Phrase Impenetrability and Wh-Intervention

Gereon Müller
IDS Mannheim

Final version, July, 2003

1. Introduction and Overview

This paper takes as its starting point the observation that something is wrong with a Minimal Link Condition (MLC) in a derivational grammar. Brody (2001) argues that a derivational approach to syntax should minimize search space, its representational residue; thus, the amount of structure that is visible and accessible to syntactic operations at any given step should be as small as possible. Given this tenet, it follows that constraints that minimize search space should be strengthened in a derivational grammar; in contrast, constraints that presuppose search space should be abandoned. A constraint that minimizes search space is the Phase Impenetrability Condition (PIC; see Chomsky (2000), Chomsky (2001b)); in contrast, the Minimal Link Condition (MLC; see Fanselow (1991), Ferguson and Groat (1994), Chomsky (1995), Chomsky (2000), Chomsky (2001b), among many others) is a constraint that presupposes search space. In line with this, I will argue that *wh*-intervention effects usually attributed to the MLC (more specifically, superiority effects as they arise with *wh*-movement in German and English), as well as certain superiority-like *wh*-intervention effects that the MLC has nothing to say about, can be derived from a strengthened version of the PIC – one that holds for *phrases* rather than *phases*.

I will proceed as follows. Section 2 provides some background assumptions, introduces standard versions of the PIC and the MLC, and lays out conceptual arguments against the MLC, and for a version of the PIC that is based on a more local domain. Section 3 develops an approach to syntactic movement operations that dispenses with the MLC and relies on a more restrictive version of the PIC. The resulting approach is then shown to account for standard superiority effects in English, the absence of standard superiority effects in German, as well as a priori unexpected instances of superiority and superiority-like effects in both languages. Finally, section 4 draws a conclusion.

2. Phase Impenetrability

2.1. The Standard Approach

Throughout this paper, I presuppose an incremental-derivational approach to movement as developed in Chomsky (2000) and Chomsky (2001b). In this kind of ap-

For helpful comments and discussion, I would like to thank Chris Collins, Gisbert Fanselow, Silke Fischer, Hubert Haider, Fabian Heck, Michal Starke, two careful reviewers, and the participants of the workshops on Minimal Link Effects at Universität Potsdam (March 2002), and on Tools in Linguistic Theory (TiLT) at Universiteit Utrecht (April 2002). This paper was inspired by a talk by Michael Brody at Universität Tübingen (February, 2001).

proach, two constraints prove particularly relevant; they reduce derivational search space by imposing strong restrictions on what counts as an active, accessible part of the derivation. First, the Strict Cycle Condition (SCC), arguably indispensable in any derivational approach to syntax, restricts possible positions for the probe (i.e., features of a head that drive movement operations and create the target for movement); second, the PIC significantly reduces the positions in which the derivation can look for a goal (i.e., the item that is to be moved). For present purposes, the SCC can be formulated in a classical way, as in (1) (see Chomsky (1973), Perlmutter and Soames (1979)).[1]

(1) *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.

A first version of the PIC is given in (2) (see Chomsky (2000, 108), Chomsky (2001b, 13)).[2]

(2) *Phase Impenetrability Condition*$_1$ (PIC$_1$):
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.

The notions of (i) "edge" and (ii) "phase" need to be clarified. (i) The edge of a head X is the left-peripheral minimal residue outside of X′; it includes specifiers of X, of which there can in principle be arbitrarily many (irrelevantly for the purposes of this paper, it also comprises adjuncts to XP); see Chomsky (2001b, 13). (ii) The propositional categories CP and vP are phases; other XPs (except perhaps for DP) are not. With this in mind, let us look abstractly at syntactic derivations, and determine the search space available to the derivation at any given point. Thus, suppose that ZP, XP, and UP are phases in (3). Then, in (3-a), an operation can have a probe only in YP (because of the SCC), and an operation can look for a goal only in YP or in the residue or head of XP (because of the PIC$_1$). In the subsequent step (3-b), the probe must be in ZP, and the search space for a goal grows as indicated.

(3) *Search space under PIC*$_1$:

a. $\overbrace{[_{\mathrm{YP}}\ldots\mathrm{Y}\,[_{\underline{\mathrm{XP}}}\ldots[_{\mathrm{X'}}\mathrm{X}}^{\mathrm{PIC}_1}\ [_{\mathrm{WP}}\ldots\mathrm{W}\ [_{\underline{\mathrm{UP}}}\ldots\mathrm{U}\ldots\ ]]]]]$ (SCC)

b. $\overbrace{[_{\underline{\mathrm{ZP}}}\ldots\mathrm{Z}\,[_{\mathrm{YP}}\ldots\mathrm{Y}\ [_{\underline{\mathrm{XP}}}\ldots[_{\mathrm{X'}}\mathrm{X}}^{\mathrm{PIC}_1}\ [_{\mathrm{WP}}\ldots\mathrm{W}\ [_{\underline{\mathrm{UP}}}\ldots\mathrm{U}\ldots\ ]]]]]$ (SCC)

[1] For more recent versions of the SCC, see Chomsky (1995), Chomsky (2001b), Collins (1997), Kitahara (1997), Bošković and Lasnik (1999), and Freidin (1999), among others.

[2] Here and henceforth, I write "PIC$_n$" when I refer to a specific version of this constraint (there will be three all in all), and "PIC" when I do not discriminate between the different versions.

Crucially, the PIC_1 does not allow an operation involving Y and an item in WP. Chomsky (2001b) argues that such operations are in fact attested, though, and he gives the following example: Suppose that YP = TP, XP = vP, and WP = VP. The PIC_1 then precludes an operation involving T and NP in VP; but such an operation must arguably be legitimate for instances of long-distance agreement with VP-internal nominative NPs, attested in a number of languages. Chomsky's solution is to weaken the phase impenetrability requirement in such a way that a phase is evaluated with respect to the PIC at the next phase level; PIC_1 is accordingly replaced by PIC_2 (see Chomsky (2001b, 14)).

(4) *Phase Impenetrability Condition*$_2$ (PIC_2):
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.

As a consequence, the derivational 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. This is shown in (5). Agreement operations involving T and VP-internal nominative NPs are now predicted to be legitimate.

(5) *Search space under PIC*$_2$:

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

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

Since the empirical focus of the present paper will be on superiority (-like) effects with *wh*-movement, let me now address the mechanics of *wh*-movement in the SCC/PIC-based approach. Movement in general is viewed 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 phase edges, i.e., Specv and SpecC. However, the need for successive-cyclic movement does not automatically provide a trigger for such movement (given that the grammar is not equipped with look-ahead capacity). If we assume that all movement operations must be triggered by certain kinds of features, it is clear that there must be such features on heads of phases that trigger intermediate movement steps to phase edges. These features must be optional (so as to prevent derivations without *wh*-movement or other unbounded dependencies from crashing); ideally, they should only occur when they are needed. To this end, the requirement in (6) is proposed in Chomsky (2000, 109), Chomsky (2001b, 34); I will refer to this as the Optional EPP Feature Condition.

(6) *Optional EPP Feature Condition*:
The head X of phase XP may be assigned an EPP-feature (after the phase XP is otherwise complete), but only if that has an effect on outcome.

It is by no means evident how "having an effect on outcome" can be understood in a strictly local way, without look-ahead. However, for the moment, I will simply presuppose here that the Optional EPP Feature Condition can indeed be checked locally.[3] On this basis, consider the (simplified) derivation of a *wh*-question involving clause-bound *wh*-movement in English. EPP features show up obligatorily on T and on C marked [wh]; in addition, there is an optional EPP feature on v that is inserted in accordance with the Optional EPP Feature Condition.

(7) (I wonder) what John read

a. $[_{VP}\ \text{read}_3\ \text{what}_1\ ]$
b. $[_{vP}\ \text{what}_1\ \text{John}_2\ \text{read}_3\ [_{VP}\ t_3\ t_1\ ]]$ (EPP on v)
c. $[_{TP}\ \text{John}_2\ \text{T}\ [_{vP}\ \text{what}_1\ t_2\ \text{read}_3\ [_{VP}\ t_3\ t_1\ ]]]$ (EPP on T)
d. $[_{CP}\ \text{what}_1\ \text{C}\ [_{TP}\ \text{John}_2\ \text{T}\ [_{vP}\ t'_1\ t_2\ \text{read}\ [_{VP}\ t_3\ t_1\ ]]]]$ ([wh], EPP on C)

A further assumption that is usually made in this kind of approach is that syntactic operations like movement are subject to a Minimal Link Condition (MLC), as in (8) (see Chomsky (2000, 123), Chomsky (2001b, 27)).

(8) *Minimal Link Condition* (MLC):
If β and γ both match a probe α and β asymmetrically c-commands γ, a syntactic operation cannot involve α and γ.

The MLC is essentially a feature-based version of the Superiority Condition in Chomsky (1973); in cases of potential ambiguity where two items could act as goals for a given probe, only the higher one can in fact participate in the operation. The MLC has a number of interesting consequences (for superiority and other effects); but there are also several well-known problems with a simple version of this constraint. An obvious problem is that subject raising from a vP-internal position to SpecT is wrongly expected to be blocked by the MLC if object movement to Specv has occurred.[4] Thus, *what*$_1$ is closer to T in (7-c) than t_2, and should therefore have precluded movement of *John*$_2$ to SpecT. Several solutions to this problem have been proposed. Chomsky (1995) envisages a way out in terms of the concept of "equidistance," which plays a role instead of the notion of "asymmetrical c-command" in the formulation of the MLC. The equidistance approach is abandoned again in Chomsky (2000), Chomsky (2001b) in favour of the stricter formulation of the MLC in (8). The problem that the MLC poses for subject raising in (7) is then addressed by observing that after *wh*-movement of *what*$_1$ to SpecC, the subject NP is the closest

[3] Following Heck and Müller (2000), I will suggest a slightly different constraint to replace the Optional EPP Feature Condition in subsection 3.1 below; this latter constraint can be locally evaluated.

[4] At least, this holds as long as we assume that object movement must end up in a position in vP that is higher than the base position of the subject; but see Richards (2001) for a different view.

goal for T after all (the intervening object having left its position). At first sight, it seems that an execution of this idea implies giving up the SCC: Movement in TP would have to follow movement in CP, in violation of strict cyclicity. Still, Chomsky suggests that there is a way out of this dilemma that respects both the SCC and the MLC in strict versions: The idea is that the MLC is not evaluated at each step of the derivation; rather, it is only evaluated at the phase level. Thus, subject raising in (7-c) would indeed violate the MLC; but TPs are not phases, and the MLC is therefore not operative at this stage. The MLC does apply to the output in (7-d) because CP is a phase. However, at this point, there is no overt NP in Specv left that would separate the subject trace and T, and, given some obvious adjustments, it follows that the MLC is respected. Of course, there is now a change of perspective that is non-trivial: The MLC cannot be conceived of as a derivational constraint on operations anymore; it acts as a representational constraint on certain kinds of structures (viz., trees with phases at the root).

This concludes the sketch of movement operations in the incremental-derivational approach developed in Chomsky (2000), Chomsky (2001b). In the next subsection, I will argue that both the MLC and the $\text{PIC}_{1,2}$ emerge as suboptimal from a point of view that takes the task of reducing derivational search space seriously; and I will argue that the MLC should be dispensed with completely in favour of a more restrictive version of the PIC.

2.2. Conceptual Considerations

It is an attractive feature of incremental-derivational approaches to syntax that complexity can be reduced, compared to representational approaches. Such reduction of complexity becomes manifest in three different domains. First, the system does not permit look-ahead: At any given stage of the derivation, operations in later cycles and their effects cannot be considered. Second, the system relies on cyclicity: At any given stage of the derivation, the SCC makes it impossible to target a position (i.e., locate a probe) by a syntactic operation that is not included in the minimal XP. And third, the system incorporates a phase impenetrability requirement ($\text{PIC}_{1,2}$) that significantly reduces the search space for the goal of an operation. In effect, all syntactic material in the domain that the PIC renders opaque can (and must) be ignored for the remainder of the derivation.[5] So far, so good. However, closer inspection reveals conceptual problems with both the MLC and the two versions of the PIC: First, the MLC inherently depends on a certain amount of search space to work on. And second, it turns out that the PIC_1 and, in particular, the PIC_2 could reduce search space even more radically. More specifically, given the overall goal of search space reduction, the MLC/$\text{PIC}_{1,2}$-based approach to movement creates three conceptual problems.

[5]This consequence is particularly obvious if we assume the concept of cyclic spell-out, according to which domains that have been rendered inaccessible via the PIC are immediately sent off to the phonological and semantic interfaces; see Chomsky (2001a, 4).

2.2.1. Weak and Strong Representationality

In his comparison of derivational and representational approaches to syntax, Brody (2001) observes that a representational approach can be strictly non-derivational. In contrast, a derivational approach is usually representational to some extent, by adhering to the very concept of syntactic structure. Brody calls a derivational approach *weakly representational* if "derivational stages are transparent (i.e., representations), in the sense that material already assembled can be accessed;" and he calls it *strongly representational* if it "is weakly representational and there are constraints on the representations." On this view, the approach sketched in the previous subsection is strongly representational: This is not the fault of the SCC or the PIC (in either version); these are derivational constraints on operations. In the formulation given in (8), the MLC is also a derivational constraint; however, this is not the case anymore if we re-interpret the MLC in the way suggested at the end of the previous section to account for the existence of subject raising in examples like (7). Here, the MLC is a representational constraint that is evaluated at the phase level; it checks the legitimacy of structures rather than operations. Brody concludes from this (and from related observations) that a representational approach has an inherent advantage over a derivational approach in this domain. Let us assume that the argument is correct. Then, given a derivational approach, the task will be to reduce its representational residue – ideally, a derivational theory should not even be weakly representational. This implies abandoning all constraints that presuppose too much structure (in a sense to be made precise); a good candidate for exclusion then is the MLC.

2.2.2. A Redundancy

Interestingly, a simultaneous adoption of the MLC and the PIC leads to redundancies: As noted by Chomsky (2001b, 47, fn. 52), "the effect on the MLC is limited under the PIC, which bars 'deep search' by the probe." Thus, the MLC can only become relevant in the relatively small portions of structure permitted by PIC_1/PIC_2; it thus loses much of its original empirical coverage. Against the background of Brody's argument involving (weak or strong) representationality of derivational approaches, this can be viewed as further evidence that derivational approaches should dispense with the MLC in toto. I would like to contend that, 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, and then chooses between two items that may in principle participate in a given operation (as is done by the MLC). Rather, minimality effects should emerge as epiphenomena of constraints that reduce the space in which the derivation can look for items that may participate in an operation (as is done by the PIC); ideally, all competition among items (that a priori qualify for some operation) that must be resolved is in fact independently resolved if the search space is sufficiently small.

2.2.3. An Asymmetry

The SCC and the PIC have complementary tasks and look like two sides of the same coin. Therefore, it is a potentially suspicious property of the system laid out above that the two constraints rely on syntactic domains of such a different size. In one case (SCC), it is the phrase, in the other, it is the phase (PIC). In an optimally designed system, we would expect more symmetry in domains for probe and goal localization: Either 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).

My goal in what follows is to develop a derivational approach that evades these three conceptual problems by exhibiting the following properties: First, the material that can be accessed at any given step of the derivation is an extremely small bundle of categories with virtually no internal structure that can hardly be called a representation anymore. Hence, the approach to be developed will not even be weakly representational.[6] Second, the MLC is dispensed with in favour of a strengthened version of the PIC. Third, the new version of the PIC has the same kind of local domain as the SCC: the phrase.

3. Phrase Impenetrability

3.1. Assumptions

Following Sternefeld (2000), I assume a system in which two types of features participate in movement operations. On the one hand, there are [*F*] features that trigger movement as probes (to specifier positions, for the cases considered in this paper, and directly, without recourse to additional generalized EPP features). On the other hand, there are corresponding [F] features on items that turn them into goals for a movement operation triggered by [*F*]. The constraint that brings about movement is the Feature Condition; the constraint that requires all movement to be feature-driven is Last Resort.

(9) *Feature Condition*:
An [*F*] feature on X requires movement of an item marked [F] to the edge of X.

(10) *Last Resort*:
Movement requires matching [F] and [*F*] at an edge.

The SCC remains the same; (1) is repeated here as (11). However, the PIC is now restricted to phrases; see the PIC_3 in (12).[7]

[6] Whether or not one still insists on calling these objects representations is no more than a terminological issue. What counts is the extreme reduction of representations to small, virtually unstructured objects, which leads to a system in which Brody's conceptual objection loses its force.

[7] As such, it closely resembles the Head Constraint developed by van Riemsdijk (1978) (see also the Bounding Condition proposed by Koster (1978)). Note that this denies a special role of CP and vP for the purposes of movement theory (contra Chomsky (2000), Chomsky (2001b), Fox (2000), Nissenbaum

(11) *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.

(12) *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.

A comparison of the abstract derivations in (3) (under PIC$_1$) and (5) (under PIC$_2$) with the abstract derivation in (13) shows that the new PIC$_3$ is more restrictive in the sense that derivational search space is minimized.

(13) *Search space under PIC*$_3$:

a. $\overbrace{\underbrace{[_{\underline{\text{YP}}}\ldots\text{Y}\,[_{\underline{\text{XP}}}\ldots[_{\text{X}'}\text{X}}_{\text{PIC}_3}}^{\text{SCC}}\ [_{\underline{\text{WP}}}\ldots\text{W}\ [_{\underline{\text{UP}}}\ldots\text{U}\ldots\ ]]]]]$

b. $\overbrace{\underbrace{[_{\underline{\text{ZP}}}\ldots\text{Z}\,[_{\underline{\text{YP}}}\ldots\text{Y}}_{\text{PIC}_3}}^{\text{SCC}}\ [_{\underline{\text{XP}}}\ldots\ [_{\text{X}'}\text{X}\ \ [_{\underline{\text{WP}}}\ldots\text{W}\ [_{\underline{\text{UP}}}\ldots\text{U}\ldots\ ]]]]]]$

Finally, recall that so far, intermediate movement steps required by the PIC were triggered by optional EPP features demanded by the Optional EPP Feature Condition. I would now like to suggest that the role of the Optional EPP Feature Condition is played by the constraint Phrase Balance. This constraint is a straightforward adaptation of the constraint Phase Balance developed in Heck and Müller (2000). Ph(r)ase Balance arguably captures the underlying idea of the Optional EPP Feature Condition; and it does so without running into the danger of invoking look-ahead – the pieces of information that must be taken into account for the purposes of Phrase Balance at any given stage of the derivation are locally available, either in the present tree, or in the workspace of the derivation, which must be accessible throughout (I will address this concept immediately).

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

The concept of potential availability of a feature remains to be defined; this can be done as in (15).

(2000), Bruening (2001), Barbiers (2002), and others). However, the revised approach is of course compatible with the all the evidence suggesting that SpecC and Specv are used by successive-cyclic movement. Moreover, the concept of phase does not necessarily have to be abandoned: Phases are independently motivated (semantically, as propositional objects), and may or may not figure as special derivational units in other parts of the theory. Note finally that the present approach is therefore not as radical as the one pursued in Epstein and Seely (2002) (where the relevant move is not from phase to phrase, but from phase to derivational step).

(15) *Potential availability*:
A feature [F] is potentially available if (i) or (ii) holds:
(i) [F] is on X or edgeX of the present root of the derivation.
(ii) [F] is in the workspace of the derivation.

The workspace of a derivation D comprises the numeration N and material in trees that have been created earlier (with material from N) and have not yet been used in D. Phrase Balance triggers movement without feature matching in cases where the Feature Condition does not force movement (viz., to intermediate positions). However, Last Resort clearly prohibits such movement. In view of this state of affairs, I will assume (following again Heck and Müller (2000)) that Last Resort is minimally violable if this is the only way to fulfill the inviolable constraints Feature Condition, SCC, PIC_3, and Phrase Balance.[8]

As a consequence of Phrase Balance, *wh*-movement must proceed via every XP on the way to its ultimate target position (the $\text{C}_{[*wh*]}$ node that attracts it, because of the Feature Condition).[9] The reason is this: As long as there is a C bearing the feature [*wh*] in the numeration, and no [wh] feature on either another item in the numeration, or in a tree that has been formed earlier, a root XP of the current derivation can only be balanced if non-feature-driven *wh*-movement takes place to its specifier. The derivation of a simple *wh*-question under these assumptions is given in (16) (compare (7)); material that is crossed out has been rendered inaccessible by the PIC_3, and is thus not available anymore for further operations in the derivation.[10]

(16) (I wonder) what John read
a. $[_{\text{VP}}\ \text{what}_1\ \text{read}_3\ \text{t}_1\ ]$
$\rightarrow$ workspace: $\{\text{C}_{[*wh*]}, \text{John}, \text{T}_{[*D*]}, \text{v}\}$
b. $[_{\text{vP}}\ \text{what}_1\ \text{John}_2\ \text{v+read}_3\ [_{\text{VP}}\ \text{t}'_1\ \text{t}_3\ \cancel{\text{t}_1\ ]]}$
$\rightarrow$ workspace: $\{\text{C}_{[*wh*]}, \text{T}_{[*D*]}\}$
c. $[_{\text{TP}}\ \text{what}_1\ \text{John}_2\ \text{T}\ [_{\text{vP}}\ \text{t}''_1\ \text{t}_2\ \text{v+read}_3\ \cancel{[_{\text{VP}}\ \text{t}'_1\ \text{t}_3\ \text{t}_1\ ]]]}$
$\rightarrow$ workspace: $\{\text{C}_{[*wh*]}\}$
d. $[_{\text{CP}}\ \text{what}_1\ \text{C}\ [_{\text{TP}}\ \text{t}'''_1\ \text{John}_2\ \text{T}\ \cancel{[_{\text{vP}}\ \text{t}''_1\ \text{t}_2\ \text{v+read}_3\ [_{\text{VP}}\ \text{t}'_1\ \text{t}_3\ \text{t}_1\ ]]]]}$

[8] This can be encoded in an optimality-theoretic manner by a ranking {Feature Condition, SCC, PIC_3, Phrase Balance} $\gg$ C $\gg$ Last Resort. Note that an additional constraint C would be needed to ensure that the higher-ranked constraints are in fact never violable in a well-formed output: C punishes the candidate derivation 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. See Müller (2000) and Fanselow and Féry (2002) for discussion of these and further options in optimality-theoretic syntax.

[9] The resulting system is thus close to analyses in Sportiche (1989), Sportiche (1998), Takahashi (1994), and Agbayani (1998), among others. It also bears a certain resemblance to GPSG analyses that rely on SLASH feature percolation (Gazdar (1981), Gazdar et al. (1985)), to the approach in terms of gap marker percolation developed by Stechow and Sternefeld (1981), and to Koster (2000)'s analysis based on feature percolation in gap phrases.

[10] Note that Phrase Balance forces movement of the *wh*-phrase within VP already, so as to displace [wh] to the edge of V.

→ workspace: {–}

Given that Phrase Balance forces intermediate, non-feature-driven *wh*-movement only if there is otherwise no potentially available [wh] feature, and given that Last Resort can only be violated if this is the only way to fulfill constraints like Phrase Balance, the prediction is that presence of an accessible [wh] feature in the workspace should make non-feature-driven *wh*-movement of a *wh*-phrase impossible, and the PIC_3 should then block any further operations applying to this *wh*-phrase. The next subsection shows that this prediction is borne out, and that it offers a simple account of superiority effects in English, without recourse to a constraint like the MLC.

3.2. Superiority Effects in English

Examples illustrating superiority effects in English are shown in (17) (for subject and object) and in (18) (for two objects): Given two *wh*-phrases that qualify in principle as goals for movement operations targetting a single $C_{[*wh*]}$ probe, only the higher *wh*-phrase can undergo movement to the target position.

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

(18) 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] ?

These superiority effects can be derived under the assumptions adopted so far. In both cases, the lower *wh*-phrase NP_2 has a chance to leave the (right-peripheral) complement position of the VP that it is merged in only if it first moves to the (left-peripheral) SpecV position; this is so because of the PIC_3. A priori, there are two conceivable ways to move NP_2 to SpecV. First, the Feature Condition might trigger [*F*]-driven movement to SpecV (Specv, SpecT, ...). This is not an option in English, which has neither object shift nor scrambling. Second, movement of NP_2 to SpecV might be triggered by Phrase Balance, as in (16). However, this is not an option either in (17) and (18): VP is balanced because there is another *wh*-phrase in the workspace, viz., NP_1. The vP and TP categories that dominate this VP are balanced in (17) because the *wh*-phrase NP_1 occupies the respective specifiers (as a result of Merge and [*D*]-driven movement, respectively); they are balanced in the embedded clause in (18) because the *wh*-phrase NP_1 is still in the workspace. Thus, any attempt to derive a sentence like (17-b) will automatically result in a sentence like (17-a); as the derivation in (19) shows, the decision against *wh*-movement of the object NP_2 is made very early, at the first stage, where NP_2 cannot move to SpecV (similarly for (18-a) vs. (18-b)).

(19) a. [$_{VP}$ $bought_3$ $what_2$]
→ workspace: {$C_{[*wh*]}$, $who_{1[wh]}$, $T_{[*D*]}$, v}

b. $[_{vP}$ who$_1$ v+bought$_3$ $[_{VP}$ t$_3$ ~~what$_2$]]~~
→ workspace: $\{C_{[*wh*]}, T_{[*D*]},\}$

c. $[_{TP}$ who$_1$ T $[_{vP}$ t$_1$ v+bought$_3$ ~~[$_{VP}$ t$_3$ what$_2$]]]~~
→ workspace: $\{C_{[*wh*]}\}$

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

Double object constructions provide an interesting testing ground for approaches to superiority effects. Constructions with ditransitive verbs basically come in two varieties in English: the prepositional object construction, and the dative shift construction. As far as the prepositional object construction is concerned, it has been observed by Chomsky (1973, 246) and Fiengo (1980, 123) that either object (NP or PP) can move if both are *wh*-phrases (see (20-ab)); however, preposition stranding (which is legitimate if the remaining NP object is not a *wh*-phrase) becomes impossible in this context (see (20-c)).

(20) 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$_1$ $[_{PP_3}$ to t$_2$] ?

The situation is different in dative shift constructions with two *wh*-objects. Barss and Lasnik (1986, 349) note that the second object NP can never move in this context; the first, dative-shifted object NP can marginally move.[11]

(21) a. (?)Who$_2$ did you give t$_2$ what$_1$?
b. *What$_1$ did you give who$_2$ t$_1$?

To account for these data, we need to say something about the structure of double object constructions in English, and about pied piping in *wh*-PPs. Modifying the proposal in Larson (1988), I assume that the direct (i.e., THEME) object is merged in a complement position of V, where it remains in both types of double object constructions (unless it undergoes movement to the clausal periphery). The indirect (i.e., GOAL) object is at the edge of V if it has undergone dative shift (see (22-b)), and in a right-peripheral V′ sister position if it is prepositional (see (22-a)).[12] Note that this right-peripheral position does not belong to the edge of V.[13]

(22) a. $[_{VP}$ $[_{V'}$ $[_{V'}$ V NP$_1$] $[_{PP_3}$ P NP$_2$]]]
b. $[_{VP}$ NP$_2$ $[_{V'}$ V NP$_1$]]

[11] The marginality is due to a general weak ban on A-bar movement of dative-shifted objects in English and thus independent of superiority; see Stowell (1981, ch. 4) and Larson (1988), among others.

[12] To accomodate evidence from binding theory, we must then assume that linear order is relevant; see Barss and Lasnik (1986) and Jackendoff (1990).

[13] However, if we follow Chomsky (2002, 133-136), this position will invariably be a specifier (i.e., non-first Merge) position. Thus, as will be shown momentarily, at this point it is crucial that (15) refers to edgeX rather than to SpecX.

With respect to pied piping, I assume that there is optional percolation of the feature [wh] in *wh*-PPs; for present purposes, this percolation operation can be conceived of as an actual feature displacement.[14]

Consider now first the examples involving prepositional object constructions in (20). Given that feature percolation of [wh] from NP to PP is optional, we have to take into account two possibilities. First, suppose that [wh] percolation has taken place, and PP bears [wh]. The two objects are merged in VP-internal non-edge positions; see (22-a). Hence (given that there is no [wh] waiting in the workspace), Phrase Balance forces movement of one *wh*-phrase to SpecV so as to balance the VP (there is a [*wh*] on C in the numeration). It does not matter which of the two *wh*-phrases moves to SpecV, but whichever *wh*-phrase moves 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. This way, (20-a) and (20-b) can both emerge as grammatical. Consider now the second option: [wh] percolation from NP to PP has not taken place. Then, PP cannot move to SpecV – if it moves, Phrase Balance will not be satisfied because [wh] is not potentially available at the VP level since it is not part of the edge of V (it is dominated by an edge element – PP – but not on an edge element itself). NP_2 in PP cannot move either, though: To leave PP, NP_2 must move to SpecP, given the PIC_3. However, this operation is not legitimate because there is no [*F*] that might trigger it (English does not have an independent PP-internal preposing operation), and because Phrase Balance is independently satisfied (with another *wh*-phrase in the workspace). Therefore, the superiority effect in (20-c) is correctly predicted.

Turning next to dative shift constructions as in (21), it follows from (22-b) that NP_2 is in SpecV for independent reasons.[15] Hence, Phrase Balance can be fulfilled without a Last Resort violation, and any such violation incurred by movement of the lower *wh*-phrase will be fatal. Consequently, NP_2 can undergo *wh*-movement (see (21-a)), but NP_1 cannot undergo such movement, because of the PIC_3 (see (21-b)).

The analysis makes a further prediction: If both *wh*-phrases are embedded in PPs, preposition stranding is predicted to be blocked throughout: It is impossible for an embedded *wh*-phrase to move to SpecP in this context because Phrase Balance is always satisfied without such movement. Here is why. For the first *wh*-phrase NP_i that is merged with P (be it NP_1 or NP_2), the PP is balanced without local inversion of NP_i to the edge of P because there is a [wh] feature on another *wh*-item left in the numeration. For the second *wh*-phrase NP_j that is merged with P, movement to the edge of PP will also be blocked because there is now invariably a tree in the

[14] Heck (2001) develops a more elaborate theory of pied piping that does without feature percolcation. This approach can be reconciled with the present analysis, but I will refrain from doing so, for reasons of space and coherence.

[15] There are two possibilities: Either NP_2 is merged in SpecV, or it is moved there because of some [*F*] feature that triggers dative shift to that position. The present analysis is compatible with both a base-generation and a movement approach to dative shift constructions.

workspace that contains (or is) a *wh*-phrase bearing a [wh] feature. Consequently, no *wh*-phrase can move to SpecP in this context, and subsequent movement of such a *wh*-phrase from its base position will fatally violate the PIC_3. By and large, this prediction seems to be tenable, as the data in (23) illustrate.[16]

(23) a. ?*Who$_2$ did you give [$_{NP}$ pictures of t_2] [$_{PP}$ to whom$_1$] ?
b. ?*Who$_1$ did you give [$_{NP}$ pictures of whom$_2$] [$_{PP}$ to t_1] ?
c. ?*Who$_2$ did you talk [$_{PP}$ to t_2] [$_{PP}$ about whom$_1$] ?
d. ?*Who$_1$ did you talk [$_{PP}$ to whom$_2$] [$_{PP}$ about t_1] ?

To sum up this subsection, the present approach accounts both for standard superiority effects in English, and their absence in certain kinds of double object constructions, without invoking the MLC, by the interaction of Phrase Balance and the PIC_3. I will now turn to the situation in German.

3.3. The Lack of Superiority Effects in German

It has often been observed that German does not exhibit superiority effects with *wh*-phrases that are clause-mates; see Haider (1983), Haider (1993), Haider (2000), Grewendorf (1988), and Bayer (1990), among many others. A relevant pair of examples involving a *wh*-subject NP and a *wh*-object NP is given in (24).

(24) 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

Similarly, German does not exhibit superiority effects with control infinitives; see Fanselow (1991), Kim and Sternefeld (1997), and Haider (2000).[17] This is shown in (25):

(25) 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

[16] It should be noted, however, that there is some disagreement about the status of these examples. Sentences like (23-d) are classified as ill formed in Jackendoff (1990, 433)), and as well formed in Fiengo (1980, 124). Furthermore, (23-a) and (23-c) are classified as acceptable by Jackendoff; but note that these examples are in fact expected to involve an additional violation of the *Clause Nonfinal Incomplete Constituent Constraint*; see Kuno (1973, 379), Lasnik and Saito (1992, 91)). This constraint is operative independently of multiple-*wh* (superiority) contexts; see (i-a) vs. (i-b).

(i) a. Who$_2$ did you give [$_{NP}$ pictures of Mary] [$_{PP}$ to t_2] ?
b. ?*Who$_1$ did you give [$_{NP}$ pictures of t_1] [$_{PP}$ to John] ?

[17] However, see Haider (2000, 239) for an additional dissimilarity requirement on the two *wh*-phrases.

Various accounts of the lack of superiority effects with two *wh*-phrases that share a minimal finite clause have been given in the literature. I will here adopt an analysis that has been suggested by Fanselow (1996) and Grohmann (1997) (who assume that the MLC underlies superiority effects):[18] German has scrambling: A lower *wh*-phrase can independently be moved to a higher position, by *wh*-scrambling. Thus, a lower *wh*-phrase cannot move across another *wh*-phrase merged in a higher cycle by *wh*-movement, given Phrase Balance; but it can do so by scrambling. To implement this analysis in the present approach, I assume that scrambling is triggered by a designated optional feature (or feature bundle) that we can refer to as [*Σ*]; accordingly, scrambled items bear [Σ] features (see Müller (1998), Sauerland (1999), Grewendorf and Sabel (1999)). For our present concerns, it is immaterial whether Σ is a formal feature that is not interpreted, or can in fact be shown to be related to contentful notions that are sometimes viewed as triggers for scrambling (definiteness, specificity, animacy, focus, and the like). The derivation of a sentence like (24-b) can then proceed as in (26), where the *wh*-object NP_2 first undergoes Phrase Balance-driven movement to SpecV (because of [*Σ*] on v, not because of [*wh*] on C), and then Feature Condition-driven movement to Specv (because of [*Σ*]). At the vP level, both *wh*-phrases show up at the edge; hence, NP_1 and NP_2 are both in principle eligible for further movement (given the PIC_3); such further movement is triggered by Phrase Balance on the TP cycle, and by the Feature Condition in the final step (CP).[19]

(26) a. $[_{VP}$ was$_{2,[\Sigma]}$ $[_{V'}$ t_2 gesagt]]
→ workspace: {$C_{[*wh*]}$, $wer_{1[wh]}$, T, $[_{v}$ hat $]_{[*\Sigma*]}$}

b. $[_{vP}$ was$_{2,[\Sigma]}$ wer_1 $[_{VP}$ t'_2 ~~$[_{V'}$ t_2~~ gesagt]] $[_{v}$ hat]]
→ workspace: {$C_{[*wh*]}$, T}

c. $[_{TP}$ was$_{2,[\Sigma]}$ $[_{vP}$ t''_2 wer_1 ~~$[_{VP}$ t'_2 $[_{V'}$ t_2 gesagt]]~~ $[_{v}$ hat]] T]
→ workspace: {$C_{[*wh*]}$}

d. $[_{CP}$ was$_{2,[\Sigma]}$ C $[_{TP}$ t'''_2 ~~$[_{vP}$ t''_2 wer_1 $[_{VP}$ t'_2 $[_{V'}$ t_2 gesagt]] $[_{v}$ hat]]~~ T]
→ workspace: {–}

This concludes the account of the lack of superiority effects in German in the present approach.[20] The prediction is that the PIC_3 should give rise to superiority effects

[18] The basic idea can already be found in the analysis of the lack of weak crossover effects in German that is developed in Grewendorf (1988, 320). Other accounts of the lack of superiority effects in German include Haider (1983), Noonan (1988), Bayer (1990), Haider (1993), Müller (1995), Richards (2001), Kim and Sternefeld (1997), Haider (2000), Pesetsky (2000), and Grewendorf (2001). Fanselow (1991), Wiltschko (1997), Grohmann (1998), and Featherston (2001) suggest re-evaluations of the empirical evidence.

[19] As noted in Fanselow (1990), Müller and Sternefeld (1993), and elsewhere, *wh*-scrambling often leads to reduced acceptability (but not strict ungrammaticality), which is not attested in cases like (24-b) and (25-b). However, reduced acceptability may result not from the application of *wh*-scrambling as such, but from the surface position of a scrambled *wh*-phrase. Since *wh*-scrambling is subsequently undone in a derivation like (26), this restriction will not apply.

[20] English does not have scrambling; but it does exhibit topicalization. Thus, it has to be ensured that the account of superiority effects in English is not undermined by intermediate *wh*-topicalization.

after all in German if the lower *wh*-phrase cannot be moved to the domain occupied by the higher *wh*-phrase because scrambling is not available (for whatever reason). The following three subsections highlight three contexts where *wh*-scrambling is impossible in German; and it is in these contexts that superiority effects do indeed occur.

3.4. Superiority Effects with Long-Distance Movement in German

The first such context is well known: As observed by Frey (1993), Büring and Hartmann (1994), Fanselow (1996), Heck and Müller (2000), Pesetsky (2000), and others, German does exhibit superiority effects with long-distance movement. This is shown by the contrast in (27).

(27) a. Wer$_1$ hat t_1 geglaubt [$_{CP}$ dass der Fritz wen$_2$ mag] ?
who$_{nom}$ has believed that the Fritz whom$_{acc}$ likes
b. *Wen$_2$ hat wer$_1$ geglaubt [$_{CP}$ dass der Fritz t_2 mag] ?
whom$_{acc}$ has who$_{nom}$ believed that the Fritz likes

The analysis is straightforward. First, as before, NP_2's [wh] feature in (27) does not permit movement: Phrase Balance is satisfied by the presence of NP_1 in the workspace; therefore, movement of NP_2 for the purposes of [wh] will fatally violate Last Resort. Second, and more importantly in the present context, NP_2 cannot move by scrambling either: Scrambling cannot leave a finite CP in German. Consequently, an embedded *wh*-phrase is correctly predicted to be stuck in the embedded clause if there is another *wh*-phrase in the workspace that is eventually merged in the matrix clause.[21]

3.5. Superiority Effects with Subject Raising in German

The second context in which there are a priori unexpected superiority effects in German involves subject raising.[22] NP raising to subject position is optional in German (see Diesing (1992)). In the present approach, this implies that the EPP feature [*D*] is optional on T. However, as shown in Haider (1993, ch. 8), the evidence cited in Diesing (1992) and much related work in favour of subject raising to SpecT (based on phenomena like particle placement) is far from conclusive. As far as I can see, there is only one context where it is clear that subject raising to SpecT must

Indeed, *wh*-topicalization is independently excluded in English (and other languages); see Epstein (1992) and Müller and Sternefeld (1996) for analyses and further references.

[21] As a matter of fact, only v and V can tolerate [Σ] in their edge domains; no other kind of head provides a scrambling domain (see Müller (1995)). Hence, there can be no [*Σ*] feature on T or C that could trigger movement beyond vP. Still, something extra will ultimately have to be said to derive the ban on long-distance scrambling in German in toto: It must be ensured that a [*Σ*] feature on a matrix V or v cannot attract an XP bearing [Σ] in the lower clause. There are various ways of achieving this; but I will not pursue the matter here.

[22] To the best of my knowledge, this observation is new; an informal survey suggests that the data are quite robust.

have occurred in German (see Müller (2001, 296)): Unstressed pronouns must be at the *phonological* border of vP (in the sense of Chomsky (2001b, 34)), i.e., they cannot be preceded by non-pronominal material within TP (in contrast, stressed pronouns behave like non-pronominal NPs). There is but one exception: The subject NP, and only the subject NP, can optionally precede these pronouns within TP. This strongly suggests a special position that is available only for subject NPs. Hence, we can conclude that if a subject NP precedes unstressed pronouns, it must have undergone optional movement to SpecT.

Interestingly, there is a clear superiority effect in exactly this context. Since we need an unstressed object pronoun to ensure that subject raising has taken place, relevant examples involve ditransitive verbs. The contrast in (28) shows that a dative *wh*-object NP cannot undergo *wh*-movement to SpecC if a *wh*-subject occurs in front of an unstressed accusative object pronoun.

(28) a. Wem_2 hat [$_{\text{vP}}$ es t'_2 wer_1 t_2 gegeben] ?
whom_{dat} has it_{acc} who_{nom} given
b. ?*Wem_2 hat wer_1 [$_{\text{vP}}$ es t'_2 t_1 t_2 gegeben] ?
whom_{dat} has who_{nom} it_{acc} given

The contrast in (29) shows the same for an accusative *wh*-object NP and an unstressed dative object pronoun.

(29) a. Was_2 hat [$_{\text{vP}}$ ihm t'_2 wer_1 t_2 gegeben] ?
what_{acc} has him_{dat} who_{nom} given
b. ?*Was_2 hat wer_1 [$_{\text{vP}}$ ihm t'_2 t_1 t_2 gegeben] ?
what_{acc} has who_{nom} him_{dat} given

This superiority effect follows under present assumptions: Suppose that a subject $\text{NP}_{[wh]}$ and an object $\text{NP}_{[wh]}$ are both in Specv at some stage of the derivation, and that T has an optional [*D*] feature. Then, TP is balanced (for [*wh*]) by feature-driven subject raising, and movement of the object NP incurs a fatal Last Resort violation.[23]

If this analysis is on the right track, we expect that a non-*wh*-subject NP should, ceteris paribus, not block movement of an object *wh*-phrase. This is the case: Only a *wh*-subject NP in SpecT blocks *wh*-movement of an object NP; see (30).

(30) a. Wem_2 hat t''_2 der Fritz_1 [$_{\text{vP}}$ 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 [$_{\text{vP}}$ ihm t'_2 t_1 t_2 gegeben] ?
what_{acc} has the Fritz_{nom} him_{dat} given

Similarly, replacing the unstressed object pronoun with a non-pronominal object NP should void the superiority effect (other things being equal). The reason is that the

[23] Haider (2002) argues that Icelandic has optional subject raising, and that it exhibits superiority effects with subject NPs only when the subject NP is in SpecT, not when it is in Specv. This generalization can be derived in the same way.

wh-subject does not have to be in SpecT in this context (non-pronominal NPs do not have to be at the phonological border of vP). (31) shows that this prediction is borne out, too.

(31) a. Wem$_2$ hat [$_{vP}$ t$'_2$ wer$_1$ t$_2$ das Buch gegeben] ?
whom$_{dat}$ has who$_{nom}$ the book$_{acc}$ given
b. Was$_2$ hat [$_{vP}$ t$'_2$ wer$_1$ dem Fritz t$_2$ gegeben] ?
what$_{acc}$ has who$_{nom}$ the Fritz$_{dat}$ given

3.6. Superiority Effects with Scrambling from Wh-XP in German

A third context in which superiority effects arise in German has been noted in Fanselow (1996). The construction involves a configuration where the two *wh*-phrases are initially not in a c-command relation (as in all the examples discussed thus far); rather, one dominates the other. More specifically, suppose that a *wh*-phrase PP_1 is dominated by a *wh*-phrase NP_2 (it has been merged with NP_2's head), as in *wieviele Bücher über wen* ('how many books about whom'). Suppose furthermore that PP_1 can be moved out of NP_2 without violating locality constraints. This implies that NP_2 is in object position when extraction takes place (otherwise, the *Condition on Extraction Domain* (CED) would be violated that permits extraction from XP only if XP occupies a complement position). It also implies that NP_2 is embedded by a certain kind of verb (verbs like *lesen* ('read') permit extraction from NP, verbs like *zerstören* ('destroy') do not); that NP_2 is sufficiently non-specific (highly specific NPs like *welches Buch* ('which book') tend to block extraction, non-specific NPs like *wieviele Bücher* ('how many books') do not); etc. Then, a *wh*-PP_1 can be scrambled from a *wh*-NP_2 if there is a [Σ] feature on PP_1 and a [*Σ*] feature on a higher head (V or v). PP_1 moves to the edge of V, driven either by the Feature Condition (if [*Σ*] is on V) or by Phrase Balance (if [*Σ*] is on v). This stage of the derivation is depicted in (32).

(32) [$_{VP}$ [$_{PP_1}$ über wen] [$_{V'}$ [$_{NP_2}$ t$'_1$ wieviele Bücher t$_1$] lesen]]
about whom how many books read

Here, PP_1 occupies an edge position of VP, and NP_2 a complement position. As shown by the contrast in (33), it is indeed the case that only PP_1 can undergo further movement, as one might expect: Movement of NP_2 on the next (vP) cycle will have to violate the PIC_3, or so it seems.

(33) a. (Ich weiß nicht) [$_{PP_1}$ über wen] er [$_{NP_2}$ wieviele Bücher t$_1$] lesen
I know not about whom he how many books read
will
wants

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

However, there is a gap in this reasoning: (32) closely mirrors the situation found with two *wh*-phrases that are co-arguments in German, which do not normally exhibit superiority effects. Thus, if NP_2 can undergo scrambling in (32), it should be able to undergo further *wh*-movement to SpecC after all. Fanselow (1996) solves this problem by showing that a derivation of (33-b) that involves intermediate scrambling of NP_2 will invariably violate another constraint: the principle of *Unambiguous Domination* (see Müller (1998, 271)). Unambiguous Domination is essentially a constraint on the movement of remnant XPs, i.e., XPs from which movement has taken place. This constraint states that α-traces must not be α-dominated (in the domain of the head of the chain). For the case at hand, this means: A scrambling trace like t_1 must not be dominated by a category that has itself undergone scrambling. This precludes intermediate scrambling of NP_2 in (32). Consequently, any derivation of (33-b) will have to violate either Unambiguous Domination or PIC_3, depending on whether intermediate scrambling of NP_2 does or does not take place.

No such effect is predicted to occur if PP_1 is not a *wh*-phrase. Now, NP_2 can (in fact, must, given Phrase Balance) move to SpecVP in (32); this movement is not an instance of scrambling because there is no [*Σ*] involved (be it directly or indirectly). Compare (33-b) with (34).

(34) (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

To sum up, German does exhibit superiority effects in certain contexts. These contexts have in common that intermediate scrambling of the second *wh*-phrase is not available, for independent reasons (scrambling in German cannot leave a finite clause, cannot target TP, and cannot apply to XPs from which scrambling has taken place). The effects are then derivable from the PIC_3.

3.7. Superiority-Like Effects with Remnant Movement in German

Let me make a brief digression at this point. Recall that it is a major goal of this paper to show that the MLC can be dispensed with in a derivational grammar because typical MLC effects follow straightforwardly from a strict version of the PIC that is independently motivated by conceptual considerations. As we have seen in the last subsection, a constraint like Unambiguous Domination proves necessary to account for one such effect (whether we adopt the MLC, as in Fanselow (1996), or the PIC_3). Interestingly, however, it has been argued that Unambiguous Domination can itself be derived from a version of the MLC (defined in terms of closeness

rather than asymmetric c-command; see Takano (1994), Koizumi (1995), Kitahara (1997), Müller (1998), and Sauerland (1999)). In a nutshell, the idea is this: 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; and that is β, not γ. Hence, β must move first, and subsequent movement of γ must incur a violation of the CED (because γ-extraction takes place from β in a non-complement position, which β must be in after movement), and, if γ-movement is to a position that follows β, an additional violation of the general ban on lowering (which is arguably derivable from the SCC, given some minor modification; see Müller (1998)). Thus, Unambiguous Domination effects (as they show up in (35-a) vs. (35-b) in German) turn out to be derivable from the MLC. On this view, the only relevant difference between typical Unambiguous Domination configurations (as in (35-a)) and typical superiority configurations is that the two items that compete for movement (because they have the same [F] feature attracted by a higher [*F*]) are in a dominance relation in the first case, and in a c-command relation in the second.

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

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

The question arises of whether the present system based on the PIC$_3$ also directly accounts for dominance-related MLC effects, in addition to the c-command-related MLC effects discussed so far. The answer is no: The illformedness of (35-a) does not follow from the PIC$_3$. To see this, suppose that there are two [*Σ*] features, one for NP$_1$, one for VP$_2$. Then, there should be a well-formed derivation for (35-a), with NP$_1$ undergoing Phrase Balance-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, this does not imply that Unambiguous Domination must be stated as such. Its effects can be derived from a more general constraint: a simple version of the A-over-A Condition.

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

This version of the A-Over-A Condition forces VP$_2$ movement to apply first in (35-a); subsequent NP$_1$ lowering then violates (at least) the CED.[24]

[24]Note that the distinction between head and edge element in the definition of the A-Over-A Condition is the only case where a minimal structural differentiation of the bundle of categories accessible for further operations seems necessary; recall the discussion in subsection 2.2.

3.8. Intervention Without C-Command in German

The three types of superiority effects in German that were discussed in subsections 3.4, 3.5, and 3.6 as such do not differentiate between PIC-based and MLC-based analyses. However, it is worth noting that, in stark contrast to what is the case with an MLC account, there is nothing in the PIC_3-based account that would tie the intervention effect incurred by a *wh*-phrase wh_1 for another *wh*-phrase wh_2 to a c-command (or dominance) relation between the two. All that is needed for an intervention effect to arise in the PIC_3-based analysis is that wh_1 enters the derivation that wh_2 is part of at a later stage, and wh_2 cannot end up in the same edge domain as wh_1 by some independently motivated movement operation. Consequently, we expect that there should be *wh*-intervention effects without c-command.

As noted in Heck and Müller (2000), it is indeed the case that non-c-commanding *wh*-phrases in a matrix clause block long-distance *wh*-movement in German. This superiority-like effect without c-command is exemplified by the contrast in (37). In (37-a), there is clause-bound *wh*-movement of NP_1 across an adverbial CP that contains another *wh*-phrase NP_2, and that is merged later; here, an intervention effect can be avoided because NP_1 can reach a position in the same edge domain as the adverbial CP by scrambling. However, the option of intermediate scrambling is not available for long-distance *wh*-movement; scrambling must stop in the embedded vP domain. Thus, the presence of NP_2 in the workspace blocks Phrase Balance-driven movement of NP_1, and (37-b) emerges as ungrammatical because of the PIC_3. In (37-c), it is NP_2 rather than NP_1 that undergoes *wh*-movement; the result is also ill formed. As in (37-b), a PIC_3 violation cannot be avoided here: No matter whether the adverbial CP is created before or after the object CP, Phrase Balance cannot trigger successive-cyclic movement of NP_2 because NP_1's [wh] feature is potentially available for $C_{[*wh*]}$ in the workspace. In addition, sentences like (37-c) are ruled out by the CED: Movement of NP_2 takes place from an adverbial CP that does not occupy a complement position.[25] The overall result is that the

[25]This means that, if nothing else is said, examples of the type in (37-c), where a *wh*-phrase XP_2 is embedded in some other phrase that c-commands the *wh*-phrase XP_1, are predicted to be ill formed even if movement of XP_2 to $SpecC_{[*wh*]}$ does not violate the CED or another locality constraint, as long as XP_2 cannot reach the main branch by some other movement operation like scrambling. Relevant examples are hard to find, though. In most pertinent cases, XP_2 will have to cross an island, and in the few well-formed constructions where locality constraints can be respected, XP_2 can usually undergo scrambling first (given some proviso concerning the coherence/incoherence distinction with control infinitives). Also recall the discussion of the examples in (23) in English. However, should there turn out to be clear cases of well-formed instantiations of the structure in (i) (where α is not an island and wh_2 cannot reach γ by an independently available non-*wh*-movement operation), the present approach would be in need of a modification.

(i) ... wh_2 ... [$_\gamma$ [$_\alpha$ ... t_2 ...] ... [$_\beta$ ... wh_1 ...] ... ...] ...

One possibility would be to make the definition of workspace of a derivation sensitive to the distinction between main and minor branches (such that features on the main branch would not count as potentially available when a derivation proceeds in a minor branch). Then, structures like (37-c) would not (have to) violate the PIC_3 anymore, and *wh*-movement in (i) would be predicted to be legitimate if α is not an

numeration underlying (37-b) and (37-c) cannot yield a well-formed output.

(37) a. Wen$_1$ hat Fritz [$_{CP}$ nachdem er was$_2$ gemacht hat] t$_1$ getroffen ?
whom has Fritz after he what done has met

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

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

Exactly the same reasoning applies in (38), where NP$_2$ shows up in a relative clause CP that is in turn dominated by an NP:

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

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

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

Yet another set of examples that illustrates the same pattern is given in (39); here the intervening *wh*-phrase that blocks long-distance *wh*-movement is embedded in a simple NP.

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

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

island. – It might also be worth noting at this point that replacing requirement (ii) in the definition of potential availability in (15) by the stricter requirement (ii)′ would lead to an approach that is very close to an MLC-based system, with *wh*-intervention effects reduced to c-command environments.

(ii)′ [F] is on X or edgeX of a root in the workspace of the derivation (lexical items are trivial roots).

c. *[$_{PP_2}$ Von wem] hat Fritz [$_{NP}$ einem Freund t_2] gesagt [$_{CP}$ dass
of whom$_{acc}$ has Fritz a friend$_{dat}$ said that
Maria wen$_1$ liebt] ?
Maria whom$_{acc}$ loves

The prediction is that the same kind of superiority-like effect without c-command should be detectable in subject raising constructions. The contrast between (40-a) and (40-b) may not be one of perfect wellformedness vs. absolute ungrammaticality; but the tendency is clear enough, and conforms to expectations: (40-a) is much better than (40-b). In (40-a), the subject NP does not have to be in SpecT, and the *wh*-object NP can therefore move to a position in front of it by scrambling; this option is not available in (40-b), where the subject NP must be in SpecT (because of the presence of the unstressed object pronoun), i.e., in a domain that cannot be reached by scrambling. (40-c) is also excluded by the PIC$_3$; in addition, it is blocked by the CED because the subject NP does not occupy a complement position.[26]

(40) a. ?Wem$_1$ hat [$_{NP}$ die Frau [$_{PP_2}$ mit welchem Mantel]] t_1 ein Buch
whom$_{dat}$ has the woman with which coat a book$_{acc}$
gegeben ?
given

b. ?*Wem$_1$ hat [$_{NP}$ die Frau [$_{PP_2}$ mit welchem Mantel]] es t_1
whom$_{dat}$ has the woman with which coat it$_{acc}$
gegeben ?
given

c. *[$_{PP_2}$ Mit welchem Mantel] hat [$_{NP}$ die Frau t_2] es wem$_1$
with which coat has the woman it$_{acc}$ whom$_{dat}$
gegeben ?
given

3.9. Intervention without C-Command in English

The system developed so far makes yet another prediction: There is no clause-bound intervention effect in the examples in (37-a), (38-a), (39-a), and (40-a) because German has scrambling. Since English does not have scrambling, we expect clause-bound intervention effects with non-commanding *wh*-phrases to occur. At first sight, this seems to contradict the standard view that argument *wh*-in situ in English does not obey any island constraints (see Chomsky (1981), Huang (1982), Huang (1995), Lasnik and Saito (1992), and Hornstein (1995), among others). However, it is worth noting that most of the pertinent examples in the literature do not involve intervention without c-command: The typical kind of multiple *wh*-question

[26] In principle, one would expect the same kind of superiority-like effect to also occur with examples involving scrambling from *wh*-XPs, as in (33). However, relevant examples that would show this are difficult to construct because they would have to involve multiple embedding within NP, which creates difficulties of various kinds in multiple questions, for (presumably) independent reasons.

that is taken to argue for non-island-sensitivity of *wh*-in situ in English has one *wh*-phrase embedded in an island, and a second *wh*-phrase merged in a higher position. This latter *wh*-phrase then undergoes movement to SpecC, as in the examples in (41), where a *wh*-phrase dominated by an object NP or by an adjunct PP does not block *wh*-movement of a subject *wh*-phrase merged later.

(41) a. Who_1 t_1 saw [$_{\text{NP}}$ the man that bought what_2] ?
b. Who_1 t_1 likes [$_{\text{NP}}$ books that criticize who_2] ?
c. Who_1 t_1 bought [$_{\text{NP}}$ the books on which table_2] ?
d. Who_1 t_1 met [$_{\text{NP}}$ friends of whom_2] ?
e. I wonder who_1 t_1 heard [$_{\text{NP}}$ the claim that John had seen what_2]
f. I wonder who_1 t_1 heard [$_{\text{NP}}$ John's stories about what_2]
g. Who_1 t_1 left [$_{\text{PP}}$ despite which warning_2] ?

Similarly, a *wh*-phrase that is part of a subject NP does not block movement of a *wh*-phrase that is merged in a higher clause, as in (42).

(42) Who_1 t_1 thinks that [$_{\text{NP}}$ pictures of who_2] are on sale ?

All this is expected under present assumptions: NP_2 cannot undergo Phrase Balance-driven movement in (41) or (42) early in the derivation because there is another item bearing [wh] left for [*wh*] of C in the workspace, and the PIC_3 precludes Feature Condition-driven movement at the end of the derivation. NP_1, in contrast, undergoes movement from the edge of v to the edge of T in accordance with Phrase Balance, and is then forced to end up at the edge of C by the Feature Condition.

However, consider now the case where *wh*-intervention without c-command does occur. In the examples in (43), an object *wh*-phrase that is merged first (NP_2) moves across a subject NP containing another *wh*-phrase (NP_1). Such movement results in significantly reduced acceptability, as predicted under the present PIC_3-based approach.[27]

(43) a. ?*Who_2 did [$_{\text{NP}}$ the man that bought what_1] see t_2 ?
b. ?*Who_2 did [$_{\text{NP}}$ books that criticize who_1] impress t_2 ?
c. ?*What_2 did [$_{\text{NP}}$ the books on which table_1] cost t_2 ?

[27] The data in (43) were checked with various native speakers, who unanimously declared them to be ill formed, and who all found a sharp contrast in the minimal pairs that can be formed on the basis of (41) and (43). However, I am aware of one exception to the apparent general neglect of constructions like those in (43) in the literature: Such examples are discussed in Fiengo et al. (1988) and, following them, Fitzpatrick (2002), and judged grammatical. I have nothing to say here about the source of the diverging judgements, except for the observation that Fiengo et al. (1988) are primarily concerned with contrasting the construction in (43), with a *wh*-phrase embedded in a subject NP and an object *wh*-phrase ending up in front of it, with one in which the subject NP-internal *wh*-phrase undergoes movement (in violation of the CED) and the object *wh*-phrase stays in situ – and not with one in which a *wh*-phrase is embedded in an object NP and a subject *wh*-phrase undergoes movement. In other words: One might speculate that judgement differences arise in this domain because different kinds of minimal pairs are taken into account, and judgements are taken to be relative rather than absolute.

d. ?*Who$_2$ did [$_{NP}$ friends of whom$_1$] meet t$_2$?
e. *Who$_2$ did [$_{NP}$ friends of whom$_1$] say that we should invite t$_2$?

The explanation is completely analogous to that given for the ungrammatical German examples involving intervention without c-command in the last subsection: Non-feature-driven movement of NP_2 to the edge of V at an early stage in the derivation is not forced by Phrase Balance (because the VP is balanced anyway, with NP_1's [wh] feature matching the matrix C's [*wh*] feature in the workspace), and therefore excluded by Last Resort. Consequently, any movement operation applying to NP_2 at later stages of the derivation fatally violates the PIC_3.[28]

3.10. Further Refinements

The approach developed in this paper imposes severe restrictions on *wh*-movement; as a matter of fact, it turns out to be slightly too restrictive in two domains.

3.10.1. Multiple $C_{[*wh*]}$ Domains and Intervention

The first problem concerns sentences like (44) in German (see Heck and Müller (2000)).

(44) [$_{NP}$ Die Frage [$_{CP_6}$ wer$_1$ C t$_1$ was$_2$ mitbringt]]] ist relevant für die
the question who what brings is relevant to the
Frage [$_{CP_5}$ wie$_3$ C Fritz denkt [$_{CP_4}$ t$'_3$ dass die Party t$_3$ wird]
question how Fritz thinks that the party will be

(44) is clumsy, but well formed. Consider the underlying numeration. There are two C heads bearing [*wh*] features (C_6 and C_5), and there are three *wh*-pronouns bearing [wh] features – hence, one of the two $C_{[*wh*]}$ heads will have to give rise to a multiple question. The example has been designed in such a way that CP_6 is the multiple question, and CP_5, which is merged earlier, is a simple question. The task now is to ensure that the *wh*-phrase *wie* ('how') can undertake steps of successive-cyclic movement until it reaches the edge of T of CP_5, where it is attracted by $C_{[*wh*]}$. Unfortunately, successive-cyclic movement of *wie*$_3$ turns out to be blocked at the very first stage under present assumptions: At the point where it must be decided whether *wie*$_3$ can move by violating Last Resort, the phrase is wrongly predicted to be balanced: There are two C heads bearing [*wh*] features, and there are two remaining items in the workspace that bear corresponding [wh] features. Closer inspection reveals that the same kind of problem also shows up in simpler sentences in German (see (45-a)) and in English (see (45-b)).

[28] In contrast, the MLC would not make the right predictions. If the MLC is defined in terms of asymmetric c-command (see (8)), all sentences in (43) are ceteris paribus predicted to be well formed; if it is defined in terms of closeness (see section 3.7), it will also wrongly permit *wh*-movement of NP_2 in (43), at least in those cases where NP_1 is deeply embedded.

(45) a. [$_{CP_5}$ Wer_1 hat t_1 wen_2 gefragt [$_{CP_4}$ was_3 Fritz t_3 mag]] ?
who_{nom} has $whom_{acc}$ asked $what_{acc}$ $Fritz_{nom}$ likes
b. [$_{CP_5}$ Who_1 t_1 asked $whom_2$ [$_{CP_4}$ $what_3$ C John likes t_3]] ?

As it stands, the *wh*-phrase NP_3 (*was*/*what*) cannot reach the edge of T, from where it can be attracted by C_4 bearing [*wh*] in accordance with the PIC_3. The highest position that NP_3 can be in prior to *wh*-movement to SpecC is the edge of v in (45-a) in German (due to this language's scrambling options), and the complement position of V in English (due to this language's lack of scrambling options).

Intuitively, the problem with (44), (45-a), and (45-b) is clear: A *wh*-phrase that is part of the workspace must not interact with a *wh*-phrase in a given derivation if the two *wh*-phrases target different $C_{[*wh*]}$ domains. Following Heck and Müller (2000), this problem can be solved by minimally enriching the representation of *wh*-features. Thus, suppose that both [*wh*] and [wh] features are accompanied by scope indices in the numeration, and that *wh*-phrases can only be interpreted with a given C node if they share a scope index. Under this assumption, a feature $[wh]_i$ on a *wh*-item can never count as potentially available for a feature $[*wh*]_j$ on a C in the workspace, due to feature mismatch. For the cases at hand, this means that the [wh] feature of NP_2 must be accompanied by the same scope index as the [*wh*] feature of C_6 in (44), and of C_5 in (45-a) and (45-b), in order to be interpretable as part of the multiple question. Therefore, at the point where the question of non-feature-driven movement of NP_3 must be decided, NP_2 does not intervene anymore: A feature like $[wh]_5$ on NP_2 in (45-ab) can never satisfy Phrase Balance for a feature like $[*wh*]_4$ on C_4. Only a feature like $[wh]_4$ on NP_3 can do so; accordingly, Phrase Balance forces successive-cyclic movement of NP_3.

3.10.2. D-Linking and Intervention

Wh-phrases that qualify as D(iscourse)-linked behave differently from other *wh*-phrases in a number of respects; see Pesetsky (1987) and much subsequent literature. One well-known pecularity of D-linked *wh*-phrases is that they do not induce intervention effects in English if they show up in situ; see the contrast in (46).[29]

(46) a. *I know [$_{NP_2}$ which books] who_1 read t_2
b. I know $what_2$ [$_{NP_1}$ which people] read t_2

The present analysis can accomodate standard accounts of this phenomenon straightforwardly. Thus, assume that D-linked *wh*-phrases in English can optionally lack a (proper) [wh]-feature that would make them accessible for a [*wh*] feature on C. Then, if a D-linked *wh*-phrase lacks the [wh] feature in the workspace, Phrase Balance can only be fulfilled by movement of the remaining *wh*-phrase, and an intervention effect can be avoided. (Of course, a [wh] feature must be present in

[29] The situation is different in German, where D-linking does not seem to have such effects; see, e.g., the examples in (40).

those cases where the D-linked *wh*-phrase itself undergoes *wh*-movement.)

4. Conclusion

Let me summarize the main results of this paper. First, I have argued that 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). Second, it follows from this move that the PIC_3 accounts for typical MLC effects in English without further ado. The MLC can therefore be dispensed with (except for a residue, the A-Over-A Condition). Third, given that German has scrambling of *wh*-phrases, superiority effects are predicted to be absent, except under those circumstances where scrambling is independently excluded (long-distance effects, subject position effects, remnant movement effects). Fourth and finally, unlike the MLC, the system based on Phrase Balance and the PIC_3 predicts superiority-like intervention effects without c-command, which are indeed attested.

Needless to say, the PIC_3 has important consequences for many other phenomena outside the domain of *wh*-constructions, especially if we we adopt the following hypothesis, which the approach assumed here lends itself to:

(47) 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 (47) effectively implies a derivational approach to semantic interpretation, i.e., cyclic semantic spell-out (see note 5; and Sternefeld (1996) for the sketch of such a model of interpretation). (47) also suggests that there is no reason left to assume the existence of traces (neither as t, nor as a copy): Given the PIC_3, these are not accessible for semantic interpretation, and there are no derivational constraints that apply to them. The hypothesis also raises interesting problems for binding of anaphors (at least for those cases that are not strictly local, and that therefore cannot be covered by the reflexivity constraints of Reinhart and Reuland (1993)) and pronouns; for control; for long-distance agreement; etc. In general, apparently non-local relations must decomposed into a succession of local steps, as proposed in Gazdar et al. (1985). More specifically, non-local relations could be accounted for by successive-cyclic local [F] feature movement from head to head (required by constraints of the Phrase Balance 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 LF movement of anaphors (see Chomsky (1986)).[30] For (obligatory) control, the strategy would amount to a decomposition of Landau (2000)'s Agree relation into small steps of feature move-

[30] Compare the account of A-chain condition effects in Reuland (2001). Also see Fischer (2002) for a derivational analysis of binding phenomena along these lines.

ment (or, indeed, a version of Hornstein (2001)'s A-movement approach). However, carrying out such analyses is beyond the scope of the present paper.

References

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

Barbiers, Sjef (2002): Remnant Stranding and the Theory of Movement. In: *Dimensions of Movement*. Benjamins, Amsterdam, pp. 47–67.

Barss, Andrew and Howard Lasnik (1986): A Note on Anaphora and Double Objects, *Linguistic Inquiry* 17, 347–354.

Bayer, Josef (1990): Notes on the ECP in English and German, *Groninger Arbeiten zur Germanistischen Linguistik* 30, 1–51.

Bošković, Željko and Howard Lasnik (1999): How Strict is the Cycle?, *Linguistic Inquiry* 20, 691–703.

Brody, Michael (2001): Some Aspects of Elegant Syntax. Ms., University College London.

Bruening, Benjamin (2001): Syntax at the Edge: Cross-Clausal Phenomena and the Syntax of Passamaquoddy. PhD thesis, MIT, Cambridge, Mass.

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

Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

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

Chomsky, Noam (1986): *Knowledge of Language*. Praeger, New York.

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

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

Chomsky, Noam (2001a): Beyond Explanatory Adequacy. Ms., MIT, Cambridge, Mass.

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

Chomsky, Noam (2002): *On Nature and Language*. Cambridge University Press, Cambridge.

Collins, Chris (1997): *Local Economy*. MIT Press, Cambridge, Mass.

Diesing, Molly (1992): *Indefinites*. MIT Press, Cambridge, Mass.

Epstein, Samuel David (1992): Derivational Constraints on $\overline{\text{A}}$-Chain Formation, *Linguistic Inquiry* 23, 235–259.

Epstein, Samuel David and T. Daniel Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. In: S. D. Epstein and T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 65–89.

Fanselow, Gisbert (1990): Scrambling as NP-Movement. In: G. Grewendorf and W. Sternefeld, eds., *Scrambling and Barriers*. Benjamins, Amsterdam, pp. 113–140.
Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.
Fanselow, Gisbert (1996): The Proper Interpretation of the Minimal Link Condition. Ms. Universität Potsdam.
Fanselow, Gisbert and Caroline Féry (2002): Ineffability in Grammar. In: G. Fanselow and C. Féry, eds., *Resolving Conflicts in Grammars*. Buske, Hamburg, pp. 265–307.
Featherston, Sam (2001): Universals and Grammaticality: Wh-Constraints in German and English. Ms., Universität Tübingen.
Ferguson, Scott and Erich Groat (1994): Defining 'Shortest Move'. Ms., Harvard University.
Fiengo, Robert (1980): *Surface Structure*. Harvard University Press, Cambridge, Mass.
Fiengo, Robert, Cheng-Teh James Huang, Howard Lasnik and Tanya Reinhart (1988): The Syntax of Wh-in-situ. In: H. Borer, ed., *Proceedings of WCCFL 7*. CSLI Publications, Stanford, pp. 81–98.
Fischer, Silke (2002): Bindung in lokal-derivationellen Theorien. Ms., Universität Stuttgart.
Fitzpatrick, Justin (2002): On Minimalist Approaches to the Locality of Movement, *Linguistic Inquiry* 33, 443–463.
Fox, Danny (2000): *Economy and Semantic Interpretation*. MIT Press, Cambridge, Mass.
Freidin, Robert (1999): Cyclicity and Minimalism. In: S. D. Epstein and N. Hornstein, eds., *Working Minimalism*. MIT Press, Cambridge, Mass, pp. 95–126.
Frey, Werner (1993): *Syntaktische Bedingungen für die Interpretation*. Akademieverlag, Berlin.
Gazdar, Gerald (1981): Unbounded Dependencies and Coordinate Structure, *Linguistic Inquiry* 12, 155–184.
Gazdar, Gerald, Ewan Klein, Geoffrey Pullum and Ivan Sag (1985): *Generalized Phrase Structure Grammar*. Blackwell, Oxford.
Grewendorf, Günther (1988): *Aspekte der deutschen Syntax*. Narr.
Grewendorf, Günther (2001): Multiple Wh-Fronting, *Linguistic Inquiry* 32, 87–122.
Grewendorf, Günther and Joachim Sabel (1999): Scrambling in German and Japanese, *Natural Language and Linguistic Theory* 17, 1–65.
Grohmann, Kleanthes (1997): German Superiority, *Groninger Arbeiten zur Germanistischen Linguistik* 40, 97–107.
Grohmann, Kleanthes (1998): Syntactic Inquiries into Discourse Restrictions on Multiple Interrogatives, *Groninger Arbeiten zur Germanistischen Linguistik* 42, 1–60.
Haider, Hubert (1983): Connectedness Effects in German, *Groninger Arbeiten zur Germanistischen Linguistik* 23, 82–119.

Haider, Hubert (1993): *Deutsche Syntax – generativ*. Narr, Tübingen.
Haider, Hubert (2000): Towards a Superior Account of Superiority. In: U. Lutz, G. Müller and A. von Stechow, eds., *Wh-Scope Marking*. Benjamins, Amsterdam, pp. 231–248.
Haider, Hubert (2002): Superiority Revisited: Dutch, English, German, Icelandic Contrasts. A Representational Account. Ms., Universität Salzburg.
Heck, Fabian (2001): Pied Piping Without Feature Percolation. Ms., Universität Stuttgart.
Heck, Fabian and Gereon Müller (2000): Successive Cyclicity, Long-Distance Superiority, and Local Optimization. In: R. Billerey and B. D. Lillehaugen, eds., *Proceedings of WCCFL*. Vol. 19, Cascadilla Press, Somerville, MA, pp. 218–231.
Hornstein, Norbert (1995): *Logical Form*. Blackwell, Oxford.
Hornstein, Norbert (2001): *Move. A Minimalist Theory of Construal*. Blackwell, Oxford.
Huang, Cheng-Teh James (1982): Logical Relations in Chinese and the Theory of Grammar. PhD thesis, MIT, Cambridge, Mass.
Huang, Cheng-Teh James (1995): Logical Form. In: G. Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*. Blackwell, Oxford, pp. 125–175.
Jackendoff, Ray (1990): On Larson's Account of the Double Object Construction, *Linguistic Inquiry* 21, 427–454.
Kim, Shin-Sook and Wolfgang Sternefeld (1997): Superiority vs. Crossover. Ms., Universität Tübingen.
Kitahara, Hisatsugu (1997): *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, Mass.
Koizumi, Masatoshi (1995): Phrase Structure in Minimalist Syntax. PhD thesis, MIT, Cambridge, Mass.
Koster, Jan (1978): *Locality Principles in Syntax*. Foris, Dordrecht.
Koster, Jan (2000): Variable-Free Grammar. Ms., University of Groningen.
Kuno, Susumo (1973): Constraints on Internal Clauses and Sentential Subjects, *Linguistic Inquiry* 4, 363–385.
Landau, Idan (2000): *Elements of Control*. Kluwer, Dordrecht.
Larson, Richard (1988): On the Double Object Construction, *Linguistic Inquiry* 19, 335–391.
Lasnik, Howard and Mamoru Saito (1992): *Move* α. MIT Press, Cambridge, Mass.
Müller, Gereon (1995): *A-bar Syntax*. Mouton/de Gruyter, Berlin.
Müller, Gereon (1998): *Incomplete Category Fronting*. Kluwer, Dordrecht.
Müller, Gereon (2000): *Elemente der optimalitätstheoretischen Syntax*. Stauffenburg, Tübingen.
Müller, Gereon (2001): Order Preservation, Parallel Movement, and the Emergence of the Unmarked. In: G. Legendre, J. Grimshaw and S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 279–313.

Müller, Gereon and Wolfgang Sternefeld (1993): Improper Movement and Unambiguous Binding, *Linguistic Inquiry* 24, 461–507.
Müller, Gereon and Wolfgang Sternefeld (1996): A-bar Chain Formation and Economy of Derivation, *Linguistic Inquiry* 27, 480–511.
Nissenbaum, Jon (2000): Covert Movement and Parasitic Gaps. In: M. Hirotani, A. Coetzee, N. Hall and J.-Y. Kim, eds., *Proceedings of NELS 30*. GLSA, Amherst, Mass, pp. 542–555.
Noonan, Máire (1988): Superiority Effects: How do Antecedent Government, Lexical Government and V/2 Interact?, *McGill Working Papers in Linguistics* pp. 192–214.
Perlmutter, David and Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.
Pesetsky, David (1987): Wh-in-Situ: Movement and Unselective Binding. In: E. Reuland and A. ter Meulen, eds., *The Representation of (In)Definiteness*. MIT Press, Cambridge, Mass, pp. 98–129.
Pesetsky, David (2000): *Phrasal Movement and Its Kin*. MIT Press, Cambridge, Mass.
Reinhart, Tanya and Eric Reuland (1993): Reflexivity, *Linguistic Inquiry* 24, 657–720.
Reuland, Eric (2001): Primitives of Binding, *Linguistic Inquiry* 32, 439–492.
Richards, Norvin (2001): *Movement in Language*. Oxford University Press, Oxford.
Sauerland, Uli (1999): Erasability and Interpretation, *Syntax* 3, 161–188.
Sportiche, Dominique (1989): Le Mouvement Syntaxique: Contraintes et Paramètres, *Langages* pp. 35–80.
Sportiche, Dominique (1998): *Partitions and Atoms of Clause Structure*. Routledge, London.
Stechow, Arnim von and Wolfgang Sternefeld (1981): A Modular Approach to German Syntax. Ms., Universität Konstanz.
Sternefeld, Wolfgang (1996): A Minimalist Semantics for Questions. Ms., Universität Tübingen.
Sternefeld, Wolfgang (2000): Syntax. Eine merkmalsbasierte Analyse. Book ms., Universität Tübingen.
Stowell, Tim (1981): Origins of Phrase Structure. PhD thesis, MIT, Cambridge, Mass.
Takahashi, Daiko (1994): Minimality of Movement. PhD thesis, University of Connecticut.
Takano, Yuji (1994): Unbound Traces and Indeterminacy of Derivation. In: M. Nakamura, ed., *Current Topics in English and Japanese*. Hituzi Syobo, Tokyo, pp. 229–253.
van Riemsdijk, Henk (1978): *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Foris, Dordrecht.
Wiltschko, Martina (1997): Scrambling, D-linking and Superiority in German, *Groninger Arbeiten zur Germanistischen Linguistik* 41, 107–142.