superiority results from the MLC. German has variable base generation of arguments (Kim & Sternefeld (1997), based on Fanselow (2001)).
7. Superiority results from the MLC. The English C attracts all wh-phrases, but only the first one that is attracted by C is PF-realized (Pesetsky (2000)). German has a different C that requires only one wh-phrase in SpecC and attracts the wh-features of the others. If feature attraction precedes phrase attraction, superiority violations occur.
8. Superiority results from the MLC. German has scrambling, which may independently move a lower wh-phrase to a higher position, by wh-scrambling (Fanselow (1996; 1997), Grohmann (1997)).

Analysis

Suppose that the last approach is the correct one: A lower wh-phrase cannot move across a wh-phrase that is generated in a higher XP by wh-movement, given PB; but it can do so by scrambling. Scrambling is triggered by [*Σ*] on a probe, and [Σ] on the scrambled item (Müller (1998), Sauerland (1999), Grewendorf & Sabel (1999)).

(29) Derivation of (27-b):

a. $[_{\mathrm{VP}}$ was$_{2,\Sigma}$ $[_{\mathrm{V'}}$ t$_2$ gesagt]]

$\rightarrow$ workspace: {C$_{[*wh*]}$, wer$_1$, T, $[_{\mathrm{v}}$ hat $]_{[*\Sigma*]}$}

b. $[_{\mathrm{vP}}$ was$_{2,\Sigma}$ wer$_1$ $[_{\mathrm{VP}}$ t$'_2$ ~~$[_{\mathrm{V'}}$ t$_2$~~ gesagt]] $[_{\mathrm{v}}$ hat]]

$\rightarrow$ workspace: {C$_{[*wh*]}$, T}

c. $[_{\mathrm{TP}}$ was$_{2,\Sigma}$ $[_{\mathrm{vP}}$ t$''_2$ wer$_1$ ~~$[_{\mathrm{VP}}$ t$'_2$ $[_{\mathrm{V'}}$ t$_2$ gesagt]]~~ $[_{\mathrm{v}}$ hat]] T]

$\rightarrow$ workspace: {C$_{[*wh*]}$}

d. $[_{\mathrm{CP}}$ was$_{2,\Sigma}$ C $[_{\mathrm{TP}}$ t$'''_2$ ~~$[_{\mathrm{vP}}$ t$''_2$ wer$_1$ $[_{\mathrm{VP}}$ t$'_2$ $[_{\mathrm{V'}}$ t$_2$ gesagt]] $[_{\mathrm{v}}$ hat]]~~ T]

$\rightarrow$ workspace: {–}

Superiority Effects with Long-Distance Movement in German

Observation (Frey (1993), Büring & Hartmann (1994), Fanselow (1997), Heck & Müller (2000), Pesetsky (2000)):
German does exhibit superiority effects with long-distance movement.

(30) Superiority effects with long-distance movement:

a. Wer$_1$ hat t_1 geglaubt [$_{\mathrm{CP}}$ dass der Fritz wen$_2$ mag] ?
who$_{nom}$ has believed that the Fritz whom$_{acc}$ likes

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

Analysis:
NP_2's [wh] feature in (30) does not permit movement at all (PB being fulfilled by the presence of NP_1 in the workspace). Scrambling cannot leave a finite CP in German (only v and V tolerate [Σ] at its edge). Hence, a potential [Σ] feature cannot trigger movement beyond vP.

Superiority Effects with Subject Raising in German

Assumption (Müller (2001)):
The EPP feature of T is optional in German (Diesing (1992)). There is one context where it is clear that subject raising to SpecT must have occurred: Unstressed pronouns must be at the phonological edge of vP. Only a subject NP can optionally precede these pronouns. In that case, the subject NP must have undergone movement to SpecT.

(31) Superiority effects with subject raising 1:

a. Wem$_2$ hat [$_{\mathrm{vP}}$ es t'_2 wer$_1$ t_2 gegeben] ?
whom$_{dat}$ has it$_{acc}$ who$_{nom}$ given

b.?*Wem$_2$ hat wer$_1$ [$_{\mathrm{vP}}$ es t'_2 t_1 t_2 gegeben] ?
whom$_{dat}$ has who$_{nom}$ it$_{acc}$ given

(32) Superiority effects with subject raising 2:

a. Was$_2$ hat [$_{\mathrm{vP}}$ ihm t'_2 wer$_1$ t_2 gegeben] ?
what$_{acc}$ has him$_{dat}$ who$_{nom}$ given

b.?*Was$_2$ hat wer$_1$ [$_{\mathrm{vP}}$ ihm t'_2 t_1 t_2 gegeben] ?
what$_{acc}$ has who$_{nom}$ him$_{dat}$ given

Analysis

If a subject $NP_{[wh]}$ and an object $NP_{[wh]}$ are both in Specv, and T has an optional EPP feature, TP is balanced by feature-driven subject raising, and movement of the object incurs a fatal LR violation.

(33) Examples with non-wh-subjects:

a. Wem$_2$ hat t''_2 der Fritz$_1$ [$_{vP}$ es t'_2 t_1 t_2 gegeben] ?
whom$_{dat}$ has the Fritz$_{nom}$ it$_{acc}$ given

b. Was$_2$ hat t''_2 der Fritz$_1$ [$_{vP}$ ihm t'_2 t_1 t_2 gegeben] ?
what$_{acc}$ has the Fritz$_{nom}$ him$_{dat}$ given

(34) Examples with non-pronouns:

a. Wem$_2$ hat [$_{vP}$ wer$_1$ das Buch t'_2 t_1 t_2 gegeben] ?
whom$_{dat}$ has who$_{nom}$ the book$_{acc}$ given

b. Was$_2$ hat [$_{vP}$ wer$_1$ dem Fritz t'_2 t_1 t_2 gegeben] ?
what$_{acc}$ has who$_{nom}$ the Fritz$_{dat}$ given

Superiority Effects with Scrambling from Wh-XP in German

Observation (Fanselow (1997)):
A wh-phrase PP_1 can be scrambled from a wh-phrase NP_2. However, once it has done this, the MLC forces movement of PP_1 to SpecC – if NP_2 also undergoes scrambling so as to be in a position for wh-movement to SpecC, a violation of Unambiguous Domination (UD; Müller (1998)) occurs; according to UD, an α-trace must not be α-dominated.

(35) Superiority effects with scrambling from wh-XP:

a. (Ich weiß nicht) $[_{PP_1}$ über wen] er $[_{NP_2}$ wieviele Bücher t_1] lesen will
I know not about whom he how many books read wants

b.?*(Ich weiß nicht) $[_{NP_2}$ wieviele Bücher t_1] er $[_{PP_1}$ über wen] t_2 lesen
I know not how many books he about whom read
will
wants

c. (Ich weiß nicht) $[_{NP_2}$ wieviele Bücher t_1] er $[_{PP_1}$ über die Liebe] t_2
I know not how many books he about love
lesen will
read wants

Analysis

Movement of PP_1 from NP_2 must take place while NP_2 is still in situ, a sister of V (otherwise, a CED effect would result); PP_1 is scrambled to SpecV (triggered by [*Σ*], either on V directly, or indirectly, by PB). In situ, NP_2 is not visible for further operations, because of the PIC_3. However, if NP_2 moves to SpecV (triggered by [*Σ*] on V or v, and by [Σ] on NP_2), UD is violated.

Superiority-Like Effects with Remnant Movement in German

Observation (Koizumi (1995), Takano (1996), Kitahara (1997), Müller (1998), Sauerland (1999)):
UD effects can be derived from the MLC: In a configuration ... [$_\beta$ ... γ ...], where β and γ both qualify as a goal for a β-external probe α, the MLC forces movement of the item that is closer to α: β.

(36) Unambiguous Domination effects:

a. *dass [$_{\mathrm{vP}}$ [$_{\mathrm{VP}_{2,[\Sigma]}}$ t_1 zu lesen] [$_{\mathrm{NP}_{1,[\Sigma]}}$ das Buch] keiner t_2 versucht hat]
that to read the book$_{acc}$ no-one$_{nom}$ tried has

b. dass [$_{\mathrm{vP}}$ [$_{\mathrm{VP}_{2,[\Sigma]}}$ das Buch$_1$ zu lesen] keiner t_2 versucht hat]
that the book$_{acc}$ to read no-one$_{nom}$ tried has

Analysis

The illformedness of (36-a) does not follow from the PIC_3. Suppose that there are two [*Σ*] features, one for NP_1, one for VP_2. Then, there should be a well-formed derivation for (36-a), with NP_1 undergoing PB-driven movement to SpecV first, followed by feature-driven movement of NP_1 to Specv, and then of VP_2 to Specv – both movements are compatible with PIC_3. However, UD can be derived by a version of the A-over-A condition.

(37) A-Over-A Condition (AOA):
If [*F*] can be checked either with a head, or with an edge element, it must be checked with the head.

Note:
This forces VP_2 movement to apply first; subsequent NP_1 lowering then violates (at least) the CED.

Long-Distance Intervention without C-Command

Observation (Heck & Müller (2000)):
Non-c-commanding wh-phrases in a matrix clause block long-distance wh-movement in German.

(38) Long-distance wh-movement across a wh-item in an adjunct clause:

a. Wen$_1$ hat Fritz [$_{\mathrm{CP}}$ nachdem er was$_2$ gemacht hat] t_1 getroffen ?
whom has Fritz after he what done has met

b. *Wen$_1$ hat Fritz [$_{\mathrm{CP}}$ nachdem er was$_2$ gemacht hat] gesagt [$_{\mathrm{CP}}$ dass
whom has Fritz after he what done has said that
Maria t_1 liebt] ?
Maria loves

c. *Was$_2$ hat Fritz [$_{\mathrm{CP}}$ nachdem er t_2 gemacht hat] gesagt [$_{\mathrm{CP}}$ dass
what has Fritz after he done has said that
Maria wen$_1$ liebt] ?
Maria whom loves

Further Cases of Long-Distance Intervention without C-Command

(39) Long-distance wh-movement across a wh-item in a relative clause:

a. Wen$_1$ hat Fritz [$_{\mathrm{NP}}$ einem Mann [$_{\mathrm{CP}}$ der was$_2$ kennt]] t$_1$
whom$_{acc}$ has Fritz a man$_{dat}$ that what knows
vorgestellt ?
introduced

b. *Wen$_1$ hat Fritz [$_{\mathrm{NP}}$ einem Mann [$_{\mathrm{CP}}$ der was$_2$ kennt]] gesagt [$_{\mathrm{CP}}$
whom$_{acc}$ has Fritz a man$_{dat}$ that what knows said
dass er t$_1$ einladen soll] ?
that he invite should

(40) Long-distance wh-movement across a wh-item in an NP:

a. Wen$_1$ hat Fritz [$_{\mathrm{NP}}$ einem Freund von wem$_2$] t$_1$ vorgestellt ?
whom$_{acc}$ has Fritz a friend$_{dat}$ of whom introduced

b.?*Wen$_1$ hat Fritz [$_{\mathrm{NP}}$ einem Freund von wem$_2$] gesagt [$_{\mathrm{CP}}$ dass
whom$_{acc}$ has Fritz a friend$_{dat}$ of whom said that
Maria t$_1$ liebt] ?
Maria loves

Analysis

The (a)-examples are well formed because NP_1 can always reach a phrase edge by scrambling. The (b)-examples are ill formed because, given another wh-item in the workspace, PB does not force movement of NP_1 to the edge, and LR prohibits such movement. Independent scrambling must stop in the embedded vP domain. (The (c)-example in (38) is ill formed because of thse CED).

Problem

(41) A potential problem:

a. [$_{\mathrm{NP}}$ Die Frage [$_{\mathrm{CP}_9}$ wer$_1$ C t$_1$ was$_2$ mitbringt]]] ist relevant für die
the question who what brings is relevant to the
Frage [$_{\mathrm{CP}_7}$ wie$_3$ Fritz denkt [$_{\mathrm{CP}_5}$ t$'_3$ dass die Party t$_3$ wird]
question how Fritz thinks that the party will be

b. Who$_1$ C t$_1$ told whom$_2$ [$_{\mathrm{CP}}$ what$_3$ C John likes t$_3$] ?

Solution:
Wh-features are accompanied by scope indices on items in the numeration.
Hence, [wh]$_i$ on a wh-phrase can never be potentially available for [*wh*]$_j$ on a C in the workspace, due to feature mismatch.

Intervention without C-Command in English

Prediction:
Since English does not have scrambling, we expect clause-bound intervention effects with non-commanding wh-phrases. At first sight, this seems to contradict the standard view that argument wh-in situ in English does not obey any island constraints (Chomsky (1981), Huang (1982; 1995), Lasnik & Saito (1992), Hornstein (1995)). However, most of the pertinent examples in the literature do not involve intervention without c-command. Where such intervention does occur, acceptability seems to be significantly reduced.

(42) Wh-in situ in an object or adjunct does not block subject wh-movement:

a. Who$_1$ t$_1$ saw [$_{\mathrm{NP}}$ the man that bought what$_2$] ?
b. Who$_1$ t$_1$ likes [$_{\mathrm{NP}}$ books that criticize who$_2$] ?
c. Who$_1$ t$_1$ bought [$_{\mathrm{NP}}$ the books on which table$_2$] ?
d. Who$_1$ t$_1$ met [$_{\mathrm{NP}}$ friends of whom$_2$] ?
e. I wonder who$_1$ t$_1$ heard [$_{\mathrm{NP}}$ the claim that John had seen what$_2$]
f. I wonder who$_1$ t$_1$ heard [$_{\mathrm{NP}}$ John's stories about what$_2$]
g. Who$_1$ t$_1$ left [$_{\mathrm{PP}}$ despite which warning$_2$] ?

More Examples

(43) Wh-in situ in a subject does not block matrix subject wh-movement:
Who$_1$ t$_1$ thinks that [$_{\mathrm{NP}}$ pictures of who$_2$] are on sale ?

(44) Wh-in situ in a subject may block object wh-movement:
a.??Who$_2$ did [$_{\mathrm{NP}}$ the man that bought what$_1$] see t$_2$?
b.?*Who$_2$ did [$_{\mathrm{NP}}$ books that criticize who$_1$] impress t$_2$?
c.?*What$_2$ did [$_{\mathrm{NP}}$ the books on which table$_1$] cost t$_2$?
d.?*Who$_2$ did [$_{\mathrm{NP}}$ friends of whom$_1$] meet t$_2$?
e. *Who$_2$ did [$_{\mathrm{NP}}$ friends of whom$_1$] say that we should invite t$_2$?

D-Linking

Problem:
Why does an intervening D-linked wh-phrase not induce a superiority violation in English?

(45) D-linking effects:

a. *I know [$_{\mathrm{NP}_2}$ which books] who$_1$ read t$_2$

b. I know what$_2$ [$_{\mathrm{NP}_1}$ which people] read t$_2$

Suggestion:
D-linked wh-phrases can optionally lack a (proper) [wh]-feature; if they lack this feature, PB can only be fulfilled by movement of the other wh-phrase.

Conclusion

Results:

- There are independent reasons for strengthening the standard PIC in a derivational grammar, from a condition on phases ($PIC_{1,2}$) to a condition on phrases (PIC_3).
- It follows from this move that the PIC_3 accounts for typical MLC effects in English. The MLC can therefore be dispensed with (except for a residue, AOA).
- Given that German has scrambling of wh-phrases, superiority effects are predicted to be absent, except for those circumstances where scrambling is independently excluded.
- Unlike the MLC, the system based on PB and the PIC_3 predicts superiority-like intervention effects without c-command.

Outlook

As it stands, the PIC_3 has important consequences for many other phenomena, especially if we pursue the strongest possible hypothesis in a derivational approach:

(46) A strong hypothesis:
Once rendered inaccessible by the PIC_3, syntactic structure does not become accessible again when the syntactic derivation terminates ("at LF"). Hence, there can be no constraints on representations ("bare output conditions").

Hypothesis (46) implies that there is no reason left to assume the existence of traces (neither as t, nor as a copy), which presupposes a derivational approach to semantic interpretation (Sternefeld (1996)). It also raises interesting problems for binding of anaphors (at least those cases that are not strictly local and can be covered by Reinhart & Reuland's (1993) reflexivity constraints) and pronouns, control, long-distance agreement, etc. It seems that apparent non-local relations must be accounted for by successive-cyclic local [F] feature movement from head to head (required by constraints of the PB type or motivated by independent features; see Pesetsky (2000) on the viability of feature movement). [F] must encode the relevant properties of the in-situ element; e.g.: anaphor, PRO. For binding, this strategy would be a natural extension of proposals like Chomsky's (1986) LF movement of anaphors, and a version of it is in fact pursued by Reuland (2001) in his account of A-chain condition effects (also cf. Fischer (2004)). For (obligatory) control, the strategy would amount to a decomposition of Landau's (2000) Agree relation into small steps of feature movement (or, indeed, a version of Hornstein's (2001) A-movement approach).