

REVIEWING PHASES:
ON SOME CONCEPTS AND CONSEQUENCES
OF PHASE-BASED MINIMALISM

JACEK WITKOŚ
Adam Mickiewicz University, Poznań

ABSTRACT

This paper presents and reviews basic theoretical concepts and empirical consequences of the recent minimalist view of derivations, which are supposed to run in cycles smaller than the clause and decidedly smaller than an entire complex sentence. The first part of the paper presents the new proposals that reduce the significance of the specifier/head relations and direct long-distance relations in grammar, proposing, in their place, the relation Agree and derivational phases. In subsequent parts, the paper focuses on some theoretical and empirical consequences of these assumptions. It appears that while the relation Agree, as a new checking configuration, is a tenable option, derivational phases, as defined by Chomsky, face considerable empirical challenges.

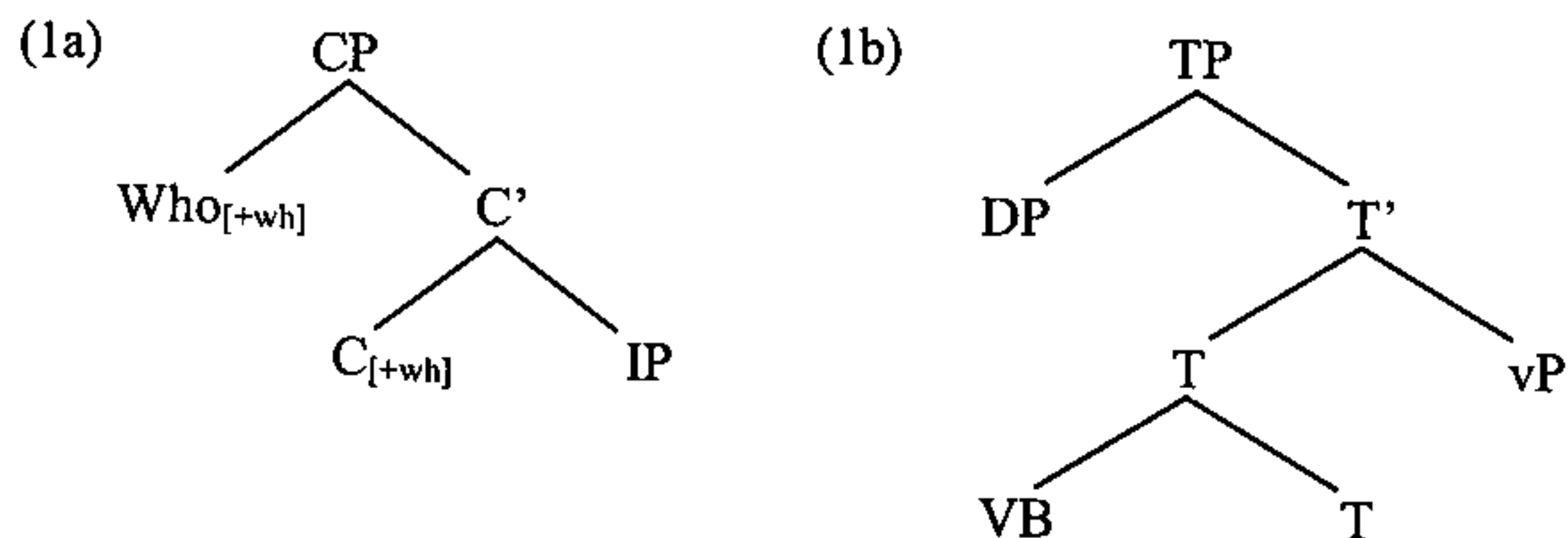
1. Introduction

The aim of this paper is to discuss the novelty and appeal of the recent stage in the development of the mainstream minimalist program (Chomsky 1998, 1999, 2001)¹ and review some its empirical consequences. This paper is organized as follows: its first part focuses on the review of certain novel assumptions and their empirical consequences, whereas the final sections are devoted to a critical discussion of a number of solutions proposed by Chomsky as well as a number of counterproposals put forward by his critics.

¹ All the page number references made to Chomsky (1998, *Minimalist inquiries: The framework*) and Chomsky (1999, *Derivation by phase*) correspond to their early *MIT Working Papers in Linguistics* editions. Page references to Chomsky (2001, *Beyond explanatory adequacy*) concern the manuscript version.

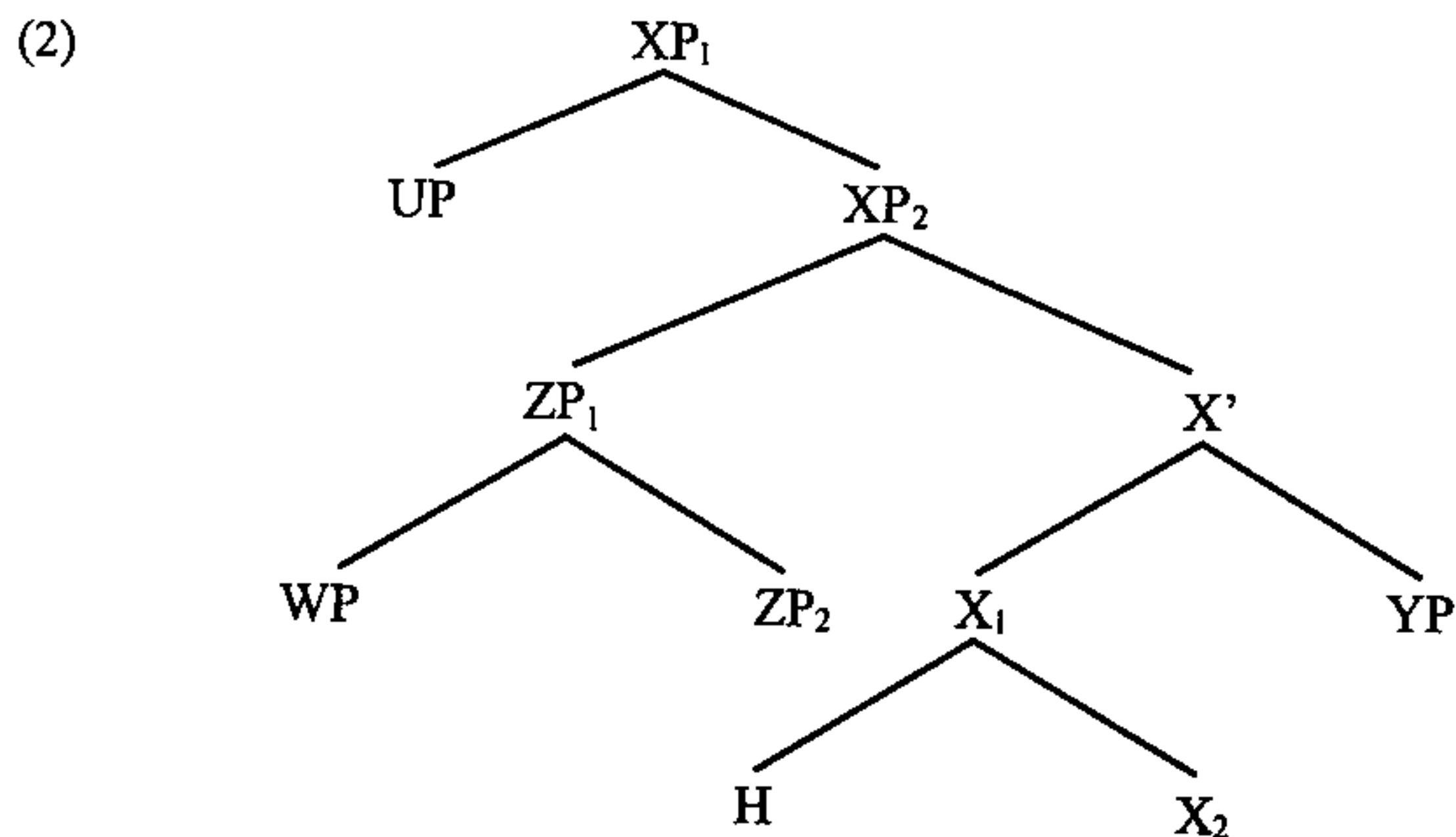
2. The relation of feature checking

One of the fundamental assumptions of the minimalist program is the notion of the checking configuration. The program started with two relations as the primitive configurations in syntax: spec/head and head/head:



Throughout the development of the program there has been a gradual shift of prominence in these two relations; while in Chomsky (1995, ch. 3) the former relation seemed to provide more empirical coverage (overt and covert substitution), the latter has become more prominent at stage two (covert feature movement).

Chomsky (1995, ch. 3) provides the following formal definition of the checking domain:



The maximal projection XP in the above structure has a specifier ZP, a head X and a complement YP. We have three instances of adjunction here: UP to XP (maximal projection to a maximal projection), changing XP into a two-segment category $\langle XP_1, XP_2 \rangle$; similarly WP to ZP resulting again in a two-segment category $\langle ZP_1, ZP_2 \rangle$, making the internal structure of the specifier more complicated; H to X (head-

to-head adjunction) turning X into $\langle X_1, X_2 \rangle$. Before we present definitions of domains, let us introduce two relevant concepts, namely *dominate* and *contain*:

- (3a) A category α dominates β if every segment of α dominates β .
 (3b) A category α contains β some segment of α dominates β .

In the above structure XP only contains UP, but does not dominate it, the same holds for ZP with respect to WP and X with respect to H; XP dominates ZP, WP, X, H and YP. These relations are crucial when defining domains. Following Chomsky (1995, ch.3), we give definitions of domains and necessary related concepts:

- (4a) Max (α) is the smallest maximal projection dominating α .
 (4b) The domain of a head α is the set of nodes contained in Max(α) that are distinct from and do not contain α .
 (4c) The complement domain of a head α is a subset of the domain reflexively dominated by the complement of the construction.
 (4d) The domain of a head α minus its complement domain will be called the residue of α .
 (4e) The minimal domain is the smallest subset K of the domain of α such that for any γ belonging to the domain, some β belonging to K dominates γ .
 (4f) The minimal complement domain of α is called the internal domain and the minimal residue of α is called its checking domain.

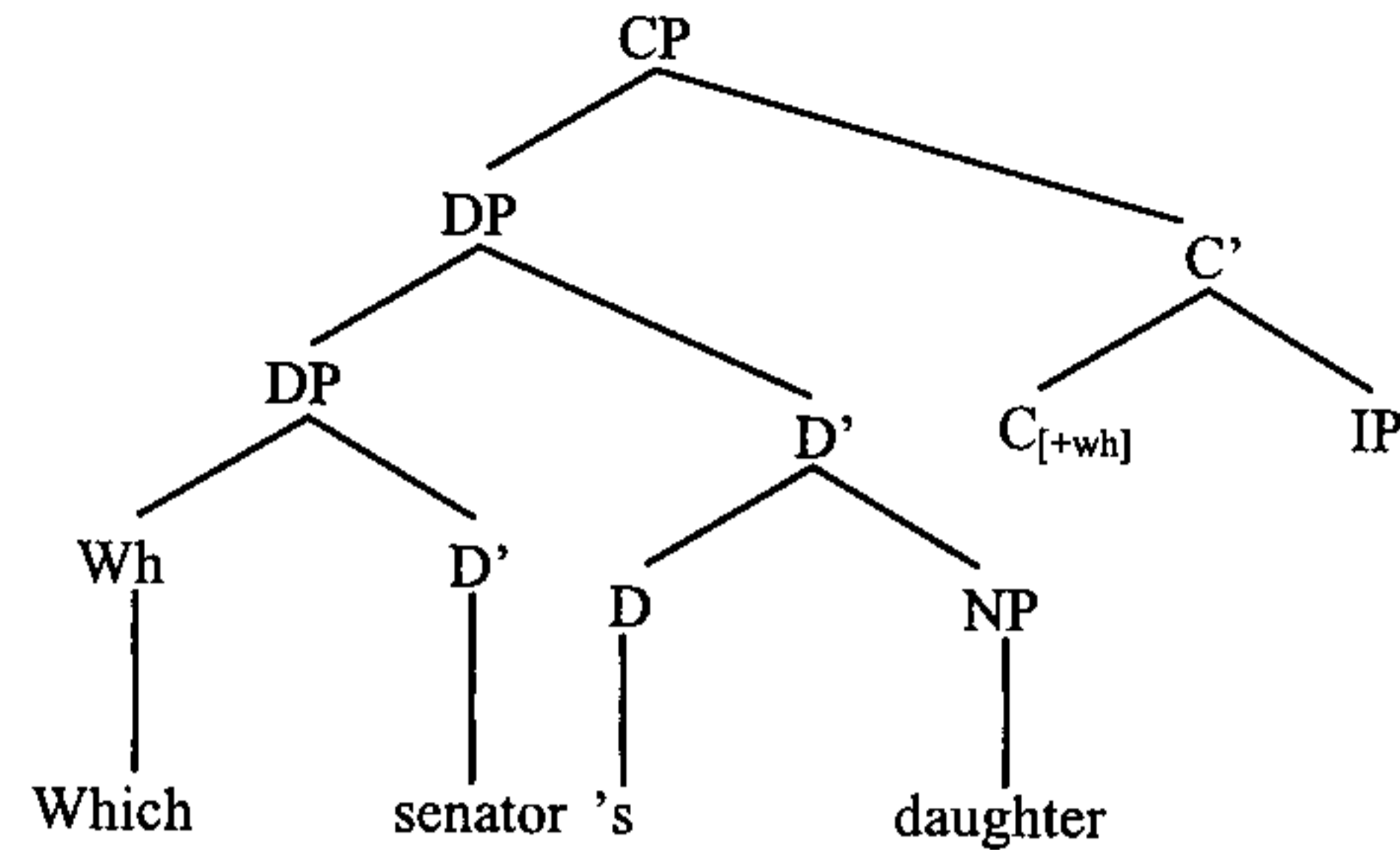
Thus the checking domain of a head includes its specifier and elements adjoined to the head but effectively excludes complements.

Still, both relations of spec/head and head/head are not free from conceptual problems and Chomsky (1998, 1999, 2001) sets out to refine the concept of checking.

2.1. Some problems with the specifier/head and head/head relations of feature checking

One thorny issue concerns feature percolation. Let us ponder over the issue of the spec/head configuration as a configuration of feature checking. As we see in (1a) the Wh phrase forms the specifier of the phrase whose label contains features facilitating movement and matching those of the specifier. The issue is less clear once a more complex syntactic object has moved into the specifier position:

(5a)



(5b) [CP [DP [DP which senator][D's daughter]] will you meet t]?

The [+wh] feature of the DP is quite clearly not in a spec/head configuration with the [+wh] feature on the label of CP. In (5a) the [+wh] feature is buried deep within the specifier of the specifier of C. How can the feature matching, checking and deletion take place in this configuration? This is a conceptual issue and such examples have been dealt with in the past with the aid of the notion of [+wh] feature percolation.² The [+wh] feature percolated up the tree within its own complex syntactic object and thus rendered the inclusive DP a Wh phrase. Now, however, it is not so obvious that the concept of percolation is relevant for feature checking. It certainly is relevant for pied-piping and movement of the entire DP but why would the [+wh] feature move or spread within the DP, if movement is tied to Last Resort? Note that once the complex specifier has been copied and merged in the [spec,C] position, the [+wh] feature cannot even move further to C via feature movement, as C does not c-command *which senator*.

A similar problem can be observed in (6) below, where a Wh-phrase is embedded within a PP:

(6) [PP to [DP which senator]] did John address his petition?

Here, the problem is aggravated, as the Wh-phrase is quite certainly not in the spec/head configuration with C, though the PP is. Again, the checking of the [+wh] feature can take place once the concept of feature percolation is employed. But why would the feature spread onto the PP?

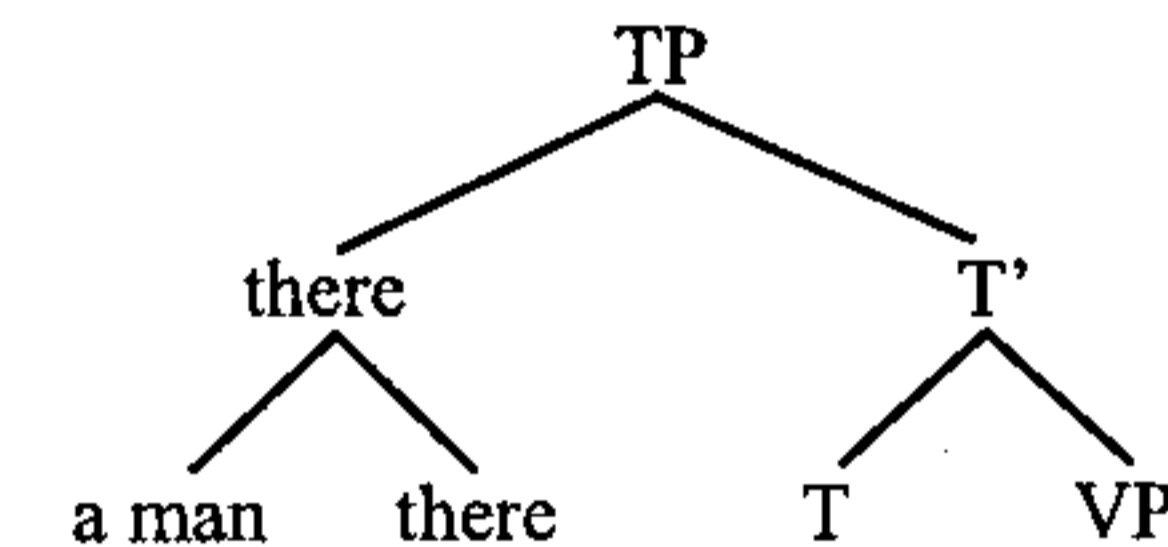
² The problem of pied-piping was discussed extensively as early as van Rimsdijk and Williams (1986).

The issue of what counts as a legitimate spec/head checking configuration is compounded by minimalist analyses even further. Lasnik (1995) observes that Chomsky's (1995, ch. 3) analysis of expletive constructions leads to a formation of an illicit extension of the spec/head checking relation; once the associate moves and adjoins to the expletive in LF, the following configuration obtains, (7b-c):

(7a) There is a man in the garden.

(7b) [TP [expl a man [expl there]] T is {a man} in the garden]

(7c)



The configuration in (7b) and (7c) allows for Nominative case checking between T and an adjunct to the specifier; clearly then the pure spec/head relation is too rigorous to capture all the relevant facts.³ Imagine what would happen in the system if an adjunct, and not only the specifier itself, could check features, for instance the [+EPP] feature of T. The example in (8b) should then be a legitimate construction in English:

(8a) Yesterday John would have arrived late.

(8b) *Yesterday would have (John) arrived John late.

Example (8b) is regularly dismissed as a violation of the requirement of the [+EPP] feature checking in overt syntax and this feature must be checked by a nominal specifier. If adjuncts could check this feature, *yesterday* could do it in example (8b).⁴

This overindulging configuration of feature checking under adjunction was revised in Chomsky (1995, ch. 4), where the features of the associate were raised to T in covert syntax and the specifier/head configuration satisfied the [+EPP] feature,

³ No other structural case is checked in such a configuration apart from the expletive construction, as Nominative and Accusative are checked through overt or covert substitution in Chomsky (1995, ch. 3). Note that *yesterday* is a NP and is in principle capable of checking the [+EPP] feature in:

(i) Yesterday was a disaster.

⁴ Note that the case feature of *John* can be checked in covert syntax, as in the expletive example in (4) above. The example includes an unaccusative verb to avoid causing ungrammaticality for additional reasons, such as the postulate that either a subject or an object must leave vP in narrow syntax.

while the head/head configuration facilitated the case feature checking. (PAGE REF).

However, the notion of covert movement is beset with another set of problems. As soon as the movement of features was allowed in the system, the question that immediately arose was how much burden of phrasal movement feature movement could take over. For example, is feature movement to be equated with A-movement? Is it able to change scope properties and binding domains? It should at least affect some of these properties, yet it generally fails to do so. Consider the following set of examples concerning the expletive/associate relation, a model case of formal feature (FF) movement in LF. First, we look at controls:⁵

- (9a) Many people aren't in the room.
- (9b) Someone may be in the room.
- (9c) Someone seems to be in the room.
- (9d) John expects someone that I do to be in the room.

The examples in (9a-c) show ambiguous scopal properties; in (9a) *many* may scope either under or over negation, in (9b) *someone* and *may* have ambiguous scope, in (9c) *someone* scopes either over or under *seem* and (9d) feeds Antecedent Contained Deletion. The wide scope properties of the indefinite pronoun in (9a-c) follow without any doubt from the fact that it may be interpreted in its overt position of the matrix subject, from which it c-commands the other scopal active elements. The narrow scope readings may result from reconstruction. Covert movement, not of formal features though, but of the entire DP, is required in (9d) to force parallel VP interpretations and resolution of the ellipsis:

- (10a) Many people aren't [_{PP} {many people} in the room].
- (10b) Someone may be [_{PP} {someone} in the room].
- (10c) Someone seems [_{IP} {someone} to be in the room].
- (10d) John [_{VP} someone that I do {{_{VP} expect [_{IP} t to be in the room]]}] [_{VP} expects [_{IP} t to be in the room]].

In brief, the point is as follows: why is there no independent empirical reflex of covert (feature) movement of the associate in the examples in (11):

- (11a) There aren't many people in the room.
- (11b) There may be someone in the room.
- (11c) There seems to be someone in the room.
- (11d) *John expects there to be someone that I do to be in the room.

⁵ These examples are discussed extensively in Den Dikken (1995), Hornstein and Witkoś (2003) and Witkoś (this volume).

The examples in (11) are unambiguous: *many* scopes under negation in (11a), so the only interpretation available is (not > many; few); *someone* scopes under *may* and *seem* in (11b-c) and ACD is not licensed in (11d), probably because feature movement of the associate *someone* to the expletive, or T, does not suffice to create parallel VPs and ellipsis cannot be resolved. In sum, it seems that despite feature movement, the associate behaves as if it were interpreted in situ at LF. Thus apart from the claim that there is covert movement of formal features of the DP, resulting in its case checking, based only on conceptual grounds, there is no independent evidence supporting this operation.⁶

In conclusion, the problem of creating the precise spec/head relation between the feature [+wh] and C⁰ and ensuing complexities presented above can be avoided if the checking is not performed in the spec/head configuration but in a different configuration, in which C⁰ could reach both simple and deeply buried Wh-phrases equally well. The relation of c-command seems to be a good candidate, as it is necessary in the system on independent grounds. Another advantage of a view of feature checking based on c-command is that the properties of examples in (9) and (11) would immediately fall in place: there is no feature movement at LF here. In all the examples in (11) the associates have their case checked in situ by the c-commanding T⁰ and are therefore 'frozen' for case-motivated LF movement of features, hence no reflex of the movement into the matrix subject position is to be expected.

2.2. The checking configuration

In Chomsky (1998, 1999, 2001), the procedure of checking uninterpretable features takes place as a result of a fundamental grammatical relation Agree holding between a Probe (functional head) and a Goal (substantive category). Assume that we are dealing with the following configuration, corresponding to TP at this stage of the derivation:⁷

- (12) T_[+φ, +EPP] [_{VP} DP_{SU} [_{+case, +φ}] v DP_{OB}]

Let us ignore the object DP for a while and focus on the well-known case and movement relation between the subject and Infl. The functional head T is the Probe and the subject DP is the Goal. The Probe is complete, that is it has full agreement features, which is a property of [+finite] T. The Probe matches the Goal for features: [+φ] features but there is no matching for the [+case] feature as in earlier stages of minimalism. Chomsky (1998, 1999, 2001) suggests that there are no case features

⁶ Note that examples (10d) and (11d) argue for phrasal covert movement.

⁷ We assume the vP internal subject hypothesis throughout.

per se on the functional head that checks them, rather the case feature of a DP is checked and deleted when valued by a set of full agreement features by a c-commanding Probe (v or T).⁸

The case feature of the DP is to be understood as literally [+structural case], underspecified for its value (Accusative or Nominative). Potentially, it can be realized as both depending on which Probe (functional head) it is checked against. It comes out as Accusative when valued by a complete v and Nominative when valued by a complete T.⁹ The interpretable [+φ] features of the DP match the uninterpretable [+φ] features of T. The syntactic relation relevant for Agree is c-command; the probe should c-command the goal within a minimal search space. The minimal search space is the c-command domain of the probe, which is the c-command domain of T (TP) in our concrete example (12).

Features are checked if they match and if the Probe and the Goal are active, that is they both still have unchecked uninterpretable features. We have thus provided an interpretation of the following definitions:

- (13) Matching and Agree (Chomsky 1998: 38):
 Matching is feature identity;
 D (P) is the sister of P;
 Locality reduces to 'closest c-command'.
- (14) Matching and Agree (Chomsky 1999: 4):
 Probe and Goal must both be active for Agree to apply;
 α must have a complete set of [+φ] features (it must be φ complete) to delete uninterpretable features of the paired matching element β.

Accusative case licensing is, in many ways, similar to Nominative case marking. It is performed in situ under case-agreement between the probe v and DP_{OB}:

- (15) [_{vP} V_[+φ] [_{VP} V DP_{OB[+φ, case]}]]

A transitive v is said to have a full set of [+φ] features, so the feature [+case] in the DP is deleted and the case is valued as Accusative. Any displacement of the object DP to the edge of vP is not attributed to the case agreement relation as such but rather to the (parametrically determined) presence or absence of the [+EPP] feature on v (or a functional head in general). The [+EPP] feature and the spec/head relation

⁸ Chomsky (1999: 4): "Structural case is not the feature of the Probes (T, v) but it deletes under agreement if the Probe is appropriate – φ complete [...]. Case itself is not matched but deletes under matching of [+φ] features."

⁹ Note that under such formulation of the feature constitution of the Probe and the Goal, structural case seems to be assigned rather than checked; the Case theory has thus turned a full circle.

in general, is now dissociated from the Case agreement relation, yet the [+EPP] feature can be satisfied only by constituents, which are visible to the head (Probe) with the [+EPP] feature. Because visible means still active, i.e. carrying an unchecked uninterpretable feature, the same head (Probe) is involved in both case agreement and forcing external merge.¹⁰ As for the timing of the case-agreement relation, valuation and elimination of uninterpretable features is supposed to take place as soon as a valid probe-goal configuration arises. In a sense then, feature checking is subject to Pesetsky's (1989) Earliness Principle rather than Chomsky's (1995) Procrastinate; as soon as features match, they should be maximally checked and deleted, subject to the following principle:

- (16) Maximize matching effects.¹¹

This principle seems an obvious and natural consequence of the cycle and the Extension condition on structure building but leads to intriguing consequences with defective probes, as we should see below.

The practical effect of (16) is that the head will force movement (copying and re-merging) of the same constituent that is the goal for Agree. Here is why: it is now impossible to have a case of two neighbouring functional heads (Probes), of which both have the [+EPP] property and cause non-cyclic movement. Let us consider a hypothetical incremental derivation in (17) with T and C and the [+EPP] feature of C is satisfied in violation of (16). Imagine a transitive construction with a Wh-object, in which T checks (values) structural case but does not force movement, (17a), and C is merged in, checks the [+wh] feature without movement, (17b), but forces the movement of the subject to form its specifier, (17c):¹²

- (17a) T [_{vP} NP_S ... Wh_{Ob}]
 ↑
 (17b) C [_{TP} T [_{vP} NP_S ... Wh_{Ob}]]
 ↑
 (17c) * [_{CP} NP_S C [_{TP} T [_{vP} {NP_S} ... Wh_{Ob}]]]
 ↑

¹⁰ By Case agreement we do not mean a necessarily successful case agreement. Defective v and T may have the [+EPP] feature and engage in agreement which does not result in case valuing.

¹¹ The following quote (Chomsky 1999: 12) defines maximal matching: "[...] If local (P,G) match and are active, their uninterpretable features must be eliminated at once, as fully as possible; there is no option of partial elimination of features under Match followed by the elimination of the residue under more remote Match. In particular, if probe P requires Move (i.e. has an EPP-feature), then the operation must be carried out as quickly as possible."

¹² At this point we abstract from the issue of locality, that is whether or not C can communicate with the Wh-phrase in its object position.

The end result in (17c) is ungrammatical but is avoided due to principle Maximize in (16); the relation between the Probe and the Goal should be exploited to the maximum, that is the [+EPP] feature of the Probe should be satisfied without any need to look for another Goal in the search domain.

Another significant aspect of the latest minimalist system is that the overt and the covert cycles in the derivation have collapsed into one, the single cycle system. In Chomsky (1995, ch. 4) a head carrying a strong feature was compared to a virus within the computational system: the virus had to be neutralized immediately through the checking of the strong feature and movement to create the specifier position. Weak features deserved a different treatment, as they could wait until the covert cycle and had to disappear before reaching the LF interface level, where, as uninterpretable, they would violate Full Interpretation. In the newest system every feature is a virus and must be checked immediately: there is only one cycle, LF and PF representations are built incrementally and backtracking is severely constrained. This solution calls for multiple application of operation Spell-out, which distinguishes between the material that can be sent to the LF-interpretive mechanisms (LF-interpretable features) and the material that can be sent to the PF interpretable mechanisms (PF-interpretable features and specified/valued uninterpretable features). We return to the issue of multiple Spell-out in further parts of the paper.

A new mode of checking and deleting features calls for a reformulation of the definition of intervention effects. These are the well-known cases of case-checked DPs that are closer to the feature checker than other case-less DPs. In the parlance of the case-agreement system, an intervener in an agreement relation is a DP, whose [+φ] features are interpretable, hence visible to the Probe, yet its uninterpretable case feature has been checked, which makes it an invalid Goal. It is then an invalid Goal, though a valid intervener (β) in the case-agreement relation between α and γ, (Chomsky 1998: 38-39):¹³

$$(18) \quad *α > β > γ$$

The issue of intervention effects is of primary importance and is related to the feature composition of β. We shall return to this issue as we proceed.

Note the crucial difference between this version of Minimalism and the previous ones: the spec/head relation is irrelevant for Agree (feature checking), the uninterpretable features on the Goal are checked and deleted in situ, without displacement. For instance, returning to example (12), the following uninterpretable features have been checked in situ without displacement: the φ features of T and the case feature of the subject DP. What remains to be checked is the feature [+EPP] on T. Chomsky

(2001) refers to this feature as [+OCC] for (unforced) occurrence of a DP. This feature states that T requires a specifier. As the (sub) array has been depleted, the only option left to satisfy this feature is movement of the subject DP from its thematic position in [spec,v] to [spec,T]:¹⁴

$$(19) \quad [TP \text{ Sub}_{[*\text{case}, +\varphi,]} T_{[*\varphi, *EPP]} [vP \{ \text{Sub}_{[*\text{case}, +\varphi]} \} v \text{ Ob}]]$$

Thus checking of case and agreement features is entirely divorced from displacement, the former is performed on the Goal in situ under c-command from the Probe within its minimal search space and the latter is forced independently by the [+EPP] feature of the Probe. At the same time the problem of complex specifiers and feature percolation dissolves; in example (2) the [+wh] feature is checked prior to the movement of the Wh-phrase into [spec,C], when C c-commands it in a local configuration.

This rudimentary sketch immediately begs several questions that have been omitted so far, as we have tried to illustrate a relatively uncontroversial case of the separation of Agree from Move and the separation of the checking theory from the spec/head relation.

Some of the initial questions are: how minimal is the minimal search space, so in the concrete case of (19), can T look past DP_{SUB} and become involved in Agree with the more distant DP_{OB}? The other question intimately related to the previous one is how long a given Goal is accessible to an active Probe.

In order to gain a better perspective on the system based on Agree and address the first question, we have to move one step back in the derivation to a stage preceding (19). At this stage a repeated application of (external) Merge has formed vP:

$$(20) \quad [vP \text{ DP}_{[+\varphi, +\text{case}]} v_{[+\varphi]} [vP \text{ V DP}_{[+\varphi, +\text{case}]}]]$$

The object DP has interpretable φ features and unchecked uninterpretable [+case] feature, and the Probe, v, has a complete set of unchecked and uninterpretable agreement features. In the configuration in (20), v can act as a Probe and the object DP as a Goal in the relation of Agree resulting in the valuing of the [+case] feature as Accusative and the checking of the [+φ] features of v. Again, the checking of features takes place in situ within the minimal search space of the Probe. A very significant point emerges from this picture of Agree and feature checking in the c-command domain of the Probe: the subject in its thematic position of [spec,v] is not placed in the minimal search space of the Probe. If so, it cannot establish the relation Agree with v. This effectively means that the concept of checking under c-command concludes the discussion of derivations in which the subject is erroneously assigned

¹³ Chomsky (1998: 38) comments on the Defective Intervention Constraint: "We therefore have the possibility of defective intervention constraints in structure [...], where > is c-command, β and γ match the probe α, but β is inactive, so that the effects of matching are blocked."

¹⁴ A clarifying note on terminology: Move is now called 'internal Merge' and simple Merge is called 'external Merge'.

Accusative and the object may be erroneously assigned Nominative, an unwelcome alternative in earlier minimalist work, where spec/head was a feature checking configuration.

The fate of the object DP in the derivation is further determined by the presence or lack of the [+EPP] feature on *v*. If there is no [+EPP] feature on *v*, the object DP remains in situ as a legitimate syntactic object whose uninterpretable features have been checked. In line with conclusions from previous work, a syntactic object with all of its uninterpretable features checked is unavailable for further Agree relations and 'frozen in place'. On the flip side of this statement, Agree can involve only the Probe and the Goal which are active through their unchecked uninterpretable features:

- (21) Probe and goal are active and available for relation Agree only when they still have some unchecked uninterpretable features.

Thus there is at least one reason for which the object DP could not be involved in the relation Agree as a Goal with T as the Probe; by the time the derivation reaches this stage, namely, the object DP has been rendered inactive.¹⁵

2.3. Agreement as residue of the specifier/head relation

There is, however, one outstanding issue that must be addressed at this point. If Agree, and chiefly *c*-command, is the relevant checking configuration, and the specifier/head relation is not relevant in itself, two theoretical and empirical consequences follow. First, every case of spec/head agreement must result from movement, because the specifier is not in the *c*-command domain of its head, and second, morphological agreement should be identical in both specifier/head and head/complement configurations.

We return to the first consequence below and as far as the second one is concerned, one would not expect the following differences, discussed among others in Bošković (1997), Hornstein and Witkoś (2003) and Witkoś (this volume). It is a well-known fact that if the whole DP moves to [spec,T], then it is the full DP that agrees with the [+ϕ] features of T. This pattern is not defective, yet agreement in non-spec head configurations can be defective:

- (22a) (?)There seems to be men in the garden.
 (22b) There is a dog and a cat on the roof.
 (22c) *Men seems to be in the garden.
 (22d) *A dog and a cat is on the roof.

¹⁵ There is one more reason concerning locality of Agree, to be discussed shortly.

These facts suggest that reducing all agreement to *c*-command, as is done in an Agree-based system, is problematic.

Chomsky (2001: 11) seems to be aware of these empirical problems and argues that the differences between the agreement patterns may be due to the difference between inner Merge and outer Merge; the associates in (22) are only merged and valued in situ, while the subject is merged and then copied and re-merged. Interestingly, this need not be a restatement of the critical observation above: specifiers have a status different from non-specifiers, thus whatever their status is it seems to result from a set of operations: Agree plus re-merge, while agreement as a reflex of pure Agree can be impoverished morphologically.

Consider the first consequence of abandoning of the spec/head configuration as a checking relationship. Incidentally, if Chomsky's comment is taken literally it will probably lead to substantial derivational complexity in simple cases of NP-internal agreement, a complexity that may nevertheless be independently required and is worth considering, if regular morphological agreement/concord is [+ϕ] feature sharing and subject to Agree. Take the following simple cases in Polish:

- (23a) pusty dom
 empty-3Sg.Msc house-3Sg.Msc
 'an empty house'
 (23b) puste domy
 empty-3Pl.Msc houses-3Pl.Msc
 'empty houses'

The modifier always fully agrees with the head noun. For the simplicity of argument let us, initially, adopt Kayne's (1994) principles, so the modifier does in fact occupy some specifier position. But the specifier cannot agree with the head upon first (external) Merge, as full morphological agreement may be an interesting reflex of Agree and re-merge (Move). This implies that a correct representation of (23) may not involve a simple structure as in (24a) but a more complex one, as in (24b):¹⁶

- (24a) [NP [AP pusty] dom]
 (24b) [NP [AP pusty] [N' dom {AP pusty}]]

The syntactic object in (24b) is formed through outer merge of the nominal head and the modifier in which the nominal projects further and the modifier is copied and re-merged in the specifier position of NP. Interestingly, there may be some independent motivation for this derivation apart from our feeble attempt at defending Chomsky's comment on the insignificance of specifier/head relations.

¹⁶ This observation is based on the discussion of LCA and BPS in Lasnik and Uriagereka (2001).

The first piece of independent motivation comes from the rationale of Bare Phrase Structure (BPS, Chomsky 1994) and its rendition of the Linear Correspondence Axiom (LCA). In the BPS system representation (24a) cannot be linearized, as the AP and the nominal are in a mutual c-command relation and empty projections and bar levels are disallowed. One of the ways of breaking the symmetry of this representation is to create, through movement, a trace/copy of the AP that can be neglected by the LCA; copies are not pronounced and are exempt from the LCA considerations. The (seemingly complex) representation in (24b) meets both: the stringent requirements of LCA and the postulate that full agreement is a reflex of inner Merge (Move).

The second piece of evidence is conceptual, shows that the nominal in (24) must at some point c-command the modifier, whose movement becomes more plausible. It is inspired by Chomsky's (2001) discussion of modifiers as constituents built on a separate plane of the derivation and imported onto the main plane at some point, inspired in turn by Uriagereka (1999). We discuss this idea in more detail in Section 7.

Let us abandon Kayne's perspective now and assume that modifiers are not in the specifier position of the head that they modify but are initially adjoined to it on a separate plane. What is relevant now is the assumption that the separate plane of derivation is invisible to the c-command relation. The argument is as follows: how does Agree work? It needs to involve a Probe and a Goal in a c-command relation. But the modifier attached on a separate plane cannot be seen by the Probe as long as it is not c-commanded. Thus in the derivation of (24) must involve a point where the modifier is imported into the main derivational plane and c-commanded by the Probe (the nominal). From this position it moves up. This is the only logical way in which modifiers can agree if such ideas of adjunction are taken seriously.

To sum up this section, relation Agree as a checking configuration seems to solve many old problems caused by the specifier/head configuration and the head/head configuration. The primitive relation of c-command, on which Agree rests, is an irreducible element of structure building. Agree seems to create problems of its own, by neglecting the unique role of the specifier/head relation for feature agreement, but there seems to be some hope that these complications are only apparent and may receive a principled treatment in the spirit of Bare Phrase Structure.

3. Motivating derivational phases

Chomsky proposes that the process of computation proceeds in stages comprising a minimum number of Lexical Items (a subarray) consisting of a lexical category and a functional category (CFC).¹⁷ The number of LIs in a subarray is sufficient to build

¹⁷ Chomsky considers only the so-called Core Functional Categories (CFCs) such as the Complementizer (Comp) and the light verb (v). In the process of his analysis it turns out that T should be treated as a

vP, the verbal predicate phrase licensing all the arguments and CP, the clause augmenting the argument structure with the tense, thematic prominence and force specifications. Once the process of construction of a phase is completed, or at the latest by the time the next higher phase is completed, the phase undergoes operation Transfer, including Spell-out, feeding PF and LF. Thus the memory load of the computation itself is lessened and once a given phase is completed, the computational machinery begins to construct another phase.¹⁸

Chomsky (1998) considers the following examples as an argument for the claim that the computation should be fed by a lexical array smaller than a sentence:

- (25a) there are questions about [_α what C [_{TP} John read t]]
- (25b) there is a possibility [_α that proofs will be discovered t]
- (25c) T is likely [_α there to be a proof discovered t]
- (25d) It's fun [_α PRO to [t go to the beach]]
- (25e) It was decided [_α PRO to be executed t at dawn]

His argument rests on the assumption that if Merge is cheaper than Move, as typically assumed (Chomsky 1995, Hornstein 2001), there would be no way of deriving most examples in (25) starting from one big Lexical Array, as a merger of the expletive should always preempt the more complex and costly Move. All derivations should behave the same as the derivation in (25c): as soon as there is an expletive in the Lexical Array, the argument should be merged only in the thematic position, otherwise the theta criterion is violated, but it should never move, with its movement pre-empted by a less costly merger of *there* or *it*. Yet, in all the examples above but (25c), arguments move overtly to their case positions and only then, on higher cycles, expletive are used. These derivations require a 'look-ahead' property; early merger of the expletive blocks raising of arguments to case positions, detectable at a later derivational step.

Reduction of the Lexical Array and forcing the derivation to run in smaller cycles (phases) seems an attractive solution to the problem at hand; if all the constituents marked α in (25) above are constructed from a separate lexical sub-array, either containing or lacking the expletive, the issue of 'look-ahead' is avoided.

Chomsky considers two basic ways of defining phases: one as the minimal converging syntactic object and the other as a strict, category based concept. Following a brief discussion of example (26) he selects the latter option:

substantive rather than functional category. He does not rule out the possibility that there should be other functional categories such as Determiner (D) or elements from the articulated left periphery of the clause in the sense of Rizzi (1996) or Cinque (1999). If such functional categories are admitted, one should wonder as to what phase status they should have.

¹⁸ Certainly several phases can be built in parallel and the syntactic objects thus constructed can be pasted together at the final stage.

(26) [_{CP} which article is there some hope [_α that John will read t]]

He notices that for Wh-movement to be a feasible operation within the probe-goal system, the Wh-phrase needs to contain an uninterpretable feature to render it active.¹⁹ Now, this uninterpretable feature is matched, checked and deleted only at the point of deriving the entire syntactic object, matrix CP. At the same time, though, the numeration includes expletive *there*, which can be ambiguously merged either at the embedded [spec,T] or at the matrix [spec,T]; again ‘look-ahead’ is required to complete the derivation. Hence, the choice of the least convergent syntactic object for a phase does not avoid the problem of ‘look-ahead’, which can be sidestepped if α is a separate phase.

Chomsky (1998: 20) motivates his choice of vP and CP as (strong) phases by saying that at the meaning side they should be ‘propositional’, i.e. a verb phrase in which all theta roles are assigned or a full clause including tense and force. Specifically, CP and transitive vP are (strong) phases, while finite TP or passive unaccusative verbal phrases are not.²⁰

Technically, once a given phase is completed, any access to it from the outside is severely limited. This property is known as the Phase Impenetrability Condition (PIC):

(27) Phase Impenetrability Condition (Chomsky 1999:10)

The domain of H is not accessible to operations at ZP, but only H and its edge.

Where: ZP is the least strong phase and the edge consists of specifiers.

PIC reflects the concept of the cycle from previous systems and forces local, phase internal checking and elimination of features.

Let us pause briefly to consider the formulation of PIC. It is clear that the condition enforces subjacency and cyclicity of movement from edge of one strong phase to another. Observe, however, that the PIC does not preclude transparency of the vP

¹⁹ This is the [+wh] feature.

²⁰ Chomsky (1998: 20): “The next step is to characterize the subarrays LA_i that can be selected for active memory. LA_i should determine a natural syntactic object SO, an object that is relatively independent in terms of interface properties. On the ‘meaning side’, perhaps the simplest and most principled choice is to take SO to be the closest syntactic counterpart to a proposition: either a verb phrase in which all theta roles are assigned or a full clause including tense and force. Call these objects ‘propositional’. Considerations on the ‘sound side’ support the choice, given properties of the kind mentioned earlier distinguishing CP from TP, which extend to vP (fronting, extraposition, pseudoclefting, response fragments, etc.). [...] Take a phase of a derivation to be a syntactic object SO derived in this way by choice of LA_i. A phase is CP or vP, but not TP or a verbal phrase headed by H lacking φ-features and therefore not entering into Case/agreement checking: neither finite TP nor unaccusative/passive verbal phrase is a phase.”

strong phase before the CP is constructed in the derivation. Consider the following example from Icelandic (modeled on Holmberg and Hroarsdottir’s 2003: 999):

(28a) Henni likuðu hestar.
her_{DAT} liked horses_{NOM}
‘She liked the horses.’

(28b) [_{TP} NP_S T [_{vP} {NP_S} v [_{vP} V NP_{OB}]]]

This a construction with a quirky case marked subject (Dative) and a Nominative case-marked object.²¹ The quirky subject has been raised to satisfy the [+EPP] feature of T. The only source of structural Nominative is T and the definition of PIC in (27) does not prohibit the relation Agree between T and a vP-internal element in a non-edge position; it is as if the strong vP phase needed some time within the derivation to congeal into an impenetrable domain. At the point of the construction of TP, vP-internal elements are accessible but as soon as C is merged in, only the edge of vP is still accessible to further operations.

4. Defective Probes, Incomplete Goals, Raising and Long-Distance Agreement

The Probe and the Goal can be either complete or incomplete and full and defective respectively. Consider a derivation, in which the four types are present. Chomsky (1999: 13-14) discusses the following example in great detail:

(29) [_β There seem to have been [_α caught several fish]].

The derivation involves two cycles, corresponding to the participial phrase and the matrix TP. On the first cycle the participle and the object become involved in the probe-goal relation, manifested in many languages through morphological agreement, sometimes including agreement for case. The participle has an incomplete set of [+φ] features: {number, gender and case}, but not [+person]. Although the participle c-commands the object, a valuation of the case on DP_{OB} cannot take place, as the participial probe is incomplete. On the other hand though, DP_{OB} has a full set of (interpretable) [+φ] features, so the incomplete [+φ] features of the participle are matched, valued and deleted. By the end of cycle α DP_{OB} is still an active goal, as its case feature has not been valued and deleted. Technically speaking, the participle should be inactive, although its [+case] feature has not been valued. Why? Because

²¹ Evidence in the form of coordination, anaphoric binding and control produced for example in Zaenen et al. (1985), shows that the Dative-marked nominal indeed functions as the subject.

only unchecked uninterpretable [+ ϕ] features render probes active.²² At stage β , a complete probe T with a full set of [+ ϕ] features engages in two case-agreement relations: with DP_{OB} and the participle. The former relation is fully legitimate: the probe and the goal are within the same strong phase, the probe c-commands the goal, the probe is complete and the goal is still active. How is the probe-goal relation between T and the participle possible? Chomsky claims that although valued and deleted, [+ ϕ] features of the participle are eliminated from the computation only at the strong phase level (matrix CP), as part of operation Transfer to PF. So at stage β they are still visible to the T probe and render the participle active as a goal of the case-agreement relation.²³ The most significant aspect of the analysis of (29) is that a syntactic object with its [+ ϕ] features valued and deleted in one case-agreement relation can be resurrected and still active in the derivation as long as the second case-agreement relation takes place within the same strong phase.

Any non-local relation of case agreement raises the issue of intervention effects. In the concrete example at hand, does the participle (an incomplete probe) screen the DP object from T (a complete probe)? The answer is no, as the intervention effect is induced only by complete probes and complete goals, the incomplete ones do not value case and allow for maximal matching effects. One illustration of this phenomenon is given in (29) above and another is provided by a simple expletive construction below:

(30) We expect [there to arrive a man].

In (30) the DP argument of the embedded unaccusative verb is said to be assigned Accusative case by matrix v. The expletive *there* is assumed to have only the feature [+person] and it does not value the set of complete [+ ϕ] features of matrix v, so they are able to probe over the (chain of the) expletive within the same phase and match and value the case feature of the object. In the process, all uninterpretable features are checked and marked for deletion (the [+ ϕ] features of v, the [+person] feature of the expletive and the [+case] feature of the object DP).²⁴

²² Chomsky (1999: 4): "For the case-agreement systems, the uninterpretable features are [+ ϕ] features for the probe [verbal elements, including participles, J.W.] and structural case of the goal N. [+ ϕ] features of N are interpretable; hence N is active only when it has structural case."

²³ It is interesting that features that are valued and deleted do not disappear from narrow syntax at once. One wonders whether deletion of valued features is necessary if it is not tantamount to their disappearance. It seems that Chomsky implicitly uses the notion of features checked and 'marked for deletion' proposed in Pesetsky and Torrego (2002).

²⁴ The expletive *there* has feature [+person], which is uninterpretable on the expletive. This assumption is necessary to account for the fact that *there* can be moved and to be moved it must be 'active', that is have an uninterpretable feature:

- (i) T seems [there_[+person] to be [a man in the garden]]
- (ii) there_[+person] T seems [t to be [a man in the garden]]

5. Long-distance Movement

Let us consider the following simple Wh construction:

(31) Which book will Betty read?

This construction involves movement across the phase boundary, so the relevant issue of phase accessibility comes into play.

At an early stage in the derivation the following vP has been formed:

(32) [_{vP} Betty_[+ ϕ ,+case] V_[+ ϕ] [_{VP} V which book_[+ ϕ ,+case]]]

In this configuration, the relation Agree can be established between v (Probe) and *which book* (Goal). They both match in ϕ features and because v is a complete Probe, it values the [+case] feature of *which book* as Accusative. At this point *which book* should be frozen in place if it were not for the [+wh] feature on it. Let us follow Chomsky (1998, 1999, 2001) and make a crucial assumption now; v has an [+EPP] feature that must be satisfied before the derivation proceeds to the next stage. The feature [+EPP] forces v to have an outer specifier.²⁵

In order to force the creation of the outer specifier through Move (inner Merge), Chomsky (1998: 23) claims that the [+EPP] feature may be added to head H, here v, once the lexical sub-array needed for the construction of the (vP) phase has been depleted. With operations of external Merge and feature placement following in this order, the first Merge of the DP subject in its thematic position in [spec,v] will not be sufficient to satisfy the [+EPP] feature. Once the [+EPP] feature finds itself on v, it forces pied piping, copying and Merge of *which book*:

(33) [_{vP} which book_[+ ϕ ,*case,+wh] [_{vP} Betty_[+ ϕ ,+case] V_[* ϕ ,*EPP] [_{VP} V {which book_[+ ϕ ,*case,+wh]}]]]

Consider this operation of displacement from the point of view of Greed; neither the Wh-phrase nor v* benefits from this operation, all their features have been checked in situ.²⁶ Yet it is crucial for the derivation, as it places the Wh-phrase at the edge of a strong phase and renders it susceptible to further movement.

At the next stage T is merged and we have the following syntactic object:

²⁵ Chomsky (2001) leaves some leeway as to whether the specifier formed through inner Merge is the inner or the outer one. In his earlier work he would have assumed the latter but here he allows for the 'tucking-in' external Merge in the spirit of Richards (1999).

²⁶ Certainly, the operation trivially satisfies Greed, as it results in feature checking but the feature checked ([+EPP]) was just introduced into the system to cause the movement.

- (34) [TP T_[+φ,+EPP] [vP which book_[+φ,*case,+wh] [vP Betty_[+φ,+case] V_[*φ,*EPP] [vP V {which book_[+φ,*case,+wh] }]]]

T is an active Probe and has the subject DP in its search space. Their [+φ] features match, check and delete and the [+EPP] feature on T forces displacement of the subject. Finally, C⁰ is merged with TP:

- (35) [CP which book_[+φ,*case,*wh] C [TP T_[+φ,+EPP] [vP {which book_[+φ,*case,+wh] [vP Betty_[+φ,+case] V_[*φ,*EPP] [vP V {which book_[+φ,*case,+wh] }]]]

C has both the [+wh] feature and the [+EPP] feature. It matches, checks and deletes the [+wh] feature of the Wh-phrase and forces its displacement to form [spec,C].

The movement in our example has crossed only one strong phase boundary and it was possible only because of the optional [+EPP] feature on v*, without it the PIC would have precluded any movement from within vP. As in some other previous systems devised by Chomsky, most notably *Barriers*, the edge of vP is the escape hatch.

It is easy to predict what happens if the movement is really long distance:

- (36) Which book did Susan [vP t⁴ say [CP t³ [John had [vP t² lost t¹]]]]]

In example (36) we indicate all intermediate copies/traces. The position marked with t³ is required by the regular [+EPP] feature of C, whereas both t² and t⁴ are forced by optional [+EPP] features on both the embedded and the main v*. On every single occasion, [+EPP] features place the Wh-phrase at the edge of the strong phase and let it pass through in accordance with the PIC.

6. Demise of Equidistance

In earlier Minimalist work the issue of crossing movements from and into argument positions gained considerable prominence. The strongly empirically motivated principles of Shortest Move (SM) and Minimal Link Condition (MLC) required an exception in the form of the principle of Equidistance to allow for crossing, rather than nesting dependencies. These issues resurface in the phase-based Minimalism in an interesting context and lead to the conclusion that the significance of Equidistance could be only illusory and its status fairly dubious.

Recall that defective Intervention arises when an element whose features have already been checked nevertheless interferes with another, longer, Probe-Goal relation. Intervention effects are more serious if the intervener is [+φ] complete.

The following examples seem to present the full paradigm of intervention effects:

- (37a) Me (DAT) thought (Pl.) [t_{me} [they (Pl.NOM) be industrious]]
 (37b) *Me (DAT) seem (Pl.) [t_{me} [John (DAT) to like horses (Pl.NOM)]]
 (37c) *John seems (Sing.) me (DAT) [t_{John} to like horses]
 (38a) (guess) what_{OB} [John_{SUB} T [vP t_{OB} [t_{SUB} read t_{OB}]]]
 (38b) John_{SUB} T [vP that_{OB} [t_{SUB} read t_{OB}]]
 (38c) [TP ... T... [vP DP_{OB} [vP DP_{SU} v ...]]]

The first three come from (Chomsky 1998: 47) and represent Icelandic constructions with quirky subjects, requiring structural case licensing as well. The latter two come from (Chomsky 1999: 21). The DIC effects show in (37b) because a DP, with its [+case] valued (by matrix T) and frozen in place but with its [+φ] interpretable features visible to the Probe, renders a further Goal (*horses*) invisible to the Probe. Example (37c) illustrates how locality works; quirky Dative with structural case is more local than a further Goal (*they*) and should qualify as a target for the [+EPP] feature of matrix T.²⁷ Example (37a) is the most interesting case from our perspective; in contrast to (37c), the Dative-marked pronoun has been raised to the matrix subject position leaving a trace (copy) behind. Interestingly, this trace (copy) is invisible to the Probe-Goal relation, here between matrix T and *they*.

Examples in (38) have further bearing on the issues of DIC effects. While the principle of Equidistance is still a valid way of avoiding Minimal Link Condition violations in Chomsky (1998: 38), it is abandoned in Chomsky (1999) in favour of a solution based on displacement of the intervener. Having discussed the status of the Thematization/Extraction (TH/EX) rule in some English unaccusative constructions²⁸, Chomsky considers the issue of the difference between the English type languages (38a) and the Icelandic type languages (38b) with respect to Object Shift and the intermediate configuration that arises in the process (38c).

The question is why the English-type languages allow for Object Shift only when the object DP vacates the outer specifier of vP, whereas Icelandic-type languages seem to allow for Object Shift with the object freezing in the position of the

²⁷ Chomsky (1998: 41):

EPP can be satisfied by:

- (i) Merge of expletive;
- (ii) Merge of associate;
- (iii) Merge of α closer to T than the associate.

It seems that the first two options are universally attested, while option (iii) is available in languages, where DPs bearing quirky case also require structural case (cf. Icelandic).

²⁸ TH/EX is a language-specific PF rule that does not feed any other movement rules in narrow syntax and is responsible for the following displacements to the left or to the right of an unaccusative VP in English:

- (i) there was [vP arrested [a man]] → there was [vP [a man] [vP arrested t]]
- (ii) there [vP came [a man from Idaho] into the room] → there [vP [vP came into the room] [a man from Idaho]]

outer specifier of *vP*. The issue is closely related to the notion of defective intervention; DP_{OB} in example (38) has its case feature checked and deleted and as such is not active for another case-agreement relation with Probe T but its interpretable and visible [+ ϕ] features can potentially cause an intervention effect between T and DP_{SU} . Both in (Chomsky 1995) and (Chomsky 1998), the answer was linked to the strength of T as an Attractor or a Probe; in Icelandic both the object and the subject were equidistant from Probe T and its agreement morphology made it strong enough to bypass the intervening DP_{OB} and establish a matching and checking relation with the subject. In contrast, T in the English-type language was a weaker Attractor/Probe and could not bypass the outer [spec,v], even though the Equidistance principle allowed for this.

Chomsky (1999: 24) proposes a different track: there is no difference between Icelandic and English in this respect, as neither can invoke Equidistance to relate the subject to probe T.²⁹ The principle of Equidistance is redundant if Icelandic has an equivalent of TH/EX, a PF driven operation of Dislocation, which moves DP_{OB} away from the outer edge of *vP* and thus there is only a trace of the object separating T from DP_{SU} .³⁰ And, crucially, traces do not induce intervention effects.

In sum, the study of these cases leads to the following generalization Chomsky (1999: 13):

²⁹ For the case of Icelandic Object Shift, which fuelled so much discussion of Equidistance, Chomsky (1999: 28) proposes the following analysis:

- (i) v^* is assigned an EPP feature only if that has an effect on outcome;
- (ii) the EPP position of v^* is assigned INT;
- (iii) the phonological border of v^*P , XP is assigned INT'.

The key element of this account is the idea that particular aspects of the semantic interpretation, such as definiteness/specificity (here INT) are read off the head of the argument chain of the object in [spec,v], while the tail position of the argument chain of the object receives the interpretation of indefinite/nonspecific only when the object has remained the sole lexical element within *vP* (at its phonological border; a term is at the phonological border of a phase HP if it is not c-commanded by phonological material within HP).

³⁰ To make the system work, Chomsky must face the case of clear countercyclicity: probe T has to access goal DP_{SU} in the inner [spec,v] prior to the actual displacement of DP_{OB} to [spec,C]. The reverse order would apparently violate the Extension Principle. The solution lies in fully exploiting the mixed derivational-representational nature of Chomsky's recent system; MLC scans the syntactic object constructed thus far for any violations of minimality only the point of Transfer, at the strong phase level. The relevant strong phase level comes at CP and the MLC is violated at the lower 'weak phase' TP level. Inner Merge at [spec,C] rescues the violation at the relevant point.

- (39) DIC Effects:
Only the head of the A-chain (equivalently the whole chain) blocks matching under the Minimal Link Condition. Traces do not interfere with matching of other Probes and Goals.³¹

7. Wh movement, Principle C and Reconstruction revisited

The assumption that the phase is a cyclic unit for both: computation in narrow syntax, multiple spell-out and semantic interpretation (SEM), works best if computation strictly observes the Extension Principle. Chomsky (1998: 53) derives the effect of this principle from the following postulate, frequently referred to as 'minimal tampering':

- (40) Given a choice of operations applying to α and projecting its label L, select one that preserves the R (L, Γ).

What is meant by R (L, Γ) is basic relations involving the label that projects, chiefly sisterhood and c-command. Let us illustrate the application of this principle with Chomsky's example. Assume that there are two syntactic objects β and α and they are to be merged in such a way that the label L of α remains unchanged and α projects.³²

³¹ The discussion of this issue in Chomsky (1999: 22-23) is slightly confusing, as he proposes the following conclusion:

- (i) EC (empty category) disallows pied-piping;
- (ii) Inactive trace disallows Match.

From his discussion of a preceding example it follows that what he means by (ii) is that the trace cannot be matched by a c-commanding probe, and matching in features is a precondition to case-agreement, movement or intervention.

"In the structure (iii), XP prevents Match of a probe P and SPEC, under PLC, only if XP has phonological content:

- (iii) [_{ZP} ... P ... [_{HP} XP [SPEC [H YP]]]]

[...] If XP with phonological content remains at the edge of (iii), then it intervenes to prevent Match (P, SPEC), hence Agree or Move of SPEC. If XP moves on to SPEC-Z, its trace is inactive and is invisible to Match by (ii), so SPEC, now at the phonological edge, is accessible to the probe P."

Chomsky ends by leaving the following related question open for further research: do traces (copies) free both Move and Agree to the same extent?

It would be surprising if they did not, as even if the subject Moves in (38c), repeated here as (iv) for convenience, case-agreement should take place with the subject in [spec,v], its base position within the search space of Probe T, across the trace (copy) of the object:

- (iv) [_{TP} ... T... [_{VP} DP_{OB} [_{VP} DP_{SU} v ...]]]

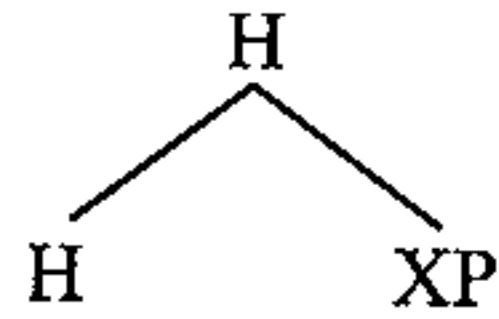
The displacement of the subject to [spec,T], due to the [+EPP] feature on T, has no bearing on valuing of Nominative its case, especially in view of the comments on the irrelevant status of the spec/head relation in Chomsky (2001).

³² The operation in question could be either pure Merge or second Move (Copy + Merge) of β .

Assume that α is a complex object of the following type and internal composition:

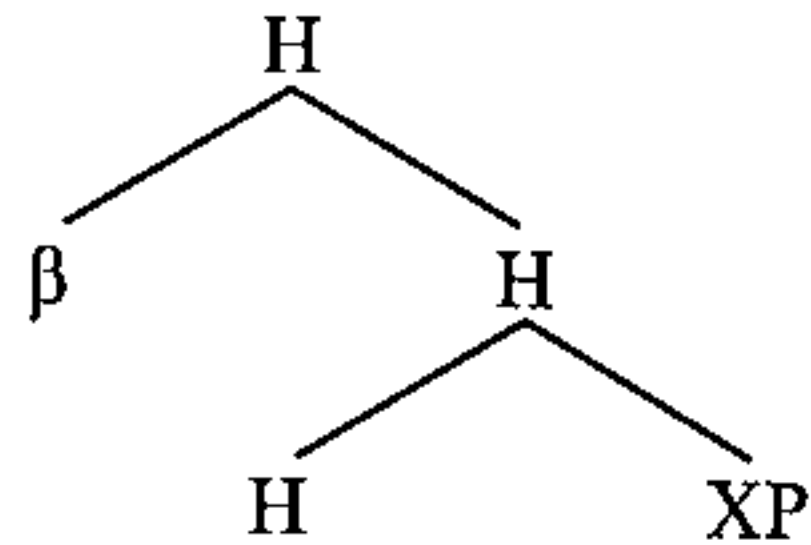
(41a) $\alpha = \{H, XP\}$

(41b)

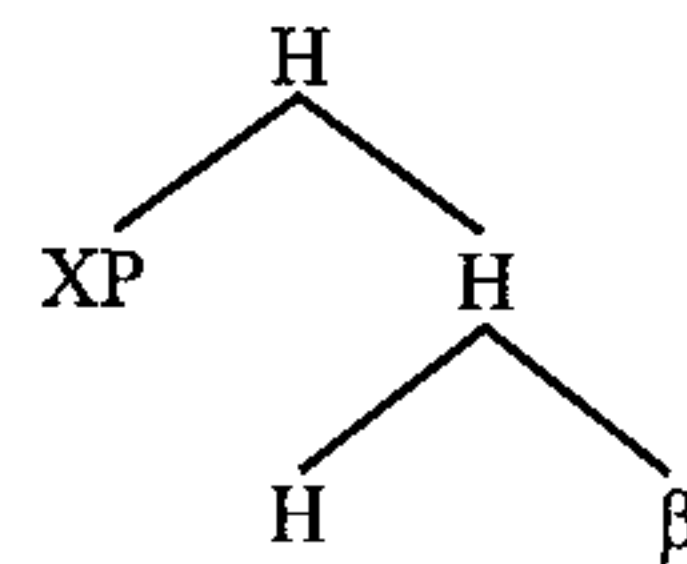


Logically speaking, there are two possible strategies of the merger of β and α in (41), both meeting plausible economy conditions. One strategy leaves the object α unchanged and β becomes the specifier of α , (42a). The other strategy involves merger of β as close to the label L as possible, (42b):

(42a)



(42b)



Trying to evaluate both strategies of merger with reference to the principle of minimal tampering, it is clear that only one, strategy (42a), consonant with the Extension Principle, obeys the principle in (40), as both before Merge in (41) and after it in (42a), XP is the sister of H and constitutes its c-command domain. The strategy of closest merge in (42b) non-trivially distorts the basic relations involving the label; in the wake of this operation XP is no longer a sister to H and is not placed in its c-command domain. Hence, Merge based on (42a) is to be preferred. The conclusion is the derivational extension of the existing tree at the root is a minimally distorting, thus a more optimal, operation.

Chomsky observes that the difference between the two strategies (extend vs. merge as close to the label as you can) can be meaningfully resolved only in the case where β is to be one specifier; as we could see above, β must be a specifier rather than a new complement of H. Both strategies appear possible when β is to become the second (or further) specifier of H. Consider the following case, where A is to be merged with the object derived through (42a):

(43a) $\{A, \{\beta, \{H, XP\}\}\}$

(43b) $\{\beta, \{A, \{H, XP\}\}\}$

Again, option (43a) blindly follows the Extension Principle; A is merged as the outer specifier of H and it certainly does not disturb the relation of sisterhood and c-command between H and XP. But now the option in (43b) is compatible with the principle of minimal tampering as well, because this principle refers only to the basic relations of the head/label of α based on c-command and sisterhood but does not make any reference to any other relations within α , chiefly the order or closeness of multiple specifiers to the head/label. Thus the merger of the outer specifier in option (43b) non-trivially affects the c-command properties of β but it does not affect the basic relations involving H.³³ This discussion legitimizes Richard's 'tucking-in'.

In a sense there is a clear message that the principle of minimal tampering sends through the realm of the minimalist enterprise: any non-cyclic Merge to a term properly contained in a larger syntactic object is a complicating factor and should ideally be excluded from the optimal design. This is, however, exactly the type of operation that was, following Lebaux (1988), assumed for adjuncts in Chomsky (1995, ch. 3) to account for the contrast between Principle C violations in Wh movement of arguments and lack thereof in Wh movement of adjuncts:

(44a) *Which claim that John₁ was asleep did he₁ deny {x claim [that John₁ was asleep]}?

(44b) [Which claim [that John₁ made]] did he₁ deny {x claim}?

To revise briefly, example (44a) feeds principle C, as the copy of the Wh phrase contains a name c-commanded and coindexed with *he*₁. In example (44b), late Merge of the relative clause (an adjunct) to the Wh phrase following its movement to [spec,C] allows for the bleeding of Principle C. The late Merge of the adjunct was justified at the stage of the Agr-based Minimalism through the assumption that arguments had to be merged cyclically observing the Extension principle, as they received their thematic roles through first Merge, and adjuncts need not. The Extension Principle was said to hold of Merge and overt substitution but not adjunction (either XP or X⁰).

This residue of non-cyclicity needs to be disposed of if strictly cyclic operations are the most optimal solution to the issue of derivational consistency. Note that head (X⁰) adjunction has been removed from the purview of the narrow syntactic derivation by the assumption that head movement is a PF phenomenon and PF is a component unrestricted by purely syntactic economy and convergence conditions.³⁴

³³ This is the difference between the strictly derivational definition of c-command and extension, as proposed in Epstein and Seely (2001) and Chomsky's mixed representational-derivational system. In the former approach 'tucking-in' is more problematic.

³⁴ In Chomsky (1995, ch. 4), the rule of WI, word internal interpretation of heads adjoined to other heads, exempts this operation from the purview of the principle of minimal tampering, whose application should force movement of VP to form the specifier of TP, rather than movement of V to T in relevant cases.

Impressed with the naturalness and conceptual plausibility of the principle of minimal tampering, Chomsky (2001) proposes to dispense with late adjunction for XPs as well. The basic idea of this proposal is that the adjunct is merged with the term it modifies cyclically but to some point it can remain invisible to the principles of narrow syntax and has input only on the semantic interpretation (SEM). Instead of an optional late merge of adjuncts there is obligatory early merge of the adjunct in a guise invisible to narrow syntax and optional conversion of the adjunct structure to a visible guise.

Chomsky argues that in a configuration of adjunction, as in (44b), the DP *which claim* behaves in some respects as if it was not modified by the adjunct at all; for example it can be selected by the verb like an unmodified DP, it can receive a theta role in a regular manner unobstructed by the presence of the adjunct and its label is still [+D], despite the clause attached to it, rather than some hybrid [+D/C] label or plain C.³⁵ He proposes that adjunction results from a specific type of Merge and distinguishes between set Merge (simple merge) producing sets, and pair Merge, producing adjunction structures:

(45a) set Merge: $\{\alpha, \beta\}$

(45b) pair Merge: $\langle \alpha, \beta \rangle$

The regular simple merge produces structures whose label is either α or β , but pair merge produces structures whose head is, say β , both before and after merge. Within narrow syntax β behaves as if α were not adjoined to it at all.³⁶

Chomsky tries to motivate the very existence of adjunction through properties of structure conditioned semantic interpretation (SEM). The interpretive phenomenon of modification cannot be expressed either by simple merge, which corresponds only to the pure predicate argument relations or by movement and edge properties, which correspond to surface scope and the general topic/focus plane. Modification is mapped to adjunction structure in syntax.

Elaborating on the details of adjunction in the base, let us consider the following case:

(46) $[\alpha [\alpha \text{ which picture of Bill}_2] [\beta \text{ that John}_1 \text{ liked}]] \text{ did he}_{1/*2} \text{ buy } t_{wh}?$

First, cyclically, β is adjoined to α through pair merge, which, for the purposes of narrow syntax does not change anything on the label of α , with β on separate plane. α is then set merged with V to produce a VP and receives a theta role from it. The

³⁵ Note that the prohibition on adjunction to arguments, frequently stressed before, is now abandoned.

³⁶ "Given the basic properties of adjunction, we might intuitively think of α as attached to β on a separate plane, with β retaining all its properties on the 'primary' plane, the simple structure" (Chomsky 2001: 15).

fact that it has β attached to it does not disturb these relations. With β on a separate plane, *he* c-commands only (the copy of) α but it does not c-command β .³⁷

This cannot be the full picture of adjunction on a 'separate plane', as the adjunct internal material can, after all, be c-commanded by terms of a more complex syntactic object. Here is a simple illustration of this fact:

(47) he asked $[\alpha [\alpha \text{ which picture of Bill}] [\beta \text{ that John liked}]] \text{ did he buy } t_{wh}?$

In example (47) matrix *he* cannot be coindexed with *John*, otherwise a Principle C effect arises.³⁸ Therefore, at some point in the derivation of example (47) the adjunct must be imported from its separate plane to the regular plane of the narrow syntactic derivation, where it is visible to c-command like other terms merged through simple set merge. Chomsky claims that there is operation SIMPL, relevant for SEM and PF processes, which is fed by adjunction structures and returns set merge structures:

(48) SIMPL: $\langle \alpha, \beta \rangle \longrightarrow \{\alpha, \beta\}$

As (47) shows, SIMPL applies at the latest at the stage of the derivation where the $\langle \alpha, \beta \rangle$ is spelled out. Chomsky assumes that SIMPL is a part of operation TRANSFER feeding both PF and LF operations.

Empirical properties of Wh movement require that the point of application of SIMPL be optional. This operation can apply either before or after Move (or in the jargon of copies, either to the copy at the edge or to the copy in the phase internal position). Here is why: for all intents and purposes, SIMPL plays the role of early/late Merge of the previous theory and it reveals the material from the separate plane to the punitive power of binding conditions holding of the cyclic derivation in the interpretive component (SEM). If SIMPL optionally applies to the copy of t_{wh} in (46), Reconstruction effects with the adjunct arise (the embedded *he* and *John* could not be coindexed).³⁹ At the latest, SIMPL applies at the edge of the phase as part of TRANSFER, so it targets the copy at the edge and consequently, the embedded *he* can be coindexed with *John* but the matrix one in (47) cannot, as it c-commands a spelled-out and SIMPLified syntactic object where both the complement and the adjunct to *picture* are on the same plane visible to narrow syntax and binding conditions.

³⁷ Chomsky (2001: 15) confirms his view that in A'-chains the reconstruction of the restriction and application of the Preference Principle is expected: "The effect (of Condition C, J.W.) for (*he*, *Bill*) is expected by (obligatory) reconstruction, but not its obviation for (*he*, *John*)."

³⁸ Chomsky (2001: 16) comments on example (47) by saying that "[...] *Bill* is subject to Condition C and hence must be c-commanded by matrix *he*". He probably means *John* rather than *Bill* as only this coindexation is relevant to the discussion of adjuncts.

³⁹ This is how Chomsky's cyclic merge of adjuncts followed by optional SIMPL capture the fact that in some constructions adjuncts cause obligatory Reconstruction effects, for example Heycock (1995).

It seems that SIMPL integrates the adjunct (α) into linear structure at the point of spell-out of the constituent (β) that the adjunct modifies. Chomsky (2001:17) captures this observation through the following general principle:

- (49) In adjunction structure $\langle \alpha, \beta \rangle$, the adjunct (α) is spelled out where the modified category (β) is.⁴⁰

As SIMPL is a part of TRANSFER leading to PF spell-out and SEM, names embedded in adjuncts to constituents interpreted in situ will always feed condition C, as the adjuncts are converted to simple merge linear structures:

- (50a) [α [α which book]][β with John's pictures]] did he find t_{wh} in a bookshop?
 (50b) Where did he find [α [α which book]][β with John's pictures]] t_{wh} ?

In this account of adjunction structures, the traditional difference in anti-Reconstruction possibilities between overt movement pied-piping adjuncts (50a) and covert movement, is the point of application of SIMPL; while in (50b) it must unambiguously apply to the Wh phrase in situ, in (50a) it can apply either to β within the copy or to β at the edge. In the former case condition C is fed, while in the latter it is bled. Note that the point of application of SIMPL before or after movement has consequences only for the interpretive mechanism SEM, where binding conditions are said to apply, but not for PF. If SIMPL applies prior to movement in (50a) and thus converts the adjunction structure to a set merge structure, PF remains unaffected, as categories moved overtly are spelled out at their landing sites. The same PF consequences follow when SIMPL applies to the edge. The scope of pied-piping is such that the adjunct is carried along with the active element of α (here the [+wh] feature on *which book*).

⁴⁰ Chomsky finds empirical support for this claim in the form of the following examples:

- (i) Bill remembered [that he heard [[α which person] [β who taught at Harvard]] had insulted him]]
 (ii) [α [α which person] [β who taught at Harvard]] (1) did Bill (2) remember [(3) that he (4) heard [t(5) had insulted him]]
 (iii) who (1) (2) remembered [(3) that he (4) heard [[α [α which person] [β who taught at Harvard]] (5) had insulted him]]

Example (i) shows the putative thematic position of the Wh phrase containing an adjunct. When this Wh phrase moves overtly to position (1) in example (ii), it passes through positions (2-5), yet the adjunct (β) can be spelled-out only in the position of the Wh phrase, α from definition (49). Example (iii) demonstrates the case of multiple Wh movement, where α is left in situ and the adjunct is spelled out in situ as well. Placing the adjunct in any other position in (ii) and (iii) causes severe degradation.

8. The great ACD hoax

The demise of late Merge as an option in the optimal system of computation requires reanalysis not only of anti-Reconstruction effects but also of another grammatical phenomenon amply documented in syntactic literature, namely the Antecedent Contained Deletion. In the context of his prohibition on late Merge, Chomsky critically reviews an approach to ACD constructions proposed in Fox (2001) and based on the concept of (string vacuous) extraposition and late adjunction. Let us first briefly consider Fox's proposal.

Fox (2001) takes up an approach to ACD proposed in Baltin (1987).⁴¹ Here, he proposes to replace his earlier analyses with the idea that construction of ACDs involves two steps: overt (Heavy Noun Phrase Shift) or covert (LF extraposition) movement of the object DP adjoining it to VP (51c), followed by late Merge of the adjunct. The first stage of overt or covert extraposition is represented below:

- (51a) You sent him_i the letter that John_i expected you would.
 (51b) You [_{VP} sent him_i the letter]
 (51c) You [_{VP} [_{VP} sent him_i the letter] {the letter}]

Following this movement, the relative clause is merged late with the QP in the landing site on the right periphery of VP. Crucially, there is no copy of the relative clause in the VP internal copy of the object DP (52c) and both VPs receive parallel interpretation at LF (52d):

- (52a) You [_{VP} [_{VP} sent him_i [the letter]] {the letter} that John_i expected you would]
 (52b) You [_{VP} [_{VP} sent him_i [the letter]] {the letter} that John_i expected you would [_{VP} send him {the letter}]]

Fox claims that the biggest advantage of this solution to the ACD problem comes in the form of the unification of anti-Reconstruction effects. In the previous account, every movement left a copy and this copy triggered Principle C effects. The only justifiable exception to this scenario was late insertion of adjuncts, as postulated in Lebaux (1988):

- (53a) Which book that John_i found in the bag did he_i like?
 (53b) [which book [that John_i found in the bag]] did he_i like {x book}

⁴¹ Baltin (1987) proposes to resolve the containment problem in ACD constructions through overt (but most often vacuous) extraposition of the QP and the relative clause to the right edge of the VP. In this way the ellipsis site is not embedded in the antecedent. This solution is critically reviewed in Larson and May (1990). Fox (2001) demonstrates how his new solution can deal with these criticisms.

The deletion of the antecedent VP internal copy of the QP in ACD constructions required to obtain structural parallelism was in conflict with the letter and the spirit of the Copy Theory of movement. This new account overcomes this discrepancy; adjuncts in both WH-questions and ACD constructions bleed Principle C through late Merge.

This new account presupposes the so-called phonological theory of QR and the idea that there are no separate cycles for overt and covert operations. In the system proposed in Chomsky (1995) no covert operations can precede overt operations within the same syntactic object, as two explicit cycles are assumed. Fox (2001) follows Nissenbaum and Fox (1999), Groat and O'Neill (1996), Brody (1995, 1997) and McGinnis (1998), and claims that both types of operations take place on the same cycle and can intermingle; the difference is that in the 'covert' ones the copy at the tail of the chain rather than at the head of the chain is pronounced.

Support for the idea that covert extraposition and late Merge of the adjunct to the NP at the right edge of the VP can bleed Principle C comes from examples discussed in Taraldsen (1981):

- (54a) I gave him₁ a book yesterday that John₁ liked.
 (54b) ?? I gave him₁ a book that John₁ liked yesterday.

Example (54a) is more acceptable than (54b) because the placement of the temporal adverbial *yesterday* indicates a movement operation. However, for Fox it is not the relative clause that is extraposed but the direct object *a book*, with which the relative clause is merged in its (unpronounced) landing site:

- (55) I gave [_{VP} [_{VP} him₁ a book yesterday] {a book} [that John₁ liked]].

Fox provides an interesting example to demonstrate that the bleeding of Principle C cannot be due to the extraposition of the relative clause only but to covert extraposition of the object followed by late Merge. Compare these two constructions:

- (56a) ?I [_{VP} [_{VP} told him₁ about your new argument the other day] that supports John's₁ theory].
 (56b) *I [_{VP} [_{VP} told you about his₁ new argument the other day] that supports John's₁ theory].

If the bleeding of Principle C were due only to the rightward movement of the relative clause, as in Baltin's original proposal, there should be no difference between (56a) and (56b), as in both cases the name embedded in the relative clause is not c-commanded by either *him* or *his*, both firmly embedded within the lower segment of the VP. Once the idea of covert extraposition and late Merge is adopted, there is a fundamental difference between the two examples:

- (57a) ?I [_{VP} [_{VP} told him₁ about your new argument the other day] {your new argument} that supports John's₁ theory].
 (57b) *I [_{VP} [_{VP} told you about his₁ new argument the other day] {his₁ new argument} that supports John's₁ theory].
 (57c) [DP his₁ [_{NP} [_{NP} new argument] [_{CP} that supports John's₁ theory]]]

Example (57a) is expected to be grammatical on either theory but if the relative clause is taken to modify the NP and be adjoined to it and *his* is a D that c-commands the NP in (57c), example (57b) is a clear Principle C violation.

If this general solution to both QR and ACD is adopted, the question of unacceptability of (58) below arises; why is the following derivation of this example based on rightward extraposition impossible:

- (58a) ??*Someone introduced him₁ to every friend of John₁'s.
 (58b) Someone [_{VP} introduced him₁ to every friend]
 (58c) Someone [_{VP} [_{VP} introduced him₁ to every friend] {every friend}]
 (58d) Someone [_{VP} [_{VP} introduced him₁ to every friend] {every friend} of John₁'s]

The idea is in fact simple and related to Lebaux's and Chomsky's arguments: the relational noun *friend* requires an argument and arguments cannot be added counter-cyclically, as this violates the Projection Principle and the Theta Criterion.⁴²

Notice that on the face of it, Fox's new account may be taken to support the picture of phase-based derivations; for instance before a larger syntactic object (phase) is constructed, all syntactic processes, both covert and overt, within a smaller one, here (vP), must be completed.

Also, when this general account of ACD constructions is adopted, the strongest argument for the position that Principle C applies only at LF collapses. Within derivations progressing in phases in which both overt and covert movement processes must take place phase internally without a separate LF cycle, the indirect object internal to the antecedent VP does not c-command *John* at any point in the derivation of (51-52).

⁴² If QR and ACD are based on covert extraposition, why is there a contrast between the cases of vacuous and non-vacuous extraposition, as in (54b) above and (ii) below:

- (i) I gave him₁ a new argument the other day that supports John's₁ theory.
 (ii) ??I gave him₁ a new argument that supports John's₁ theory.
 (iii) *I gave him₁ a new argument that John's₁ theory is incorrect.

Fox claims that this set shows strong contrasts. Example (i) is more acceptable because the parser has evidence for covert extraposition in the form of the placement of the relative clause. The derivation of (ii) is not blocked at all and is in fact available but the resulting structure is dismissed in performance. He follows Phillips (1996) and assumes that the parser shows a strong preference for the lowest attachment in cases of ambiguity, thus vacuous extraposition is much harder to access than visible extraposition. Example (iii) is dismissed independent of the parser; the appositive clause is an argument, it resists late Merge and the presence of the name in the VP internal copy brings about a Principle C violation.

There is an even more radical approach to ACD constructions: that is to deny their existence. This option is exercised in Chomsky (2001: 19-20), where ACD constructions are taken to be misinterpreted sentences with qualifications or afterthoughts (themselves adjuncts).

Chomsky begins by criticizing the analyses of QR, adjunct extraposition and ACD based on covert extraposition and late Merge of Nissenbaum and Fox and Fox (2001) as conflicting with the powerful postulate in (49) forcing SIMPLification, linearization and spell-out of adjunction structures at the point of the spell-out of the subject of modification. The above analyses of late Merge following covert extraposition of the host NP are successful descriptively because they explicitly follow a strategy opposite to postulate (49); the adjunct is never spelled-out in the position adjacent to the overt copy of the host NP. Chomsky observes that such scattered or discontinuous spell-out of adjunction structures calls for an unnecessary complication of operation Transfer at PF.

In place of all the previous attempts at dealing with ACD constructions, Chomsky proposes to treat them as regular cases of deletion, where the parallelism of the VP is achieved through specific assumptions concerning the constituent structure of the mainstream ACD cases, shown in (59b):

(59a) John likes every boy Mary does.

(59b) John likes [_{DP} [_{DP} every boy,] [(that is, more accurately, ...) every boy Mary likes]].

In (59b) the antecedent VP never contains the ellipsis site and there is no need for QR or extraposition to secure parallelism. The 'afterthought' segment can then undergo regular ellipsis.

The 'afterthought' hypothesis of QR and ACD, though interesting and desirable for conceptual reasons, appears to barely scratch the problem and offers no interesting insights on other properties of these constructions.⁴³ For instance den Dikken (1995), Hornstein (1995) and Hornstein and Witkoś (2003), note the following contrast:

(60a) John expects someone that I do to be in the room.

(60b) *John expects there to be someone that I do to be in the room.

Chomsky's 'afterthought' account does not distinguish between these two examples:

(61a) John expects [[someone]][(that is more accurately, ...) someone that I expect] to be in the room.

(61b) John expects there to be [[someone]][(that is more accurately, there to be) someone that I expect] to be in the room.

Both (61a) and (61b) seem to be legitimate 'afterthoughts' and input to PF deletion, yet an application of deletion to (61b) does not produce a grammatical result.

8. Multiple feature checking and multiple Agree

Hiraiwa (2001) observes that Chomsky's account of the Defective Intervention Constraint coupled with abandonment of the principle of Equidistance (Chomsky 1999: 21-22), leads to complications in empirical cases dealt with in previous stages of minimalism through multiple specifiers, multiple satisfaction of the [+EPP] feature and multiple feature attraction. Consider the following configuration, in which a probe is involved with case-agreement relations with two goals:

(62a) DIC: $*\alpha > \beta > \gamma$

(*Agree (α, γ), α is a Probe and β is a matching Goal, and β is inactive due to a prior Agree with some other Probe)

(62b) Multiple Feature Checking: $\alpha > \beta > \gamma$

(Agree (α, β) and Agree (α, γ), where α is a Probe and both β and γ are matching Goals for α .)

Defective Intervention Constraint (DIC) and Multiple Agree both occur in practically the same configuration and it is not so obvious that they could be distinguished. The system of case-agreement proposed by Chomsky assumes single Agree relations; Probe α accesses Goal β , they match in features and the case feature of β is valued, checked and deleted (or marked for deletion). Even if the [+ ϕ] features of the Probe do not check and delete and the Probe is [+multiple], it cannot access Goal γ , c-commanded by β , as the interpretable [+ ϕ] features of β are visible and block access to any Goal lower in the structure.

This is a very unwelcome result and leaves many cases of Multiple Agree unaccounted for.⁴⁴ Hiraiwa (2001: 69) proposes that, a parametrically determined, Multiple Agree option should become licit if the following principle is introduced:

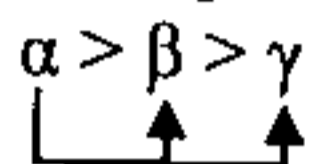
⁴³ For instance the problem of the licensing of ACD constructions in the subject position of ECM clauses and lack thereof in the subject position of regular finite embedded clauses.

⁴⁴ Here is an example of a construction involving multiple feature checking or Agree; in Japanese raising-to-subject constructions both the matrix and the embedded subjects can be marked Nominative by matrix T:

(63) Multiple Agree/Move:

Multiple Agree (multiple feature checking) with a single Probe is a single simultaneous syntactic operation; Agree applies to all the matched Goals at the same derivational point derivationally simultaneously. Multiple Move (movement of multiple Goals into multiple specifiers of the same Probe H) is also a single simultaneous syntactic operation that applies to all the Agreed Goals.

(64) Multiple Agree as a single simultaneous operation:



(Agree (α, β, γ), where α is a Probe and both β and γ are matching Goals for α)

At the point in the derivation where the Probe is merged into the syntactic object and it starts to search for the closest matching Goal within its c-command domain. Once such a Goal is located, the case-agreement relation is not established immediately but the Probe, being [+multiple], continues the search for other matching Goals within the entire active phase. Once the end point of the active phase is reached, Agree applies derivationally simultaneously, as a single operation, to all the matched Goals. Thus at the point of the derivation, where the Probe enters into the checking relation with the lower Goal, the higher c-commanding Goal is not yet inactive and does not produce the DIC effect.⁴⁵

9. Challenges for phase-based derivations

The stage of minimalism based on the probe-goal relation and case-agreement is the latest in its short history. Although it seems that it is a very promising line of research and will be followed in the future, let us put forward the following observations as problems for phase based syntax.

- (i) John-ga/ni [mada Mary ga kodomo-ni amaku] kanjita.
John-NOM/DAT still Mary-NOM children-DAT lenient-INF feel-PST
'Mary seems to John to be still lenient to children.'

⁴⁵ In order to accommodate Multiple Agree into the system, Hiraiwa (2001: 71) proposes the following reformulation of the DIC:

A syntactic operation Agree must obey a strict locality condition. Agree (α, γ) is prohibited if there is a close matching Goal that is already inactive at the point of the derivation where the Probe is merged; thus the DIC is restricted to a case where a Probe for γ and a Probe for intervening δ are derivationally distinct.

9.1. The status expletive constructions

The status of expletive constructions in the latest stage of minimalism is of primary importance. Why? Because expletive constructions show the priority of Merge over Move and lead to the dreaded cases of 'look-ahead'. This leads Chomsky to propose that derivations should run in phases. A possible problem for this account comes from an account of expletive constructions where Merge over Move does not hold, such as Hornstein and Witkoś (2003) and Witkoś (this volume). Once again, consider the cases that lead Chomsky to postulate phases and lexical sub-arrays reducing 'look-ahead':

(65a) There is someone wondering whether someone is here.

(65b) Someone is wondering whether there is someone here.

Example (65a) should be blocked by the acceptability of (65b) if both derivations started out with the same lexical array but the problem is solved if the two clauses in (65) form two different phases constructed from two separate sub-arrays. Hornstein and Witkoś (2003) and Witkoś (this volume) propose that the expletive is merged with the associate and the moves away from it in overt syntax for case reasons:

(66) [DP [DP there] someone]

This view of expletive constructions does not require phases and separate sub-arrays; if *there* does not directly merge into Spec TP, the derivations of (65) are not comparable once (66) is reached. From this point on, they have divergent numerations. The partial derivation in (67) leads to (65b).

(67) [[there+someone] is here]

The derivation that leads to (65a) proceeds through an early stage shown in (68):

(68) [someone is here]

Thus if the account of existential constructions based on (66) as the basic structure is feasible, and it seems to account for all properties of these constructions, a substantial part of the rationale for proposing phases is gone.

9.2. Disparity between the vP and CP phases

There is a puzzling disparity between the vP and the CP phase when it comes to their opacity to external probing: vP is porous to external probing, while CP is not. Consider the following configuration, representing a situation in Icelandic, in which *v* is transitive but does not assign Accusative case and the object DP is valued

against the T Probe and comes out as Nominative, our former example from Holmberg and Hroarsdottir (2003: 999):

(69a) Henni likuðu hestar.
her_{DAT} liked horses_{NOM}
'She liked horses.'

(69b) [_{TP} NP_S T [_{vP} {NP_S} v [_{vP} V NP_{OB}]]]

Crucially, for the relation of Agree between T and DP_{OB}, the strong phase boundary of vP should be porous. Incidentally, this case does not violate the PIC, although it certainly introduces the distinction between the strong/weak phase, as the relevant domains of the application of the PIC, so CP and vP are strong phases, while VP and TP are weak ones.⁴⁶ It follows from (69b) that a head of a weak phase can access, or probe into, the subjacent strong phase with impunity.

Now consider the other strong phase, CP. Interestingly, it is never transparent to the relation of Agree holding between an external probe and a constituent placed anywhere else but at the edge of CP. For example, it is not possible to have a relation of Agree between V and [spec,T] across CP:

(70a) *_{[VP V [CP [TP DP_{SU}]]]}

(70b) *_{[VP V [+CASE] [CP [TP PRO_{SU}]]]}

(70c) *_{[VP V [+CASE] [CP [TP DP_{SU}]]]}

So the verb cannot access PRO in the subject position, nor can it assign lexical case to the subject, yet the CP phase should be porous to external probing by the time the next strong phase, here vP, has been formed.

This disparity between the vP and CP becomes less of a freak coincidence if vP is a weak derivational phase, whereas only CP is the strong one. Consequently, the definition of PIC can be simplified to:

(71) Phase Impenetrability Condition (Simplified)

The domain of C is not accessible to further computational operations, but only C and its edge, where the edge consists of specifiers.

⁴⁶ In accordance with the definition of PIC in (27), it is at the next strong phase, that is CP, that only the head and the edge of vP should be accessible.

9.3. Phonology and vP phases

Chomsky claims that there is evidence in favor of vP and CP phases based on their PF properties, namely that they can be both topicalized and extraposed. Here is an example of vP topicalization:

(72a) She told him to wash the dishes and [_{vP} wash the dishes] he did t.

(72b) Maria kazała mi zmywać naczynia i [_{vP} zmywać naczynia] będę t do zmroku.
Maria told me wash-Inf. dishes and wash-Inf. dishes will-1Sg. till dusk
'Maria told to wash the dishes and wash the dishes I will till dusk.'

Strong phases, including vP, are transferred to PF operations as whole indivisible chunks and can move as units. Weak phases are not fed to PF operations independently, only as subparts of large units and are not expected to be displaced as chunks.

There is an obvious loophole in this type of argumentation; think about vPs and VPs. Consider a VP headed by a passive participle; it is unaccusative, hence a weak phase by definition. Consequently, it is not expected to allow for topicalization or extraposition, contrary to fact:

(73a) [_{VP} broken into pieces] was that book

(73b) Mówiłem ci, że będzie wyrzucony na bruk i [_{VP} wyrzucony na bruk] został właśnie wczoraj.
told-1Sg.Msc. you that will-3Sg. thrown-Msc. on street and thrown-Msc. on street became-3Sg. just yesterday
'I told you that he would be thrown onto the street and thrown onto the street he was just yesterday.'

Both (73a) and (73b) seem acceptable. But if strong and weak verbal phases can be extraposed, the topicalization test does not distinguish between the two and the argument for the PF independence of strong phases is considerably weakened.

Incidentally, Chomsky (1998: 35) uses a similar example of complex displacement, where VP is moved as a unit, (74a):

(74a) [[written t_j for children]_i, [those books]_j couldn't possibly be t_i]

(74b) ?_{[CP [written t_j for children]_i, I think [_{CP} t_i [those books]_j couldn't possibly be t_i]]}

Here, the passive VP is moved as a unit, and as (74b) shows it can even be moved long distance into another phase.

On the other hand it is possible to argue that passive VPs constitute strong phases after all. Svenonius (2002: 202) argues that in the passive the verb can still license the external argument, in the form of the *by*-phrase, and so its underlying representation should include v:

(75) [VP [PP by DP] v [VP V DP_{OB}...]]

This assumption does not introduce any serious additional complexity into the system of computation, as the passive transformation could include the intermediate step, akin to Indirect Feature Movement (IMF) in long distance movement, at which the agentive *v* would require the [+EPP] feature to cause Object Shift:

(76) [TP T [VP DP_{OB} [v [PP by DP] v_[+EPP] [VP V {DP_{OB}} ...]]]]

We can observe, however, that claiming that passive verbal phrases are strong phases is a step away from Chomsky's design. First, it plainly contradicts his definition of strong phases in (Chomsky 1998: 20) and second, it separates passives from unaccusatives and thus seems to run against the long tradition of regarding passive participles and unaccusatives as a natural class, falling under Burzio's (1986) Generalization.

9.4. Problems with motivating Long Distance Movement

One of the rather inelegant consequences of phase-based derivations is the concept of Indirect Feature Movement (IFM), whereby a constituent moves to satisfy the PIC and the system must propel the movement in small cycles through the application of features [+EPP] features on relevant heads, mostly *v*. Let us consider the relevant configuration, discussed in Chomsky (1998: 24):

(77a) Which book do you think Susan lost?

(77b) [CP [which book]_i ... [VP t_i ... v_[+EPP] ... [CP t_i ... [VP t_i ... v ... t_i]]]]

Śmiecińska (2003: 63-64) considers such cases and asks the obvious question: why should the entire system work only thanks to a purely engineering solution? To see this point more clearly, consider the stage of the derivation in the matrix clause; why should the matrix *v* have the [+EPP] feature?⁴⁷ This feature accompanies others and is justified in the case of the embedded *C* (it may accompany some [+wh] feature) and *v* (it accompanies the full set of [+φ] features valuing case). But what is its function in the matrix clause? It turns out that none but to let this derivation converge pure and simple without violating the PIC. Note that the movement through the intermediate positions cannot lead to the [+wh] feature valuation of the moving constituent, since that would automatically render it 'inactive' and frozen in place. From

the point of view of feature composition, this derivation would be much simpler if the matrix *C* could access the Wh-phrase in its Case position.⁴⁸

9.5. Reconstruction sites at the vP edge

Chomsky (1998: 22) also makes the following comment as evidence for the strong-phase status of vP:

(78) "The impenetrability condition requires that A' movement target the edge of every [strong, J.W.] phase, CP and vP. There is evidence from reconstruction effects and parasitic gap constructions that this may be true."

He credits this observation to Fox (1999, 2000). This argument for reconstruction is based on the following premise: in the process of Wh-movement the Wh phrase probably docks at the edge of vP and late Merge of an adjunct in this position allows it to bleed Principle C effects with respect to the indirect object. Consider the following illustrative case from Fox (2000: 165):⁴⁹

(79) [which of the papers {that he₁ wrote for Ms. Brown₂}] did every student₁ t get her₂ t* to grade e*?

Fox follows Lebaux (1988) and Chomsky (1993) and assumes that adjuncts can be merged with the syntactic objects they modify at any point in the overt derivation. In the example at hand, the relative clause in the curly brackets could in principle have been merged with *which of the papers* either in the thematic position (e*), in [spec,C] of the embedded infinitive (t*) or at the edge of the matrix vP (t).

Fox constructs the example in such a way, that the pronominal variable identifies the reconstruction site of the restriction on the Wh-phrase: it must be interpreted in the position c-commanded by *every student*. He observes that the well-formedness of this example means that Wh-movement proceeds through the edge of vP, as (t) is the only reconstruction site, where *he* is in the c-command domain of *every student* and *Ms. Brown* is out of the c-command domain of *her* and does not violate Principle C.

⁴⁸ In connection with the same issue Fanselow and Mahajan (2000) argue that already the embedded *C* cannot trigger the movement of Wh-phrase through a [+wh] feature, as this would violate selectional restrictions on *say*, which does not select for an interrogative complement. They suggest that a categorial [+D]-like feature should be responsible for this movement. This obviously leads to the issue of long distance questions involving Wh-phrases that are not DPs.

⁴⁹ Originally, Fox derives these properties of reconstruction from a *Barriers*-style VP adjunction.

⁴⁷ There have been numerous attempts in the minimalist literature to dispose of the [+EPP] features altogether, e.g. Grohman, Drury and Castillo (2000).

Chomsky implicitly takes these facts to mean that the EPP feature of *v* forces movement through the edge of the *vP* phase.⁵⁰ These facts may not be as strong arguments for the strong phase status of *vP* as they seem for the following three reasons.

Weak phases may also require the EPP feature, as the discussion of TH/EX shows. The EPP feature on *v* is crucial for Reconstruction effects with strong phases, so the fact that it may appear on weak ones undermines its status as a diagnostics for strong or weak phasehood.

Once *vP* is not regarded as a strong phase and movement through its edge position is not required under PIC, the question arises how the reconstruction facts in (79) should be interpreted.⁵¹ Chomsky himself hints at a possible answer in several places. He repeatedly claims (inspired by Reinhart 1995 and Fox 1999, 2000) that optional features are present in the derivation if they make an interpretative difference. Thus an optional EPP feature on *v* can facilitate a bleeding of Principle C, instrumental for receiving a particular interpretation at the LF interface.⁵²

Finally, there is evidence that most reconstruction effects can be captured without recourse to the placement in the *vP* edge position. Kuno (1997), Safir (1999) and Witkoś (2002) argue that the process of vehicle change, independently required in the grammar as argued in Fiengo and May (1993), can take over the function of late adjunction site. Thus if the name embedded within the Wh-phrase is replaced by its pronominal correlate in the VP internal position of the trace (copy) Principle C effect disappears and Principle B effect is considerably weak. Take the following example (Witkoś 2002: 332-33, fn. 19):

- (80a) [ile książek Piotra₁] Maria oddała mu₁ [x książek Piotra₁] po roku?
how many books Piotr's Maria returned him books Piotr's after year
'How many of Piotr's books did Maria return to him after a year?'
- (80b) [ile książek Piotra₁] Maria oddała mu₁ [x jego₁książek] po roku?
how many books Piotr's₁ Maria returned him₁ his₁ books after year

⁵⁰ As a matter of fact there may be no Late Adjunction at the edge of *vP* in the classical sense of Lebaux (1988). As Chomsky (1998, 1999, 2001) argues the principle of Least Tampering forces early (pair) Merge of adjuncts in a separate plane invisible to the rules of computation and their subsequent TRANSFER at any point in the structure. All the consequences of the typical Late Adjunction follow if TRANSFER applies to the copy at the edge of *vP*, rather than in the VP-internal position.

⁵¹ Abandoning the spec/head configuration as a feature checking configuration implies that the troublesome issue of checking the case feature of the WH-object is now resolved (c.f. Kitahara 1997). The case of the object is valued in situ, prior to any displacement. If *vP* is not a strong phase, the EPP feature of the probe C can access the object in its base position.

⁵² Chomsky (1998: 24): "One might also explore a generalization of the idea that operations can apply only if they have an effect on outcome. This would translate here into restricting [occurrence of the [+EPP] on the head of a phase, J.W.] to the condition in which it permits IFM or specific interpretations associated with peripheral positions, e.g. specificity and informational conditions on Object Shift."

- (80c) *Maria oddała mu₁ książki Piotra₁.
Maria returned him₁ books Piotr's₁
- (80d) Maria oddała mu₁ jego₁ książki.
Maria returned him₁ his₁ books.

Here, an account of the bleeding of Principle C effect between the two objects need not require movement through the edge of *vP*; vehicle change affects the VP internal copy to produce the result equivalent to changing (80c) into (80d) in the non-interrogative context.

To recapitulate the main point of this section, Reconstruction effects at the edge of *vP* need not correlate with the strong phase status of the *vP* and can be rendered in many alternative ways without invoking obligatory movement to this position.

9.6. Violations of the Phase Impenetrability Condition

From the perspective of Chomsky's PIC the status of the infinitive clause in various construction types is crucial. Chomsky (1998, 1999, 2001) relies on the Null Case hypothesis of Chomsky and Lasnik (1993) and Martin (1996, 1999, 2001) and stresses the fact that infinitives embedded under control predicates are full CPs with complete TP complements, whereas infinitives embedded under raising verbs are only defective TPs.⁵³ Phase-wise, CPs are strong phases while TPs are not.

This picture, and the status of phases, is challenged by a construction known as the Long Genitive of Negation (LGoN), as observed in Błaszczak (2001) and Witkoś (2003). Consider the following example:

- (81a) Maria nie chce wypić kawy.
Maria not wants drink coffee-GEN/*ACC
'Maria doesn't want to drink coffee.'

- (81b) ...[NegP Neg-P [VP ... v ... [CP [TP PRO_[+φ] [VP t_{PRO} ...v ... DP_{OB-G}[+φ]]]]]]]

⁵³ Chomsky and Lasnik (1993) observe that certain aspects of the distribution of PRO make it look very similar to other case-marked DPs:

- (i) John tried [PRO_i to be arrested t_i].
(ii) *John tried [PRO_i to seem to t_i that the problem is unsolvable].

In the passive construction (i), PRO moves in accordance with the Last Resort Condition to check case and (ii) is correctly ruled out if the chain of PRO includes two case positions, one in the complement of P and the other in [spec,T]. Ruling out example (ii) is a problem for the GB accounts of PRO placing it in an ungoverned position: PRO seems to be legally escaping the consequences of government within PP in this case.

Błaszczak (2001) follows a plausible assumption, based on previous analyses in Willim (1990), Dziwerek (1997) and Witkoś (1998), which holds that Genitive on the object should be caused by some case-agreement relation involving the object as the Goal (DP_{OB-G}) and matrix negation as the Probe (Neg-P). She observes that this is, however, exactly the type of relation prohibited by the Phase Impenetrability Condition (PIC). As example (81b) shows, the Goal DP_{OB} is not at the edge of the strong phase vP, so it ceases to be accessible to any probing past the point at which the embedded CP has been constructed. Even if Neg were assumed to be a part of the matrix vP phase, it is separated from its purported Goal by the CP phase and the embedded vP phase. Błaszczak (2001) concludes that phase-based minimalism is unable to account for the phenomenon of Long Distance GoN.⁵⁴

Witkoś (2003) adds a further observation on the status of example (81) and similar ones. There is one more problem for GoN in an approach based on single cycle; that is an approach discarding the separation of the overt and covert cycles in the derivation. Even if the entire example (81) is lumped into a single derivational phase, so that PIC is satisfied, there is the issue of the Defective Intervention Constraint (DIC), also exemplified in (81b). There should be interference in the case-agreement relation between Neg-P and DP_{OB-G} coming from PRO, whose [+φ] features are interpretable, hence visible to the Probe, yet its uninterpretable CASE feature is checked, which makes it an invalid Goal, though a valid intervener (β) in the case-agreement relation between α and γ, (Chomsky 1998: 38-39):⁵⁵

(82) *α > β > γ

The case of the LGoN in Polish seems to provide an insurmountable challenge to derivations running in Chomsky's phases.⁵⁶

Another challenge seems to come from multiple Wh-movement constructions in English. As we saw above in (29), phase-based derivations allow for countercyclic operations internal to a phase but at the phase-level the ensuing representation needs

⁵⁴ Note that the problem of Long Distance GoN remains even if it is assumed that the object is moved to the edge of vP in overt syntax (if Polish has Object Shift), as the edge of the embedded vP phase is visible only up to the CP level and for the object to be visible to the matrix Neg, it would have to be at the edge of the CP phase, an empirically untenable position.

⁵⁵ Chomsky (1998: 38) comments on the Defective Intervention Constraint: "We therefore have the possibility of a defective intervention constraint [in structure (82), J.W.], where > is c-command, β and γ match the Probe α, but β is inactive, so that the effects of matching are blocked."

⁵⁶ Witkoś (2003) proposes to solve this problem by substantially modifying Chomsky's system and assuming that vP is never a strong phase and both control and raising infinitives are TPs. This automatically solves both the PIC problem, as Negation and the object in example (81) are now within the same strong phase, and the DIC problem as traces/copies of the subject, unlike PRO, do not qualify for interveners.

to comply with the representational notion of the cycle. Consider a simple multiple Wh-construction:

- (83a) Who knows what?
 (83b) [TP T [vP who v [vP V what]]]

The derivation of (83a) passes through an intermediate stage indicated in (83b), where both Wh-phrases are members of the same phase. It is imaginable, if MLC is built into the definition of movement, that at the stage (83b) the fate of *what* as a Wh-in situ is decided without the dreaded 'look-ahead'. Because C can have only one specifier in English only one Wh-phrase can be moved and it must be *who*, because already at stage (83b) it c-commands *what* and is thus closer to C. The movement of *who* to [spec, T] places it even closer to C but the matter of which of the two Wh-phrases moves is decided already at the vP phase level.⁵⁷ Thus no illicit step is taken.

Now consider the following example of multiple Wh-movement, where 'look-ahead', even within the phase-based derivation cannot be avoided:

- (84a) Who₁ did Peter tell t₁ [PRO to read what₂]
 (84b) ?Peter told this student [PRO to read what]
 (84c) [vP PRO v [vP V what]]
 (84d) [vP [what]₁ [v' PRO v [vP V t₁]]]
 (84e) [CP [what]₁ C [TP T [vP t₁ [v' PRO v [vP V t₁]]]]]
 (84f) [vP Peter v [vP who [v' V [CP [what]₁ C [TP T [vP t₁ [v' PRO v [vP V t₁]]]]]]]]]

At the local stage (84c) the Wh-phrase can either move or not, or to be more precise, following Chomsky (1998: 45) either the entire Wh-phrase or only its abstract head is moved. But how can the decision be reached locally within the embedded vP? This decision hinges on the composition of another phase, chiefly, the presence of another c-commanding Wh-phrase. If the object Wh-phrase does not move and other sub-arrays do not contain a suitable Wh-phrase, as in (84b), the sentence may end up as gibberish or receive an echo question interpretation.

Let us say, the Wh-phrase is moved to the edge of its local phase, courtesy of the optional [+EPP] feature on v, in anticipation of further movement, (84d). In the meantime another strong phase (CP) is constructed and IFM forces the movement of *what* to [spec,C], (84e). Only at the point of constructing another phase (two phases away from the source) does it become evident that the entire effort of pied-piping *what* has been in vain (84f). At the stage (84f), another c-commanding Wh-phrase is introduced into the derivation and MLC can immediately and locally decide that it will be moved up.

⁵⁷ This decision may involve either straightforward movement and non-movement of the Wh-phrases or the movement of both Wh-operators but the pied-piping of only one phrase (*who*).

The conclusion seems to be that the dreaded 'look-ahead' property of phase-less derivations rears its ugly head despite all efforts to the contrary.

9.7. Copy deletion algorithm and application of the LCA

A serious conceptual problem for the system based on case-agreement comes from the area of the algorithm for linearization of copies at PF. Nuñez (1995, 1999) proposes a very appealing minimalist rendition of LCA: the axiom was useful in deterministically striking out all copies in chains of overt movement but the copy at the head of the chain. If every single copy in the chain participated in the calculation of the LCA, the string would never be linearized, as the ordering of terminal elements could not be total. Here is why:

(85) What did Bill eat {what}

If all copies survived to the point of application of LCA, the terminal string in (85) could not be ordered properly, because *Bill* both c-commands *what* and is c-commanded by it.

Nuñez' basic idea was that the choice of copies to be deleted was determined by the number of deletion operations necessary for pruning the excess copies. If LCA is a condition of convergence at PF, it forces the deletion of all but one copy but principles of economy determine which copy is retained. The difference between the copy at the head of the chain and the one at the bottom is that the one at the head has a different feature composition from other copies in the chain. Consider the example again, couched in the technicalities of the model in Chomsky (1995):

(86) What [-wh, -case] did Bill [what [+wh, -case]] eat what [+wh, +case]]

We have made the picture more complex and more precise by introducing the third copy of *what* in the case checking position in the mode of the Agr-less minimalism. The topmost copy has all its features checked and deleted, the middle copy has one feature checked and deleted and the bottom copy has no features deleted. Nuñez argues that if the applications of operation Delete are counted in the computation, the most economical way of selecting the right copy to be left is to leave the topmost copy. In any other case unchecked features within the lower copy must be deleted and these additional applications of deletion are uncalled for.

Now this picture of the algorithm for copy deletion cannot apply. Why? Because spec/head is not a checking configuration and the feature composition of both the topmost copy and the copy immediately below it is identical:

(87) What [-wh, -case] did Bill {what [-wh, -case]} eat {what [+wh, -case]}

In the current system C matches, checks and deletes the features of the Wh-phrase when it is at the edge of vP because only in this position is the Wh-phrase in the search space of C. Further movement to [spec,C] does not affect the feature composition of the topmost copy. Observe that the two relevant copies, the one in [spec,C] and the one in [spec,v] are indiscernible from the perspective of the number of deletion operations required by the LCA: both have all the features checked. It seems that this is the price Chomsky's system pays for discarding the spec/head as a checking configuration. As Chomsky himself has relatively little to say about the selection of copies for deletion, an interesting insight of Nuñez' into the nature of PF deletion and the role of LCA has been lost.

9.8. Feature composition and the timing of Spell-out

Epstein et al. (1998) and Epstein and Seely (2001) develop a radically derivational approach to syntactic relations and argue against Chomsky's mixed derivational-representational system.

They criticize the notion of derivational phase and essentially propose that every rule application should be treated as a separate cycle and every syntactic object formed in the derivation is a phase.

Their critique of phase-based derivations concerns several aspects of the system. They observe that choice of vP and CP as cyclic nodes (phases) in Chomsky (1998, 1999, 2001) is purely arbitrary. Epstein and Seely (2001: 78) notice that there are transitive vPs that are not prepositional, (88a-b), there are propositions that are neither vPs nor CPs (88c); there are also constituents within which all theta roles compatible with the head are discharged, yet they are unaccusative or passive verbal projections (88d):

- (88a) [_{vP} who bought what]
- (88b) [_{vP} everyone bought something]
- (88c) They consider [_{sc} John smart]
- (88d) [_{vP} arrived the boys]

In examples (88a-b) at the vP level, prior to movement, these syntactic objects are not propositions but cases of vacuous quantification and subsequent to movement, they are open sentences containing variables. In example (88c) the small clause is a proposition and in example (88d) VP should be a phase when the completeness of the discharge of theta roles available to a given head is assumed to be the relevant criterion.

They also raise three conceptual problems for the definition of phases. First, they are said to be relatively independent at the interface level. The problem with this definition of phases is that interface levels are reached at the end points of the

derivation and it would be counterintuitive and non-cyclic to define phases operational throughout the derivation with reference to its end points. This is a clear case of 'look-ahead'. Another problem is that if phases are defined in both phonological and LF terms, their intrinsic and defining properties should be relevant at both levels at the same time. It is unclear, though, why the PF representation should be concerned with the prepositional nature of the phase. Finally, phases, as privileged points in the derivation, where the split into LF and PF applies in the familiar T-shaped fashion, may be seen as an attempt at reintroducing levels of interpretation specific only to the syntactic component of grammar and thus unrelated to the interface levels and bare output conditions.

The main thrust of the anti-phase argumentation is directed against the idea that Spell-out should apply at the phase level. One of the basic postulates of minimalism is that syntactic operations are feature driven and specifically, the features propelling the derivation are uninterpretable at either the PF or the LF interfaces. The uninterpretable features must be matched, checked/valued and deleted (or stripped away) from the narrow syntax leading to LF (though many valued uninterpretable features have a PF reflex, such as agreement on T/VB or structural case on D). In Chomsky (1999, 2001), uninterpretable features lack a value (e.g. [+case] can be either Accusative or Nominative) and they are valued upon matching against a probe.⁵⁸ Thus the valued/unvalued distinction plays the role of intra-derivational distinction between interpretable and uninterpretable features.⁵⁹ The function of Spell-out in the system is to remove the valued uninterpretable features from the computation, as even valued uninterpretable features cause non-convergence at LF.⁶⁰ Chomsky (1998, 1999, 2001) argues that because features are checked and valued iteratively in the derivation, they should also be iteratively removed from the computation, unlike in earlier minimalist models (Chomsky 1995, where two distinct transformational cycles were posited). But Spell-out does not apply iteratively at every point where uninterpretable features are valued; it is delayed until the phase level (or the next-phase-up level in more complex derivations). Epstein and Seely focus on the following dependency between the timing of Spell-out and the valuation of LF uninterpretable features:

- (89a) Spell-out cannot operate prior to valuation, as unvalued features cause non-convergence at PF as well as LF interfaces; it must therefore apply after the valuation to know which feature must be spelled-out and removed from narrow syntax.

⁵⁸ Uninterpretable/unvalued features can be placed on either the Probe or the Goal, hence the unvalued [+case] feature is placed on the DP (Goal) but the unvalued [+EPP] and [+ φ] features are placed on the Probe (T).

⁵⁹ Interpretable features are valued by definition.

⁶⁰ They either violate FI at LF, as [+case] or [+EPP] features or converge as 'gibberish', as valued [+ φ] features of T.

- (89b) "The operation Spell-out removes LF-uninterpretable material from the syntactic object K and transfers K to the phonological component. It must therefore be able to determine which syntactic features are uninterpretable, hence to be removed. Prior to application of Agree, these are distinguished from interpretable features by lack of specification of value. After application of Agree, the distinction is lost. To operate without reconstructing the derivation, Spell-out must therefore apply shortly after the uninterpretable features have been assigned values." (Chomsky 1999: 5)
- (89c) "The valued uninterpretable feature can be detected with only limited inspection of the derivation if earlier stages of the cycle can be 'forgotten' – in phase terms, if earlier phases need not be inspected." (Chomsky 1999: 12)

Epstein and Seely point out that 'shortly after' the valuation of features uninterpretable features is probably too late anyhow and may lead to unnecessary and illicit cases of 'look-back'. They illustrate their point with the following example (Epstein and Seely 2001: 80-81):

- (90) [_{CP} what did Fred say that Bill said that Fred said that Bill said ... that John will [_{VP} what buy]]

In order to send *what* to Spell-out, the system must know exactly which of its features are relevant (PF interpretable and LF uninterpretable). The inspection of *what* in matrix [spec,C] does not distinguish between the interpretable and uninterpretable features because at this point all of the features of *what* have been assigned values through checking ([+wh] and [+case]). Thus the derivation must be inspected in search of occurrences of unvalued features at earlier derivational steps. This inspection, however, cannot be entirely innocent, as in order to detect the unvalued occurrence of the feature [+case] on *what*, the system needs to look all the way back at the copy at the edge of the embedded vP. Now this is in radical disagreement with Chomsky's postulate in (89c), and in violation of PIC (if the two very distant occurrences of what where to be compared), especially, as earlier phases could (and should) have been spelled out earlier and as such inaccessible to the derivation. Thus, logically, in order to be operational, Spell-out inspecting the derivation would have to access the entire derivational history or some radically anti-phasal representation.

In return, they propose to treat Spell-out as a part of transformational operations and provide for its application in the following manner, Epstein and Seely (2001: 75):

- (91a) "Spell-out must itself be a derivation-sensitive operation."

- (91b) “The null hypothesis is that it can examine only the input to a given transformational rule application (in which a feature is unvalued) and the output of that same transformational rule application (in which the feature is now valued).”
- (91c) “We propose that Spell-out operates on all and only those formal features that appear without a value in the input to a rule being applied, but appear in the output of that rule application with the “previously” unvalued feature now valued.”

With Spell-out operating after each rule application (each outer or inner Merge) they see no need for derivational phases, as each category becomes a phase in this system. There is naturally, no need for inspection for feature occurrences or ‘look-back’.

This criticism reveals architectural paradoxes of the phase-based system and questions both the empirical and conceptual justification for phases. Yet, this system itself may in certain respects turn out to be more conservative than Chomsky’s.

For instance, the radically derivational approach may have to reject certain novelties in phase-based derivations and all the vestiges of representationalism. Take Chomsky’s recent arguments against the principle of Equidistance. In the partial representation below, Chomsky allows for non-cyclic operations and a local violation of MLC in the following case provided the violations have been repaired by the time the construction of the strong phase is completed:

- (92a) (guess) what_{OB} [John_{SUB} T [vP t_{OB} [t_{SUB} read t_{OB}]]]
- (92b) John_{SUB} T [vP that_{OB} [t_{SUB} read t_{OB}]]
- (92c) # [TP ... T ... [vP DP_{OB} [vP DP_{SU} v ... t_{OB}]]]
- (92d) [CP DP_{OB} C [TP DP_{SU} T ... [vP t_{OB} [vP t_{SU} v ... t_{OB}]]]]]

At stage (92c) Probe should access the subject at the edge of the strong vP phase, yet it is screened from it by the object shifted to the edge of vP. With the concept of Equidistance abandoned in favour of the plain minimal c-command, the object should cause an intervention effect. Yet, for Chomsky (1999, 2001) MLC is a condition on representation at the strong phase level. Phase-internally (92c) is illicit but the MLC does not apply to the syntactic object at this stage but at stage (92d), where the object has been moved up to [spec,C] and its copy/trace is not a defective intervener.⁶¹

⁶¹ We explicitly refer to the case of object Wh-phrase movement to [spec,C]. In the Icelandic Object Shift construction the object at stage (92d) would have been PF moved leaving a copy/trace.

Strictly derivational approaches must immediately reject such an account, as here MLC is a condition on derivation and no violation of MLC is tolerated at the intermediate step (92c). Thus Epstein and Seely (2001) need to keep the notion of Equidistance in its formulation from Chomsky (1995, 1998)⁶² and the notion of parametric difference in the strength of the T probe between English-type languages and Icelandic-type languages.

There is also one aspect of the discussion of intraderivational Spell-out in Epstein and Seely (2001) that must be highlighted. First, consider the discussion of features involved in Wh-movement. In contrast to Chomsky (1995), Chomsky (1998: 44–45) assumes that both the [+Q] feature on C and the [+wh] features on the Wh-phrases are uninterpretable, and are in this sense analogous to structural Case. The Wh-phrase, however, also carries an interpretable feature [+Q], matching the uninterpretable feature [+Q] on C. The Wh-phrase is active in the derivation as long as its [+wh] feature has not been checked and deleted.⁶³ This short elaboration on feature types serves us to show the construction in (90) above in a new light:

- (93) [CP what did Fred [vP what [+wh] say that Bill said that Fred said that Bill said ... that John will [vP what [+wh] buy what [+case], [+wh]]]]]

In their discussion Epstein and Seely concentrate on the fact that an occurrence of the unvalued [+case] feature must be detected within the thus-far constructed syntactic object. In fact, they are concerned not with the Spell-out of the category Wh-phrase but with the Spell-out of the (valued but LF uninterpretable) feature [+case].

⁶² Chomsky (1998: 38): “Terms of the same minimal domain are ‘equidistant’ to probes.”

⁶³ This assumption raises the question of successive cyclic WH-movement in (99) above. On the assumptions of Chomsky (1995) the embedded C in long-distance WH-movement has [+Q] feature. When the probe of the embedded C finds the goal in the form of the [+wh] feature of *which book* and the EPP feature leads to overt movement, *which book* should not be accessible to further movement out of CP, as interpretable features (Q) do not drive movement. This problem was solved in Chomsky (1995) on the assumption that F_Q/wh-feature on the WH-phrase is interpretable and repeatedly accessible to operation Attract. In Chomsky (2000: 149, fn. 91) the intermediate C is supposed to be equipped with an additional EPP feature; in this way the WH-phrase is moved without having its uninterpretable features checked prematurely.

Wh-island effects are treated like cases of defective intervention illustrated for superraising in (96) and below:

- (i) *How do you wonder [CP what to repair t t]?
- (ii) *[CP [Q do] C_{[+EPP],[+Q]} you wonder [CP what_[+Q] to repair {what} how_{[+Q],[+wh]}]]]

In (ii) the matrix C has features [+Q] and EPP and still acts as a probe for the goal carrying an unchecked [+wh] feature. On the strength of the PIC, the Probe can sweep the structure for an eligible candidate only to the edge of the nearest phase, embedded CP. Now *what* has its interpretable feature [+Q] available but its uninterpretable feature [+wh] already checked and deleted, rendering it inactive and inaccessible to Agree. Although inactive, embedded *what* is still visible and blocks the Probe on the matrix C from reaching *how*. As a result, relation Agree is not established, the Probe has an unchecked uninterpretable feature ([+Q]) and the EPP feature and the derivation crashes.

But the Wh-phrase has two uninterpretable features at the beginning of the derivation and by extension both must be initially unvalued: [+case] and [+wh]. Certainly, in the derivation of (93) the phase-based Spell-out has no problem locating the unvalued occurrence of the [+wh] feature, as it is at the edge of the root vP. The problem for Chomsky is that *what* is spelled out once but the operation must involve substantial look-back to determine which features should be stripped away from narrow syntax (PF features and valued LF uninterpretable features). For Epstein and Seely, Spell-out must access *what* at least twice: once in the position of the edge of the most deeply embedded vP where the valued but LF uninterpretable feature [+case] is eliminated from narrow syntax and again in [spec,C] where the valued but LF uninterpretable [+wh] feature and other PF relevant features are eliminated.

The key differences in the conception of Spell-out between both systems are summarized below:

- (94) Chomsky's system:
- (a) Spell-out targets categories, as bundles of features, not individual features;
 - (b) PIC and the escape hatch in the form of the edge of the strong phase allow for a single application of Spell-out to a given constituent.
- (95) Epstein and Seely's system:
- (a) As is evident from the quote in (91c), Spell-out can target individual features, not entire categories;
 - (b) Spell-out can apply repeatedly to the same constituent and target individual features.

It is not at all clear which conception of Spell-out is superior. Of course, unlimited look-back in Chomsky's system is a flaw in the design but repeated Spell-out of the same category raises some fundamental questions. For example Spell-out applies to both PF relevant features and LF uninterpretable features; if 'scattered' Spell-out of the latter were allowed (as in the example above, where the [+case] feature is spelled-out in a different position from other features), one might expect 'scattered' Spell-out of the former as well. This, however, is not the case: words are not scattered all over the place but spelled-out as wholes. Observe, that this point is not an argument against multiple Spell-out, because it applies repeatedly to the syntactic object constructed in the derivation but only once to a given category. In the light of the discussion above the following notion may be worth considering:

- (96) Impatient Spell-out (Holmberg and Hroarsdottir 2003; Svenonius 2001):

A category is spelled-out when it is 'syntactically complete', in the sense that any unspecified features are assigned a value.

The notion in (96) implies that on the one hand Earliness holds and Spell-out is impatient, but on the other it must lag behind the checking operations, as in the phase-based system where the edge of the currently constructed phase is not spelled-out to allow for movement and valuation of more than one uninterpretable feature on a category before it is spelled-out as a whole. Thus, it need not be necessary to inspect the derivation for every occurrence of an unvalued feature within a given category; it may suffice to spell it out when every feature on this category is valued.⁶⁴

10. In place of conclusion

Conceptual arguments seem to favour phase-based derivations. Ideally, memory load should be reduced and the system of computation should be very primitive and simple. Yet, many empirical problems remain and it seems that the cyclic status of CP seems corroborated, though even this view can be misleading, as the CP is a complete domain, where all grammatical relations should in principle be satisfied.

Thus the fact that CP-external material may have little bearing on CP-internal processes can be as well due to the combined effect of MLC and Maximize and Minimal Tampering (CP-internal probes are closer to CP-internal goals than their external equivalents). For example Franks (2000), Fanselow and Mahajan (2000) and Śmiecińska (2003) argue against cyclic status of the intermediate CP in cases of long Wh-movement in English. Consequently, derivations could operate on large syntactic objects, amounting to several clauses. If the view of expletive constructions of Hornstein and Witkoś (2003) and Witkoś (this volume) is adopted, the supposed priority of Merge over Move in these constructions ceases to have bearing on the size of the lexical array and therefore removes a chief empirical argument for phases.

On the other hand there is mounting evidence, both conceptual (Uriagereka 1999, Epstein and Seely 2001) and empirical, for instance Fanselow (2003),⁶⁵ that

⁶⁴ The remaining problem is how to distinguish which feature to strip away from narrow syntax once they have all been valued, as Epstein and Seely observe. It may be necessary to return to the previous three way distinction of formal features to facilitate their discrimination at Spell-out: unvalued features causing crash at both PF and LF, the ones valued in Initial Numeration and the ones valued in the computation (formerly known as 'deleted but not yet erased and accessible' in Chomsky 1995).

⁶⁵ Fanselow (2003) discusses interesting cases of topicalisation in German, where the constituent moved to the preverbal position in V-2 contexts is only a subpart of the topic of the sentence. Fanselow assumes that the topic carries a feature that requires checking against the head of CP:

- (i) War er fromm?
'Was he religious?'
- (ii) In der Bibel gelesen hat er nur selten.
He has only rarely read in the Bible.'
- (iii) In der Bibel hat er nur selten gelesen.

In example (iii) *in the Bible* need not be an independent topic but can stand for the VP topic, thus examples (ii) and (iii) can have the same information structure; *in the Bible* serves as a proxy topic for the VP.

the narrow syntactic computation is fed piecemeal to the domain of interface processes. A strict derivational approach eschewing the notion phase and postulating sending each step of the derivation to the interpretive procedures PF and LF is quite promising but faces problems of its own.

The general conclusion is that our understanding of operations involved in complex derivations is still too limited and at this point complete elimination of both derivational 'look-ahead' and 'look-back' may be an impossible task.

REFERENCES

- Abraham, W., S. D. Epstein, H. Thrainsson and J. W. Zwart (eds.) 1996. *Minimal ideas*. Amsterdam: John Benjamins.
- Baltin, M. 1987. "Do antecedent-contained deletions exist?" *Linguistic Inquiry* 18: 579-95.
- Bański, P. and A. Przepiórkowski (eds.) 2003. *Generative Linguistics in Poland: Morphosyntactic investigations*. Warszawa: Instytut Podstaw Informatyki Polskiej Akademii Nauk.
- Błaszczak, J. 2001. "Covert movement and the genitive of negation in Polish." *Linguistics in Potsdam* 15.
- Bošković, Ž. 1997. *The syntax of nonfinite complementation*. Cambridge, Massachusetts: MIT Press.
- Brody, M. 1995. *Lexico-logical form*. Cambridge, Massachusetts: MIT Press.
- Brody, M. 1997. "Perfect chains." In Haegeman, L. (ed.) 139-67.
- Chomsky, N. 1993. "A minimalist program for linguistic theory." In Hale, K. and S. J. Keyser (eds.) 1-52.
- Chomsky, N. 1994. "Bare Phrase Structure." In Webelhuth, G. (ed.) 385-439.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. 1998/2000. "Minimalist inquiries: The framework." In Martin, R., D. Michaels and J. Uriagereka (eds.) 89-156.
- Chomsky, N. 1999. "Derivation by phase." *MIT Occasional Papers in Linguistics* 18.
- Chomsky, N. 2001. Beyond explanatory adequacy. Manuscript. MIT, Cambridge, Massachusetts.
- Chomsky, N. and H. Lasnik. 1993. "The theory of principles and parameters." In Jacobs, J., A. von Stechow, W. Sternefeld, and T. Vennemann (eds.) 506-69.
- Cinque, G. 1999. *Adverbs and functional heads: A crosslinguistic perspective*. Oxford: Oxford University Press.
- Den Dikken, M. 1995. "Binding, expletives and levels." *Linguistic Inquiry* 26: 347-54.
- Dziwirek, K. 1998. "Reduced constructions in Universal Grammar: Evidence from the Polish object control construction." *Natural Language and Linguistic Theory* 16: 53-99.
- Epstein, S., E. Groat, R. Kawashima, and H. Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: OUP.
- Epstein, S. D. and N. Hornstein (eds.) 1999. *Working minimalism*. Cambridge, Massachusetts: MIT Press.
- Epstein, S. D. and D. Seely (eds.) 2002. *Derivation and explanation in the Minimalist Program*. Oxford: Blackwell.
- Fanselow, G. 2003. "Surprising specifiers and cyclic Spell-out." In Bański, P. and A. Przepiórkowski (eds.) 29-46.
- Fanselow, G. and A. Mahajan. 2000. "Towards a minimalist theory of Wh-expletives, Wh-copying and successive-cyclicity." In Lutz, U., G. Müller and A. von Stechow (eds.) 195-230.
- Fiengo, R. and R. May. 1994. *Indices and identity*. Cambridge, Massachusetts: MIT Press.
- Fox, D. 1999. "Reconstruction, binding theory and the interpretation of chains." *Linguistic Inquiry* 30: 157-196.
- Fox, D. 2001. Antecedent Contained Deletion and the Copy Theory of Movement. Manuscript, Harvard University.
- Franks, S. 2000. "A PF-insertion analysis of *that*." *Syntax* 3.
- Groat, E. and J. O'Neil. 1996. "Spell-out at the LF interface". In Abraham, W., S. D. Epstein, H. Thrainsson and J. W. Zwart (eds.) 113-139.
- Grohman, K., J. Drury and J. C. Castillo. 2000. "No more EPP." *West Coast Conference on Formal Linguistics 19 Proceedings*. 153-166.
- Haegeman, L. (ed.) 1997. *Handbook of generative syntax*. Dordrecht: Kluwer.
- Hale, K. and S. J. Keyser (eds.) 1993. *The view from Building 20*. Cambridge, Massachusetts: MIT Press.
- Heycock, C. 1995. "Asymmetries in Reconstruction." *Linguistic Inquiry* 26: 547-570.
- Hiraiwa, K. 2001. "Multiple Agree and the Defective Intervention Constraint in Japanese." *MIT Working Papers in Linguistics* 40: 67-80.
- Holmberg, A. and T. Hroarsdottir. 2003. "Agreement and movement in Icelandic raising constructions." *Lingua* 113: 997-1019.
- Hornstein, N. 1995. *Logical form: From GB to minimalism*. Cambridge, Massachusetts: Blackwell.
- Hornstein, N. 2000. *Move! A minimalist theory of construal*. Cambridge, Massachusetts: Blackwell.
- Hornstein, N. and J. Witkoś. Forthcoming. "Yet another approach to existential constructions." To appear in *Festschrift for Christer Platzak*.
- Jacobs, J., A. von Stechow, W. Sternefeld, and T. Vennemann (eds.) 1993. *Syntax: An international handbook of contemporary research*. Berlin: Walter de Gruyter.
- Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.
- Kenstowicz, M. (ed.) 2002. *Ken Hale: A life in language*. Cambridge, Massachusetts: MIT Press.
- Kitahara, H. 1997. *Elementary operations and optimal derivations*. Cambridge, Massachusetts: MIT Press.
- Kuno, S. 1997. Binding Theory and the Minimalist Program. Manuscript, Harvard University.
- Larson, R. and R. May. 1990. "Antecedent containment or vacuous movement: Reply to Baltin." *Linguistic Inquiry* 21: 103-22.
- Lasnik, H. 1995. "Case and expletives revisited: On Greed and other human failings." *Linguistic Inquiry* 26: 615-33.

Fanselow's argument for the close, clause internal, cyclic interplay between syntax and phonology is based on the observation that the movable proxy of the VP topic in German is not identified morphologically but it is identified via the Nuclear Stress Rule (NSR), a device of the PF component. Thus, syntactic proxy topicalisation in (iii) cannot apply before the phonological rule of NSR scans the vP constituent for the most deeply embedded phrase (here, [in der Bibel]), assigns stress to it and hereby designates it as a feasible proxy topic. Only then can the syntactic component of grammar resume its operation and move the proxy topic to [spec,C].

- Lasnik, H. and J. Uriagereka. 2001. Essential topics in the minimalist program. Manuscript, University of Maryland, College Park.
- Lebeaux, D. 1988/2000. *Language acquisition and the form of grammar*. Ph.D. dissertation, University of Massachusetts, Amherst. Amsterdam/Philadelphia: John Benjamins.
- Lutz, U., G. Müller and A. von Stechow (eds.). 2000. *Wh-scope marking*. Amsterdam/Philadelphia: John Benjamins.
- Martin, R. 1996. A minimalist theory of PRO and control. Doctoral dissertation, University of Connecticut, Storrs.
- Martin, R. 1999. "Case, the Extended Projection Principle, and minimalism." In Epstein, S. D. and N. Hornstein (eds.). 1-25.
- Martin, R. 2001. "Null Case and the distribution of PRO." *Linguistic Inquiry* 32: 141-166.
- Martin, R., D. Michaels and J. Uriagereka (eds.) 2000. *Step by step: Essays in honour of Howard Lasnik*. Cambridge, Massachusetts: MIT Press.
- McGinnis, M. 1998. Locality in A-movement. Doctoral dissertation MIT, Cambridge Massachusetts. MITWPiL.
- Nissenbaum, J. and D. Fox. 1999. "Extraposition and scope: A case for overt QR." *WCCFL* 18.
- Nuñez, J. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland, College Park.
- Nuñez, J. 1999. "Linearization of chains and phonetic realization of chain links." In Epstein, S. D. and N. Hornstein (eds.). 217-49.
- Pesetsky, D. 1989. Earliness. Manuscript, MIT, Cambridge, Massachusetts.
- Pesetsky, D. and E. Torrego. 2002. "T-to-C Movement: Causes and Consequences." In Kenstowicz, M. (ed.). 132-194.
- Phillips, C. 1996. Order and structure. Doctoral dissertation, MIT, Cambridge, Massachusetts: MITWPL.
- Reinhart, T. 1995. "Interface strategies." *OTS Utrecht University Working Papers*. 1-109.
- Richards, N. 1999. "Featural Cyclicity and the ordering of Multiple Specifiers." In Epstein, S. D. and N. Hornstein, (eds.). 127-58.
- Riemsdijk, H van and E. Williams 1986. *Introduction to the theory of grammar*. Cambridge, Massachusetts: MIT Press.
- Rizzi, L. 1997. "The fine structure of the left periphery." In Haegeman, L. (ed.). 281-337.
- Safir, K. 1999. "Vehicle Change and reconstruction in A-bar Chains." *Linguistic Inquiry* 30: 587-620.
- Stepanov, A. 2000. Late Adjunction and Minimalist Phrase Structure. Manuscript, University of Connecticut, Storrs.
- Svenonius, P. 2001. "Impersonal passives: A phase-based analysis." *Proceedings of the 18th Scandinavian Conference on Linguistics in Lund*.
- Svenonius, P. 2002. "Icelandic case and the structure of events." *Journal of Comparative Germanic Linguistics* 5. 197-225.
- Śmiecińska, J. 2003. Bridge requirements and island conditions on Wh-movement in English: An evaluation of successive cyclic derivation of Wh questions. Doctoral dissertation, Adam Mickiewicz University, Poznań.
- Taraldsen, T. 1981. "The theoretical interpretation of a class of 'marked' extractions." *Proceedings of the 1979 GLOW Conference*. Scuola Normale Superiore, Pisa.
- Toman, J. (ed.) 2002. *Annual Workshop on Formal Approaches to Slavic Linguistics. The Second Ann Arbor Meeting 2001*. Ann Arbor, Michigan: Michigan Slavic Publications.
- Uriagereka, J. 1999. "Multiple spell-out." In Epstein, S. D. and N. Hornstein (eds.). 251-82.

- Webelhuth, G. (ed.) 1994. *Government and Binding Theory and the Minimalist Program*. Oxford and Cambridge, Massachusetts: Blackwell Publishers.
- Willim, E. 1990. "On case marking in Polish." *Papers and Studies in Contrastive Linguistics* 25: 204-220.
- Witkoś, J. 1998. *The syntax of clitics: Steps toward a minimalist account*. Poznań: Motivex.
- Witkoś, J. 2002. "Wh movement and anti-reconstruction effects: Vehicle change and late merge." In Toman, J. (ed.). 311-334.
- Witkoś, J. 2003. "Some notes on single cycle syntax and Genitive of Negation." *International Journal of Arabic-English Studies* 4. 153-173.
- Witkoś, J. (this volume). "Raising expletives."
- Zaenen, A., J. Maling and H. Thrainsson. 1985. "Case and grammatical functions: the Icelandic passive." *Natural Language and Linguistic Theory* 3: 441-83.