Phases and cyclicity

Some problems with phase theory

Günther Grewendorf
University of Frankfurt

Joost Kremers
University of Cologne

October 3, 2007

1 Introduction

In recent years, the notion of *phase* has acquired a central position in the minimalist program (Chomsky 2001, 2004, 2006, forthcoming).[1] In this paper, we examine the notion of phase as it has evolved in Chomsky's most recent papers, putting it in the broader perspective of the minimalist theory. In section 2, we first take a look at the notion of phase, why it was introduced and how it figures in the current theory. In the rest of the section, we look at several other essential points of the theory, all of which are in some fundamental way linked no the notion of phases.

In section 3, we turn to several questions that the current model raises. We first look at the notion of TRANSFER, which raises questions about the way in which the interfaces work. Further issues we discuss are the idea that the No-Tampering Condition can be violated in specific ways, expletives, Icelandic experiencer constructions, and the notion of parallel probing.

The main purpose of this paper is to analyse the framework that Chomsky develops in his latest papers, and to identify potential problems. We offer tentative solutions for some of these problems, but especially the problems surrounding expletives and Icelandic experiencers are left open as problems to be solved.

2 Overview of recent changes

2.1 The notion of 'phase'

In essence, phases are meant to account for the cyclicity of derivations. The fundamental idea of phase theory as developed in Chomsky's recent work is that the derivation of a sentential structure takes place in stages. Phases were originally introduced to solve two empirical problems and one conceptual problem. The first empirical problem involves the derivation of sentences like (1a); the second empirical problem concerns the apparent countercyclicity of Icelandic raising constructions with experiencer subjects and long distance agreement (1b):

(1) a. There is a possibility [that proofs will be discovered]
b. Mér$_i$ þóttu t$_i$ [þær vera duglegar]
me$_{DAT}$ thought$_{pl}$ they$_{NOM,pl}$ be hard-working
'I thought they were working hard'
(Sigurðsson 1996)

[1] For ease of reference, Chomsky's recent papers are referred to with the following abbreviations: MI (Chomsky 2000), DbP (Chomsky 2001), BEA (Chomsky 2004), OP (Chomsky forthcoming), and UGb (Chomsky 2006).

Let us first consider (1a). Since the economy principle Merge over Move would require Merge of the expletive rather than object movement in the embedded passivised clause, such a sentence cannot be derived. The availability of phases opens the possibility that the lexical array of a derivation is subdivided into subarrays that are assigned to particular stages of the derivation. A sentence like (1a) can then be derived if it is assumed that the expletive is not contained in the lexical subarray on the basis of which the embedded clause is derived.[2]

In (1b), the experiencer has moved to Spec,T as a dative subject. The matrix verb agrees in number with the embedded nominative. We know that the trace of the experiencer does not create an intervention effect. However, at the stage of the derivation where T is merged and enters the operation Agree with the embedded nominative, the experiencer is still *in situ* and should create an intervention effect for this operation. In order to avoid the undesirable assumption that the derivation can "look ahead" and anticipate later steps, it is assumed in DbP that the cyclicity of the derivation is evaluated at the phase level: when the matrix phase in (1b) is completed, movement of *mér* to Spec,T has already taken place.

A similar argument is developed in DbP (p.18) for Case valuation of participles in sentences such as (2):

(2) a. There seem to have been caught several fish
b. $[_{\alpha}$ Prt [catch $[_{\text{DO}}$ several fish]]

The number and gender features of the participle can be valued by the direct object *several fish*. The Case feature cannot, however: the Case feature of DO is valued when DO is probed by matrix T. If the φ-features of the participle are deleted after valuation at stage α they are invisible for Matching with probe T. The problem is overcome if Spell-Out takes place at the next strong phase level and the φ-features of the participle remain visible until this level.

In BEA (p.20), further developed in OP, Chomsky adds a conceptual motivation for the assumption of phases: what constitutes a phase is determined by the possibility of Spell-Out (TRANSFER), i.e. by the requirement that the semantic and phonological information relevant to the interfaces is available. What this means is that when valuation takes place, the structure must be spelled out immediately. Valuation of uninterpretable (i.e. unvalued) features makes them indistinguishable from interpretable (valued) features. Given that Full Interpretation requires uninterpretable features to be deleted, deletion must take place immediately upon valuation. However, uninterpretable features may have effects on the sound side (i.e. they may be phonologically realised), so that they must remain until TRANSFER.

Now, if uninterpretable features must be deleted upon valuation and at the same time must remain until TRANSFER, TRANSFER must also take place immediately upon valuation. The only way to achieve this is to assume that all (feature-valuing) operations take place at the same point in the derivation, and that TRANSFER takes place at this point as well. These points in the derivation are the phases.

In essence, phases establish a strict cyclicity in derivations: all syntactic processes take place within the phase. However, on the assumption that C and v^* are the phase heads, a single phase would be too narrow for many operations to take place. Hence, it is assumed that within phase α, the *edge* of the next lower phase β is still accessible for operations in α:

(3) $[_{\text{ZP}}$ Spec [Z ... $[_{\text{HP}}$ Spec [H Comp]]]]
edge

[2]Note that the idea of a numeration or lexical array seems to have died a silent death in recent minimalist theory. Chomsky seems to assume direct lexicon access for Merge. In such a model, this particular argument seems voided.

H and Z being phase heads, the edge of the phase HP (=β) consists of the head H and the specifier(s) of H, as indicated in (3). The elements in this edge are available for syntactic operations that originate in the phase ZP (=α).[3]

The reason for the inaccessibility of the complement of HP to operations in ZP is the assumption that spell-out applies to each phase separately, for reasons just discussed. In earlier formulations of phase theory, the phase complement (or *spell-out domain*) is sent off (*spelled out*) to the phonological component only. This, however, creates an imbalance between the two interpretational systems, the sensory-motor (SM) and the Conceptual-Intentional (C-I) system: while SM receives the structure in chunks, C-I receives it as a whole.

In recent formulations, Chomsky adopts the idea that there is only one computational cycle, which builds structures until a phase has been created. At this point, TRANSFER takes place: the phase complement is transferred to both interface systems, where it is interpreted. As a result, there is no longer a syntactic level of representation that can be called LF, nor can there be operations that take place "at LF". This result, the reduction of the computational process to a single cycle and the abolishment of LF and PF as syntactic representations, is considered by Chomsky to be in line with the *Strong Minimalist Thesis* (SMT), which states that language is an optimal solution to the interface conditions that the language faculty must satisfy (Chomsky forthcoming, p. 2).

As stated, the result of TRANSFER is that the part of the structure that is sent off to the interfaces is no longer available to further operations in the derivation. This inaccessibility has been captured by the *Phrase Impenetrability Condition* (PIC). There have been two different formulations of the PIC:

Version 1: The domain of the head H of a strong phase HP is not accessible to operations outside HP; only H and its edge are accessible to such operations (MI).

Version 2: The domain of H is not accessible to operations at the next highest strong phase ZP; only H and its edge are accessible to such operations (DbP, BEA).

The difference between the two definitions mainly lies in the abilities of non-phase heads. Consider (4):

(4) $[_{CP}$ C $[_{TP}$ β [T $[_{v*P}$ α [$v*$ YP]]]]]

According to version 1 of the PIC, the domain (i.e. complement) of $v*$ is not accessible to operations outside $v*$P. That is, T in (4) can probe α and $v*$, but not YP, nor anything in YP. As a result, elements from YP must be moved to Spec,$v*$ in order to be accessible to T. In version 2 of the PIC, this is not the case: YP only becomes inaccessible when the next higher phase head, in this case C, is merged. As a result, T can probe (an element in) YP, so that it can agree with a quirky nominative object in the lower phase, as seen in Icelandic.

In OP, Chomsky reinterprets the notion of *phase* in such a way that version 1 of the PIC automatically follows. This leads to certain problems with the Icelandic quirky nominative object constructions, as discussed in section 3.3.

2.2 TRANSFER

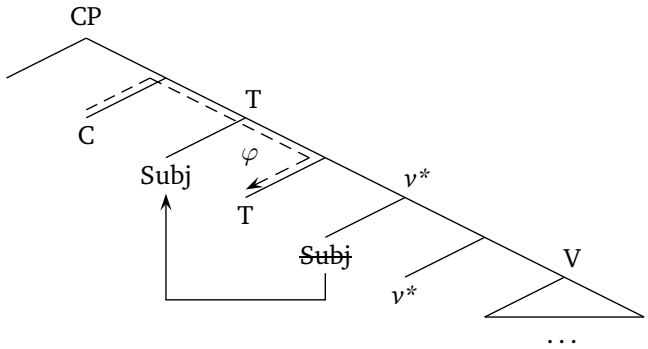
In OP, the notion of a phase is directly linked to TRANSFER operations, which operate at various stages of the computation and hand the syntactic object already

[3]Provided that there is no intervening phase head, of course.

constructed to the phonological interface, which maps it to the sensory-motor component ("spell-out" in the original sense), and to the semantic interface, which maps it to the conceptual-intentional component. As noted above, it is a consequence of these TRANSFER operations that the final internal level LF can be eliminated (OP, p. 8). Thus the TRANSFER operations constitute the phases and "in the best case", the phases will be the same for both TRANSFER operations.

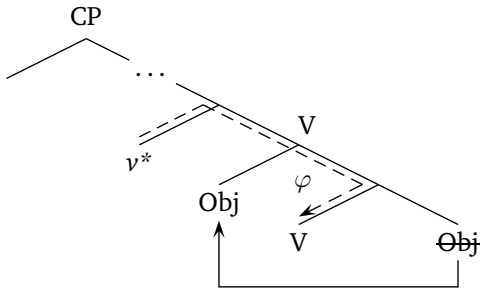
The importance of phases is increased even more in BEA and OP: Chomsky argues that *all* operations take place at the phase level. Consequently, internal Merge should be driven only by phase heads (see also BEA p. 14, where valuation of uninterpretable features is regarded to be part of TRANSFER). The T head, Chomsky claims, does not have φ-features in and of itself: it only has φ-features when it is selected by C. That is, T lacks φ-features in the lexicon, and inherits them from the C head in syntax. This means that ultimately, the Agree process that values the φ-features of T and the subject's Case feature, and optionally raises the subject, is initiated by C, not by T. Nonetheless, when the subject moves, it moves to Spec,T, not to Spec,C, because of the fact that the φ-features are passed on to T:

(5)



The relation between C and T is mirrored by the relation between v^* and V. Chomsky assumes that like C, v^* has a set of Agree features (essentially φ-features), which are inherited by V. As a result, object agreement is handled not by v^*, but by V. The object therefore moves to Spec,V rather than to Spec,v^*:

(6)



The main reason for adopting this assumption seems to be a desire to treat both C and v^* phase heads in an identical manner. Note, however, that this move does have the advantage that the base position of the subject is no longer hierarchically lower than the agreement position of the object, so that Agree between T and the subject can be established without needing to account for the absence of intervention effects. It also means that in ECM-constructions, the embedded subject can move into the matrix clause, an idea that goes back to Postal (1974). A disadvantage of this move seems to be that object movement in German cannot be analysed in terms of the operation Agree (i.e. as Case driven movement):

(7) a. weil der Student den Computer nicht repariert hat
since the student the$_{ACC}$ computer not fixed has
'because the student did not fix the computer'
b. *weil der Student nicht den Computer repariert hat. (only narrow reading)

In (7a), the object *den Computer* 'the$_{ACC}$ computer' has moved from its VP-internal position to a position higher than the negation. This movement differs from so-called Object Shift (OS) in Scandinavian languages, so it cannot be analysed in an analogous manner.[4] On the assumption that object agreement takes place in Spec,*v**, it could be argued that German object movement differs from OS because it is case-driven, but such an account is now no longer available.[5]

2.3 Edge feature and movement

Merge is, obviously, still the single most important operation in the computational system. Chomsky argues that Merge is driven by a feature of lexical items (LIs) that he calls the *edge feature* (EF). The edge feature enables a lexical item to be merged with another element. EF takes a very prominent role in the new framework, being the driving force for Merge. It is responsible for the fact that a lexical item can be merged with its complement:

> For an LI to be able to enter into a computation, merging with some SO (and automatically satisfying SMT), it must have some property permitting this operation. A property of an LI is called a feature, so an LI has a feature that permits it to be merged. Call this the edge-feature (EF) of the LI. [...] When merged with a syntactic object SO, LI forms {LI, SO}; SO is its complement. (OP, p. 6)

That is, a lexical item can be merged with its complement by virtue of having an edge feature. Further mergers are also possible: EF of *v** enables it to be merged with a second SO, which forms the subject. That is, while it is valid to say that the merger of *v** and VP *fulfills v**'s EF, this does not mean that EF is deleted or inactivated. It remains active, and may trigger additional Merge operations. The same is obviously true for C: EF enables C to be merged with its complement TP, and also enables further Merge operations that fill the Spec,C domain.

It seems reasonable to assume that this liberty extends not only to phase heads, but to non-phase heads as well, at least to lexical categories: a verb may need to merge with more than one internal argument, and nouns, if we do not analyse all noun-phrase dependents as adjuncts, occasionally do as well. Nonetheless, phase heads are different with respect to the edge feature, in that they are the only types of heads that can satisfy their EF through *Internal Merge* (IM).

Internal Merge, as opposed to *External Merge* (EM), is the merger of the root of an SO_1 with an SO_2 that is already part of SO_1. IM results in *one* SO being merged in *two* different positions in the same tree. When the structure is spelled out, only one instance is phonologically realised. Usually, though not necessarily (cf. Nunes 2004), this instance is the highest one.[6]

The movement of the subject to Spec,T and the movement of the object to Spec,V are also instances of IM, satisfying EF of T and V, respectively. However, this is only

[4]Most importantly, OS obeys Holmberg's Generalisation (Holmberg 1986), which states that OS is dependent on V movement, which is not the case for object movement in German.

[5]An account in terms of Case/Agree might be considered problematic anyway, because indefinite objects that appear before negation are interpreted as specific, so that object movement in German may in fact be a form of scrambling.

[6]Chomsky expresses some surprise as to the confusion this so-called *copy theory of movement* has caused. Although Chomsky usually speaks of "two copies" of an element, it may be better to speak of "two instances" of an element.

possible due to the Agree features that T and V inherit from the phase heads C and *v**. In this way, it becomes possible to define the A/A′-distinction derivationally: A′-movements are movements triggered by EF on the phase heads, while A-movements are movements triggered by EF "by proxy", i.e. EF on T or V after inheriting Agree features from a phase head.[7] As Chomsky states:

> For our purposes here, it will suffice to define an A′-position as one that is attracted by an edge-feature of a phase head; hence typically in SPEC-C or outer SPEC of *v**. Others are A-positions. From this point of view, A- and A′-positions are distinguished not by their structural status within a phrase-marker, but by the manner in which they are derived. (Chomsky forthcoming, p. 15)

Defining the A/A′-distinction in this way satisfies *Full Interpretation* (FI), in that no extra machinery is needed to distinguish A-positions from A′-positions. As the disinction is relevant in syntax (which it obviously is, given the different properties that A- and A′-movement show), SMT leads us to expect that it is relevant for C-I.[8] A-positions are related to Case and ϑ-roles, while A′-positions are related to a series of interpretative properties of clauses and their constituents. Obviously, then, the A/A′-distinction is relevant at C-I, so that the distinction in syntax conforms to SMT.

Thus, in OP, the A/A′-distinction provides the motivation for feature inheritance: the inherintance of Agree features from C to T establish the A/A′-distinction in syntax. However, as Richards (2007) points out, this motivation may be suggestive, but does not have "the force of necessity" (p. 564). As Richards argues, establishing the A/A′-distinction would just require two different features, which we actually have: EF and Agree features. The possibility of multiple specifiers means that there should be no need to pass the Agree features on to T: the subject could simply move to Spec,C and satisfy the Agree features, still leaving open the possibility that further elements move to Spec,C also, satisfying EF.

Richards provides an alternative motivation for feature inheritance. He argues that the need for the inheritance of Agree features from C to T follows from two standard assumptions in phase theory:

(8) a. Value and Transfer of uFs must happen together. (p. 566)
    b. The edge and nonedge (complement) of a phase are transferred separately. (p. 568)

The rationale behind (8a) is that uninterpretable features are unvalued when they enter the derivation, and must be valued during it.[9] Richards now states:

> (...) once valued by Agree, they [uFs] are indistinguishable from interpretable (i.e., lexically valued) features, without reconstructing the derivation. It follows that uFs must be spelled out (transferred) as soon as they are valued if the system is to avoid lookback. If Value takes place before Transfer, the derivation will crash at Sem; if Value takes place after Transfer, then the derivation will crash at both interfaces (since unvalued features cannot be interpreted). Value must therefore be part of Transfer (these operations are, in this sense, "simultaneous"); (Richards 2007, p.566)

[7] That is, EF replaces the EPP to a large extent. There are still some "residual" EPP effects, as Chomsky notes (OP, p. 21).

[8] Chomsky argues that the C-I interface is primary, claiming that language first and foremost developed as a means to structure and formulate thought. Language, he claims, is in many ways poorly designed for communication (cf. ambiguity, garden paths, etc.), suggesting that it is a secondary function.

[9] The notation uF is meant to imply both uninterpretable and unvalued.

In short: once uF is valued, it must be deleted and the structure containing it transferred *immediately*. Note that this consideration implies that all operations are triggered by the phase heads: if T were to value the subject's case feature and its own Agree features independently of C, TRANSFER of those features could not take place immediately.

(8b) is a direct consequence of the PIC. The derivation is cyclic, in that each phase is transferred to the interfaces separately. But since the PIC states that a phase head Ph and the edge of Ph are accessible to operations outside the phase, they can only be transferred when the next higher phase is completed. As a result, the complement of Ph is transferred when Ph is complete, but Ph and its edge are transferred in the next cycle.[10]

Richards argues that the two assumptions in (8), although both standard assumptions in phase theory, are incompatible as they stand: uninterpretable features on C are valued through the standard mechanisms, but once this has happened, C itself is *not* transferred. Only TP is. The uninterpretable features on C, now valued, therefore remain in the derivation. Feature inheritance provides a mechanism to avoid this conflict. If the uninterpretable features on C are transferred to T, they are immediately transferred upon valuation.

This brings us to another refinement in the system. All interpretative properties are assigned at the C-I interface. While this may seem a self-evident statement given the theoretical assumptions behind the minimalist program, the assumption has consequences that have hitherto not been taken into account.

For one, ϑ-role assignment is not a process that takes place in core syntax. Rather, the C-I interface interprets certain structural configurations as configurations for ϑ-assignment. These configurations are those that are created through EM with a ϑ-role assigning head. This reinterpretation of ϑ-role assignment goes hand in hand with the new definition of the A/A′-distinction: it is no longer possible to identify an A-position by the fact that a ϑ-role or Case was assigned to it, but at the same time, it is no longer necessary to identify an A-position in this way.[11]

Another domain whose structure is reconsidered is the C-domain. Given the fact that the phase heads can satisfy EF through internal merge, there is no longer any need to postulate *wh*-features, Topic or Focus features, or Q features, etc. on C to trigger movement of the relevant phrases to the C-domain. All these elements are raised merely by the fact that C has an edge feature, enabling it to merge with any number of phrases. This move eliminates a slew of features that were thought to trigger Agree operations for which there is in fact little or no morphological evidence, contrary to subject-verb (and object-verb) agreement, for which the morphological evidence is abundant.[12]

Instead of having *wh*-, Topic, Focus etc. features, there is just EF, which allows the C head to raise any element to its Spec. The interpretative properties that A′-movements to the C-domain cause are established at the C-I interface, and are not encoded in the syntactic structure. That is, as Chomsky puts it, "what is raised is identified as [e.g.] a topic by the final position it reaches, and any extra specification [i.e. a topic feature] is redundant." (OP, p. 16). One consequence of this is that minimality does not play a role anymore in A′-movement. Simply put: if minimality

[10] As we will see is section 3.3, this assumption causes problems for Long Distance Agree in e.g. Icelandic experiencer constructions.

[11] Note that Case positions are somewhat problematic in this view. They do not seem to have any import at the C-I interface. If ϑ-roles are assigned at the foot of a chain, and further interpretative features at the head (i.e. in the CP-domain), case positions seem irrelevant. Their relevance would seem to be restricted to SM, where they have the communicative function of signalling structural positions from which the hearer can reconstruct ϑ-assignment.

[12] Additionally, it has not been clear why e.g. *wh*-features on C should be uninterpretable, which they would need to be, given that C initiates the raising of *wh*-elements.

would apply, topicalisation of an object would be blocked by the presence of a subject.[13]

What Chomsky argues for is that the interpretative properties that A′-movements to the C-domain cause are established at the C-I interface, and are not encoded in the syntactic structure. It is not entirely clear, however, how the identification of the correct interpretation can take place. Either we must assume that C is a shorthand for a variety of functional heads (c.f. Rizzi 1997), and that each of these has an EF feature that can (but does not have to) attract elements from the clause. Alternatively, we may assume that the proper interpretation takes place wholly in the C-I module: the position of a raised element then merely coveys that it is a topic, or a focus, etc., and C-I has to determine which interpretation fits the context best.[14]

Another consequence of this system is that a clause containing a *wh*-element may arrive at the C-I interface without this element having moved to Spec,C. The derivation may simply have failed to raise the element to Spec,C. Since there is no unvalued feature on the *wh*-element, the derivation will not crash. In such a case, the C-I interface will still attempt to assign this structure an interpretation, and it will succeed in doing so if the context allows an echo-question.

Essentially, this means that the computational system overgenerates. It generates structures without being restricted by the interpretation those structures will eventually be assigned. Not all generated structures can be assigned an interpretation at the interfaces, but those structures cannot be filtered out by the computational system. Only the interfaces themselves can do so. One consequence of this is that it becomes possible to generate structures that are *deviant*, that is, structures that would be considered "incorrect" under normal circumstances, but may become appropriate in specific contexts.[15]

The idea that A′-movement is not triggered by features but merely by EF raises questions for the successive-cyclic nature of movement. We know that A′-movement is successive cyclic,[16] but if A′-movement is not triggered by features, the intermediate movement cannot be triggered by features either. The alternative, argued for by e.g. Boeckx (2003), is to assume that successive cyclic movement is driven by the requirement that movement has to proceed in "local steps". This idea can be seen as a return to the familiar *Shortest Move* condition (also called *Minimise Chain Links* or *Fewest Steps*) on the operation *Form Chain* (see OP p.21, fn.57). How small these local steps are remains to be clarified (for A′-movement possibly every category, as proposed in Chomsky 1986).

This account of successive cyclic movement in terms of "local steps" predicts that A′-movement passes through intermediate positions but cannot stop there to be spelled out (see OP p. 21). The account thus provides us with an explanation for an unsolved problem with Japanese long scrambling from finite clauses. Saito (1994) has shown that although VP is a possible landing site for scrambling in Japanese and although long scrambling has to pass through an intermediate VP-position of the matrix clause, long scrambling can only target the matrix IP rather than the VP (the so-called *VP-adjunction paradox*):

[13] A specific result of this is that there are no superiority effects in *wh*-movement. In languages that show such effects, e.g. English, they must be explained in some other manner.

[14] Chomsky (OP p. 9) states: "…where C is shorthand for the region that Rizzi (1997) calls the "left periphery," possibly involving feature spread from fewer functional heads (maybe only one)", which tends toward the former approach, but it seems to us that the issue is far from resolved.

[15] Perhaps this could even be taken so far as to render ungrammatical structures acceptable in the proper context, e.g. in Foreigner Talk registers.

[16] E.g., from reconstruction effects (Barss 2001), or from morphologically visible agreement at stages intermediate between probe and goal (cf. Boeckx 2003).

(9) a. [$_{IP}$ sono hon-o [$_{IP}$ John-ga [$_{VP}$ Bill-ni [$_{CP}$ Mary-ga *t* motteiru to]
that book-ACC John-NOM Bill-DAT Mary-NOM have C
itta]]] (koto)
said (fact)
'that book, John said to Bill that Mary has'

b. ??[$_{IP}$ John-ga [$_{VP}$ sono hon-o [$_{VP}$ Bill-ni [$_{CP}$ Mary-ga *t* motteiru to]
John-NOM that book-ACC Bill-DAT Mary-NOM have C
itta]]] (koto)
said (fact)
'that book, John said to Bill that Mary has'
(Saito 1994, p.265)

Now if intermediate steps are independently triggered, the VP would have to contain something like an EPP feature. But then we cannot explain why long scrambling to VP is out. However, if local steps are not triggered by features, intermediate copies cannot be spelled out so that this problem is avoided.

On the other hand, another empirical problem arises from languages such as German where intermediate *wh*-copies in a non-*wh*-Comp can be spelled out as in the case of partial *wh*-movement,[17] as in (10a), and the so-called "copy construction", as in (10b):

(10) a. Was glaubst du wen Maria liebt?
what think.2sg you who$_{ACC}$ Mary likes
'Who do you think Mary likes?'

b. Wer glaubst du, wer du bist?
who$_{NOM}$ think.2sg you who$_{NOM}$ you are
'Who do you think you are'?

A possible solution might be found here by assuming that under certain conditions multiple copies of an element can be (partially) spelled out (cf. Fanselow & Çavar 2002 for such a proposal).

2.4 Simultaneous application of operations

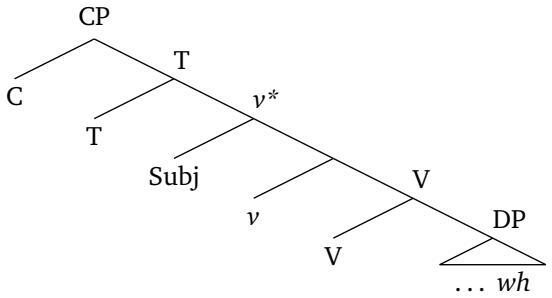
The assumption that all operations are triggered by the phase heads leads to another, in some ways rather remarkable, step: based on certain extraction cases, Chomsky argues in OP (but already hints at the possibility in BEA, p. 14) that operations (specifically, raising to Spec,T and raising to Spec,C) may proceed *in parallel*. Take the following cases:

(11) a. [of which car]$_i$ did they find they find [the driver t_i]?
b. *[of which car]$_i$ did [the driver t_i] cause a scandal?

The contrast in (11) is the standard contrast between extraction from object, which is allowed, and extraction from subject, which is ungrammatical. Structurally, the distinction is clear. (11a) has the following structure:

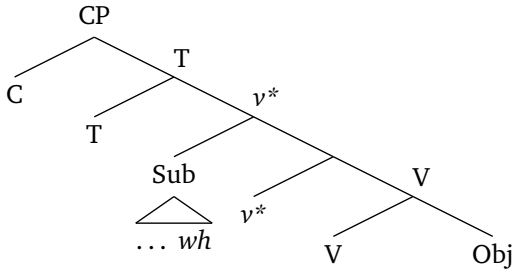
[17] Properties: anti-locality, not possible with subject clauses, sensitive to the *Complex NP Constraint*, not possible with factive verbs

(12)



Extraction of the *wh*-argument from object position is possible in (12), passing through Spec,*v**. Consider now (11b), which has the structure in (13):

(13)



Here, the *wh*-element is embedded in the subject, and although this subject is in the phase edge, extraction from it is not possible.[18] The example in (14), however, shows that the posibility of extraction depends on the base position of the element being extracted from:[19]

(14) [of which car]$_i$ was [the driver t_i]$_j$ awarded t_j a prize?

Here, the DP from which extraction has taken place is in subject position, the same position from which extraction was impossible in (13). In this case, extraction is grammatical, however. We can account for this by assuming that extraction to the CP domain proceeds from the base position:

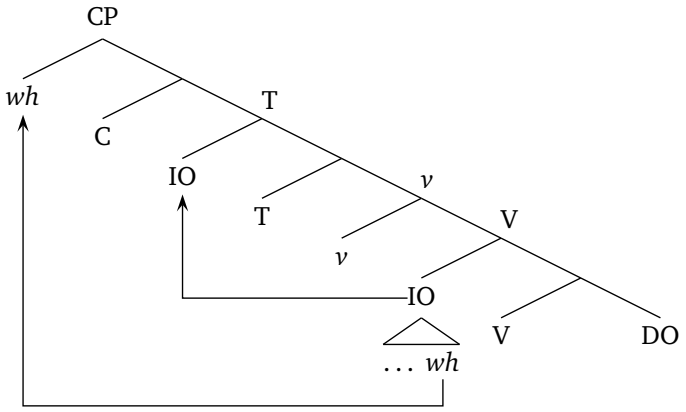
[18]Chomsky speculates that the reason for this subject-island effect lies in the fact that the "search (. . .) goes too deeply into a phase already passed" (OP p.18). The relevant phase would seem to be the DP, because the DP itself is in the edge of *v**P, and therefore can hardly be considered "too deeply" embedded. Note, however, that if DP is a phase, extraction out of the object also passes a phase boundary. We would need something like Nunes & Uriagereka's (2000) proposal for parallel workspaces to make a distinction along these lines.

[19]Note that the data is not entirely as clear as (14) suggests. Chomsky (1995a, p. 412) gives the following example to show that *wh*-movement takes place after passivisation:

(i) *Who$_i$ was [a picture of t_i] taken by Bill?

If (i) is ungrammatical, it is not immediately obvious why (14) is acceptable.

(15)



However, such a derivation would be problematic for the No-Tampering Condition: "Merge of X and Y leaves the two SOs unchanged" (Chomsky forthcoming, p. 5). The movement of the indirect object to Spec,T cannot take place after the movement of the *wh*-element to Spec,C, as the merger of IO with T would no longer be at the root of the tree. On the other hand, it is also impossible for the movement of IO to take place *before* the movement of *wh* to Spec,T. Once IO has moved, it is in subject position, from which extraction is ungrammatical. Since the movement of IO to Spec,T cannot take place before nor after the *wh*-movement, Chomsky's conclusion is that the two movement operations must proceed in parallel.

This conclusion is backed up by two theoretical motivations. First, if T inherits its Agree features from C, as Chomsky has assumed for independent reasons, T *cannot* even initiate any movement operation before C has been merged and the Agree features have been passed on to T. Or, more precisely, T cannot satisfy its EF through internal merge until it has inherited the Agree features from C that trigger the movement operation.[20]

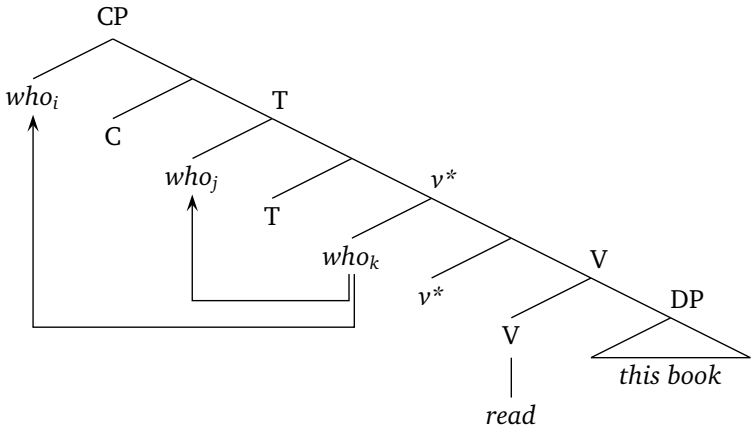
The second theoretical motivation involves the *Generalised Inactivity Condition* (OP p.16). Since the earliest stages of minimalist theory, it has been considered a natural assumption that once an element has valued all its features, it is no longer accessible for further A-movements. Chomsky now generalises this notion, and argues that the head of an A-chain is invisible for *all* types of operations: not only can it not undergo further A-movement, A′-movement is barred as well. This move obviously relates to the assumption that A′-movement is triggered by EF, not by unvalued features, which means that the head of an A-chain has all its features valued, making further movement impossible. It also relates to the copy theory of movement: once movement is construed as copy+merge, the notion of an element moving its way up the tree is no longer valid, and simultaneous probing (and copying) of an element α by two different probes can only be excluded through stipulation.

Chomsky still maintains the idea that an element can only be probed if it has unvalued features, however. As a result, A′-movement of a phrase α cannot take place once α has all its features valued and forms a complete A-chain. Therefore, A′-movement must take place before or simultaneously with the highest A-movement that α undergoes.

The idea that A-movement and A′-movement take place in parallel is adopted generally. It also holds for e.g. subject-*wh*:

[20] As already mentioned in footnote 7, there are still some residual EPP effects; specifically, cases in which a T head not embedded under a C head (e.g. an infinitival T) attracts a copy of the subject in its specifier, which is, however, not spelled out (i.e., in the old model, cases in which the subject moves *through* an infinitival Spec,T on its way up to a higher position).

(16)



(16) represents a simple subject-*wh* question. The *wh*-element *who* is raised to Spec,T from its base position. Simultaneously, C attracts the *wh*-subject. The position C targets is the base position of the *wh*-element, not Spec,T.

In its base-generated position, the *wh*-subject has an unvalued case feature. This feature is valued by the attraction to Spec,T. As a result, the *wh*-element has no unvalued features left, and is inactivated, and could therefore not raise further to Spec,C.

After the movement operations have taken place, the structure contains two A-chains and one A′-chain:

(17)

A-chains:	(who_j, who_k)
	(who_k)
A′-chain:	(who_i, who_k)

Although the A-chains are invisible and although there is no direct relation between the operator who_i and the head who_j of the A-chain (who_j, who_k), there is evidence from familiar properties of A-movement (binding, scope, Weak Crossover) that there really is a copy in SpecT which heads the two-membered A-chain in (17):

(18) a. Who_i seems to his_i friends to be preferable
b. *Who_i do you seem to his_i friends to prefer

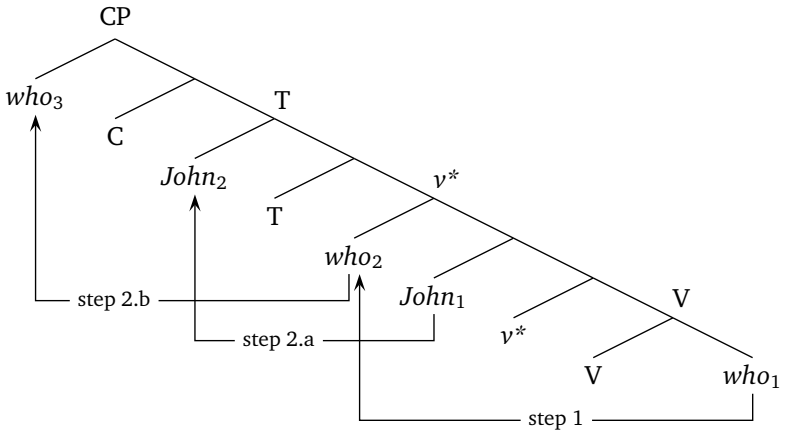
The lack of a Weak Crossover effect in (18a) is due to the fact that the pronoun is A-bound by an element in Spec,T.

A further conceptual argument for the parallel application of operations can be seen in the derivation of sentences such as (19):

(19) who did John see?

According to OP (p.18) the derivation proceeds as represented in (20):

(20)



Given PIC, the *wh*-object who_1 first has to raise to the edge of the internal phase v^* (step 1) in order to be accessible to the phase head C ("local step"-condition). The Agree-feature of C, inherited by T, probes the subject $John_1$ and raises it to Spec,T (step 2.a), and simultaneously the edge feature of C seeks the *wh*-object in the outer Spec of v^* and raises it to Spec,C (step 2.b). The argument for parallel probing by C is provided by the observation that if T were a phase head, or an independent probe (as assumed in DbP), "then raising of subject to SPEC-T would be blocked by intervention of the φ-features of *who* in the outer SPEC of v^*" (OP, p. 18). However, since T is not a phase head and since both operations are driven by the phase head C in parallel, the problem does not arise.

Recall that in DbP the problem posed by (20) was solved by assuming that the probe-goal relation is evaluated for the Minimal Link Condition at the next higher strong-phase level (after it is known whether the outer edge of v^*P has become a trace) (DbP p.27). In OP the derivation of (20) is actually intended to show that A′-chains behave like A-chains with respect to intervention. Since movement of the subject crosses the lower copy of the *wh*-chain (who_2) in the A′-position Spec,v^*, it is concluded that as in the case of A-chains, only the full A′-chain (equivalently its head) creates an intervention effect.

Note, however, that the derivation in (20) only shows that the lower copy of an A′-chain does not create an intervention effect for Agree. It does not yet show that the full A′-chain does in fact create such an effect.

Note that not only is it possible for both C and T to target an element in its base position, it is also possible for both types of operations to interweave. Take (21):

(21) of which car_j is [[the driver t_j]$_i$ [t_i likely to [t_i cause a scandal]]]?

We know that in (21), the A′-extraction of *of which car* cannot proceed from subject position (Spec,T). Because the DP originates inside the infinitival complement of the adjective *likely*, we also know that extraction of the *wh*-element cannot proceed from the DP's base position. Extraction therefore must have taken place from an intermediate position.

3 Some points of discussion

3.1 TRANSFER

The approach sketched by Chomsky in his most recent paper raises several interesting questions, which deserve further attention. One question is raised by the operation TRANSFER. At certain points during the derivation, chunks of structure are sent

off to the interfaces. We may ask how the interfaces deal with these chunks. If we look at the SM-interface, for example, we note that the chunks that are received by the interface do not necessarily correspond to any sort of unit in phonology.

The phonological unit that interacts with syntax is the phonological phrase (φ). The basic correspondence rule is that each lexical XP in syntax corresponds with a phonological phrase. There are various additional rules that deal with functional material and with lexical elements that do not project fully-fledged phrases (e.g. Nespor & Vogel 1986, Truckenbrodt 1999, in press).

Consider the Italian phrase in (22), taken from Nespor, Guasti & Christophe (1996, p. 9):

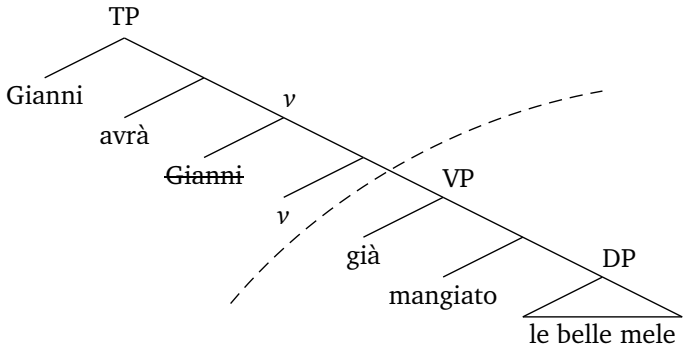
(22) [$_{IP}$ Gianni avrà [$_{VP}$ già mangiato [le belle mele]]]
Gianni will.have already eaten the good apples
'Gianni will have already eaten the good apples.'

The phonological structuring of (22) is given in (23), with phonological phrases delimited by curly braces:

(23) { Gianni } { avrà già mangiato } { le belle mele }

Given the syntactic model outlined in the previous section, the structure of (22) is roughly the one in (24), with the transfer domain (the complement of the phase head) indicated by the dashed line:

(24)



The transfer domain of the lower phase head *v* excludes the auxiliary verb *avrà*. The phonological phrase that *avrà* belongs to, however, is not the one formed from *Gianni*, but the one formed from the main verb *mangiato*. That is, the phonological structure can only be assigned after the higher phase has been transferred.

Fuß (2007) discusses this problem, and proposes a solution in terms of *phonological domains*. A phonological domain is a syntactically defined domain that serves as input for the phonological component. Fuß defines the phonological domain in such a way that it becomes possible to relate phonological phrases to syntactic structure. Yet, this definition appears *ad hoc*: it is primarily motivated by the need to capture the proper domain. No motivation in terms of SMT is provided, and none seems available.

We can envisage a tentative solution, however, that does not take recourse to any *ad hoc* syntactic (or for that matter, phonological) notion, if we take into account the correspondence rules between syntax and phonology. The VP, according to standard assumptions, corresponds to a φ. The object may either be incorporated into this φ, or, if it is complex enough, may form its own φ, which is the case here. That is, after the first TRANSFER operation, the following phonological structure is derived:

(25) già mangiato le belle mele
{ ω ω } { ω ω }

In the next phase, the subject *Gianni*, being an XP, forms a φ as well. The auxiliary *avrà* is not an XP, so does not correspond to a φ. Rather, it is a head and as such corresponds to a prosodic word (ω). Therefore, the second TRANSFER operation yields the following structure:

(26) Gianni avrà
{ ω } ω

Because a phonological structure encompasses *all* phases of the structure being derived, it seems natural enough to assume that the SM module assembles the different parts, forming (27):

(27) Gianni avrà già mangiato le belle mele
{ ω } ω { ω ω } { ω ω }

Here, *avrà* corresponds to a prosodic word that is not part of any φ. Crucially, phonological structure prohibits this: a prosodic word ω must be incorporated into a phonological phrase φ.

The prosodic word $avrà_\omega$ is adjacent to two φ's, and could in principle be incorporated into either. Basic rules of directionality[21] state that $avrà_\omega$ is to be incorporated to the φ on its right, $\{già\ mangiato\}_\varphi$:

(28) Gianni avrà già mangiato le belle mele
{ ω } { ω ω ω } { ω ω }

Such restructuring operations are common in theories on phrasal phonological structure, and can therefore be safely assumed for our purposes here. Many issues obviously remain, both on the syntactic and on the phonological side, but for fundamental cases, the phase model does not seem to pose any problems for the syntax-phonology interface.

Similar questions arise with the C-I interface. Stechow (2005) argues that LF cannot be interpreted cyclically, and that interpretation is only possible at the end of the derivation. However, he also argues that *building* the LF structure can be done cyclically.[22] With LF abolished, interpretation of a phrase structure is delegated to the C-I module. It is the task of the C-I *interface* to provide the C-I module with a structure that it can interpret. This structure is a semantic structure derived from the syntactic phrase marker.

Consider the phrase in (22). The VP, complement of *v**P, is transferred first, and corresponds to the predicate of the structure. Lambda extraction allows us to say that the C-I interface builds a structure similar to (29) out of it:

(29) $\lambda x(\mathbf{eat}(x, \iota y(\mathbf{apples}(y) \wedge \mathbf{good}(y))))$

The next TRANSFER operation sends off the rest of the structure, providing on the one hand an argument for the lambda formula in (29), and on the other tense and aspect marking. Assembling the complete structure from this information plus the formula in (29) seems straightforward enough, and should pose no serious challenge to the phase model of syntax. Once the C-I interface has constructed the semantic representation, it can be sent off to the C-I module, where it is interpreted as a whole.

[21] Which are language-dependent, cf. Truckenbrodt (1995), Nespor et al. (1996)

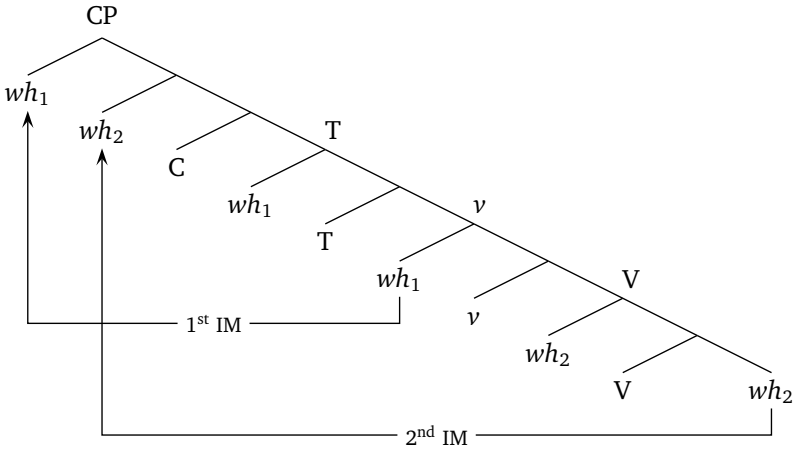
[22] One prerequisite for this, Stechow argues, is that the grammar has indices. We leave this matter aside.

3.2 No Tampering

Another issue to consider is what the idea of parallel movements means for the No-Tampering Condition. Chomsky argues that movement of the subject to Spec,T cannot take place before C is merged, because T inherits its Agree features from C. But in this configuration, C being merged in the tree, the movement of the subject to Spec,T no longer targets the root of the structure.

Chomsky himself points to Richards's (2001) analysis, and argues that his analysis can "capture some of the properties of 'tucking in' [...], taking the 'edge' to be the position as close as possible to the probe." (OP, p. 7). This remark essentially relates to the notion of multiple specifiers and observations suggesting that in multiple-specifier configurations, the element that is attracted later is positioned closer to C than elements attracted before it:

(30)

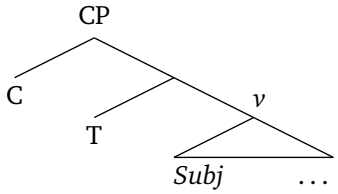


Ignoring for the moment the Agree-movements of both object and subject, we know that in languages such as Bulgarian, all *wh*-elements move to the C-domain overtly. The general observation is that the highest *wh*-element in C's search domain, i.e the element that is attracted first (by standard assumptions on probe-goal relations), is also the highest specifier of C. That is, the element that is attracted second (third etc.) is not merged at the *root* of the tree, but "tucked in" between C and the specifier(s) already present. In Richard's formulation, the second (third, etc.) specifier is merged at the "edge" and "as closely as possible" to the probe C.

That is, if the No-Tampering Condition is understood to require merger at the edge, with "edge" being defined along standard ways (the specifier domain of C and the C head itself), we obtain an account of multiple specifiers and tucking in that conforms to the NTC.

Movement of the subject to Spec,T, Chomsky suggests, can be handled in the same way. The configuration after the merger of C but before movement is (31):

(31)



Given that movement of the subject to Spec,T is triggered by Agree features present on T but inherited from (and inherent to) C, it seems reasonable to assume that the "edge" here includes both C and T, not just C. But T being the actual probe, the subject is merged as closely as possible to it, i.e. in Spec,T.

However, we must ask ourselves if this is indeed the most natural interpretation of the phrase "merging as closely as possible to the probe". Adopting the usual set-notation, the structure before movement of the subject is (32):

(32) {C, {C, {T, {T, *v*P}}}}

Similarly, in set notation, the general form of Merge is (33), with the order of A and B irrelevant:

(33) $A \oplus B = \{A, \{A, B\}\}$

Now, merging an element "as closely as possible to the probe", with the probe being T in this case, would result in the following operation (again, with the order of T and *Subj* being irrelevant):

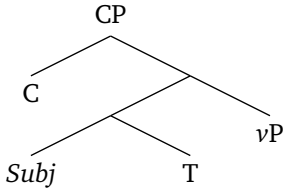
(34) *Subj* $\oplus$ T = {T, {*Subj*, T}}

Substituting the result in (34) for the head T in (32) yields the following structure:[23]

(35) {C, {C, {T, {{T, {*Subj*, T}}, *v*P}}}}

This, however, is not the structure that we want. It is equivalent to the following tree diagram, with the subject in the complement position of T, and the *v*P elevated to specifier:

(36)



Therefore, in order for "tucking in" to work, it seems necessary to include a requirement that IM does not target the head itself but always its first projection. In fact, Richards includes such a requirement when he says that a specifier must be merged "at the edge". This, however, amounts to reintroducing the specifier-complement distinction as a primitive notion of the theory, which Chomsky aims to do away with.

It seems, then, that parallel probing leads to an elaboration of the definition of Merge that is conceptually questionable, and might better be avoided.[24] In section 3.5, we suggest an alternative to the parallel probing mechanism that avoids this problem.[25]

3.3 Icelandic experiencer constructions

Further questions are raised by Chomsky's discussion of certain Icelandic dative-nominative experiencer constructions, exemplified in (37):

[23] Note that it is not clear whether such a substitution operation, which is required in some form or other if we adopt the notion of "tucking in", is even possible without augmenting the basic definition of Merge in (33).

[24] In fact, Chomsky himself discusses these issues (MI, p. 136), albeit in a slightly different context. Essentially, Chomsky argues that what he calls *local merge*, i.e. merging as closely as possible to the label of an SO, may be needed to account for head movement, but must be excluded for other forms of merge.

[25] This problem not only arises with parallel probing, of course, but also with the notion of "tucking in". An alternative analysis of multiple *wh*-fronting is the *wh*-cluster model developed in Grewendorf (2001).

(37) a. to someone seems [Olaf be intelligent]
b. [C [T [DAT [*v** NOM]]]] . . .

Chomsky points to these constructions as a further piece of empirical evidence for parallel probing, but closer examination shows that there still remain some problems that need to be resolved.

As shown in (38), if the experiencer of the raising verb remains *in situ* (which Chomsky assumes to be Spec,*v**, cf. OP p. 19) in an expletive construction, it blocks long distance agreement of T with the embedded subject:

(38) a. Það virðist/*virðast einhverjum manni [hestarnir vera seinir]
EXPL seems/seem some man$_{DAT}$ the-horses$_{NOM}$ be slow
'it seems to some man that the horses are slow.'
(Holmberg & Hróarsdóttir 2003, 998)

b. [C [EXPL T [DAT [*v** NOM . . .]]]]
*Agree

If the dative experiencer is raised to Spec,T, T-NOM agreement is permitted. Apparently, because now only the lower copy of the A-chain (Spec,T, Spec,*v**) intervenes between T and the embedded nominative subject, T-NOM agreement is possible:

(39) a. Mér virðast ~~mér~~ [hestarnir vera seinir]
me$_{DAT}$ seem$_{pl}$ the-horses$_{NOM}$ be slow
'It seems to me that the horses are slow.'
(ibid.)

b. [C [DAT T [$_{v*P}$ ~~DAT~~ [*v** NOM . . .]]]]
Agree
Agree

However, if this experiencer undergoes *wh*-movement, it again blocks long distance agreement:

(40) Hvaða manni veist þú að virðist/*virðast t$_{wh}$ [hestarnir vera seinir]
which man$_{dat}$ know you that seems/seem the-horses be slow
'To which man do you know that the horses seem to be slow.'
(ibid.)

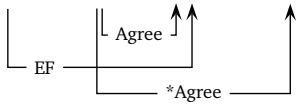
The observation in (40) is unexpected and somewhat paradoxical if *wh*-movement is preceded by A-movement to Spec,T since in that case, (40) is, as far as the relevant part of the structure is concerned, identical to (39): the element in Spec,*v** (i.e. the base position of the experiencer) is merely the lower copy of an A-chain, which should allow agreement.

Chomsky argues against Holmberg & Hróarsdóttir's (2003) suggestion that the experiencer moves directly to Spec,C in (40) (without passing Spec,T) so that its A-chain only consists of the element in base position and thus blocks agreement by intervention. He points out that the presence of the A-chain (Spec,T, Spec,*v**) is established by the familiar A-movement effects (binding, scope, Weak Crossover etc.)

Chomsky's own solution to this problem is again provided by the assumption that both operations in (40), A- and A′-movement, are driven by the phase head C and apply in parallel. As a consequence, there is no relation between Spec,C and Spec,T in the case of (40), but an operator-argument relation between Spec,C and each of the two separate A-chains (Spec,T, Spec,*v**) (formed by the Agree feature

of C, inherited by T) and (Spec,v*) (formed by the Edge feature of C).[26] The latter A-chain suffices to block T-NOM agreement and the former A-chain yields the A-movement effects. This solution is represented in (41):

(41) [DAT C [~~DAT~~ T [$_{v*P}$ ~~DAT~~ [v* NOM . . .]]]]



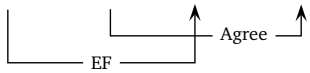
However, there is a problem with this account in view of the surprising fact, observed by Halldór Sigurðsson, that *wh*-movement of the experiencer dative does not block *overt* raising of the embedded subject:

(42) a. Hverjum hefur Ólafur virst vera gáfaður?
who$_{DAT}$ has Olaf$_{NOM}$ seem$_{sg}$ be intelligent
b. Hverjum$_i$ hefur$_j$ [$_{TP}$ Ólafur$_k$ t$_j$ virst t$_i$ [$_{IP}$ t$_k$ vera gáfaður]]]
(Holmberg & Hróarsdóttir 2003, 1004; see also Legate 2002)

So why does the goal of the edge probe not create an intervention effect in this case? The answer given by Holmberg & Hróarsdóttir (2003) that subject raising is an instance of stylistic fronting here cannot be right since the fronted subject agrees with the matrix verb and since there is intervening phonetic material (which usually blocks stylistic fronting). So an alternative explanation compatible with parallel probing should be sought.

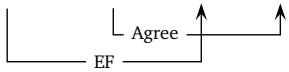
Chomsky's tentative solution (suggested in a personal letter) is that C simultaneously probes with the Edge feature (raising the *wh*-element *hverjum* directly from the base position) and with Agreement (raising the embedded subject *Ólafur* to Spec,T):

(43) [DAT C [NOM T [$_{v*P}$ ~~DAT~~ [v* ~~NOM~~]]]]



According to this proposal, the intervention effect of the single A-chain (Spec,v*), left behind by *wh*-movement, is obviated by parallel application of Edge- and Agree-probing to the experiencer and the embedded subject. Notice that there is nothing wrong with simultaneous probing of two different goals. The same kind of parallel operation applies in the case of *wh*-movement of the object, where the Edge feature of matrix C probes the *wh*-object in the outer Spec of v* and the Agree feature of matrix C (inherited by T) probes the subject in the base position.

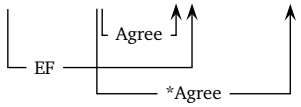
(44) [who C [John T [$_{v*P}$ ~~who~~ [~~John~~ [v* [V ~~who~~]]]]]]



However, Chomsky's proposal raises the question as to why the kind of parallel probing assumed in (44) is not possible in the case of (41). If Multiple Agree is permitted (as Chomsky assumes, following Hiraiwa), matrix C should be able to simultaneously probe the experiencer with the Edge feature and, through T, the experiencer as well as the embedded subject with Agree, as shown in (41), repeated here:

[26] At first sight, this configuration seems to be problematic for the Bijection Principle: an operator cannot bind more than one variable. Note, however, that the Bijection Principle may not be required to account for Weak-Crossover effects (Ruys 2000).

(41) [DAT C [~~DAT~~ T [$_{v^*P}$ ~~DAT~~ [v^* NOM . . .]]]]



If parallel application of operations obviates the intervention effect of the singleton A-chain (DAT) in (43), it should have the same effect in (41).

There are further problems with Chomsky's account. The first one has to do with the PIC. Recall that in the matrix clause of raising examples such as (42) there is the internal phase v^*. Agree-probing into the complement of this phase should therefore violate the PIC. To avoid a violation of PIC, it was proposed in BEA that the Minimal Link Condition be evaluated at the level of the phase and that within a phase the computation operates freely. Since traces are assumed to be irrelevant for this evaluation, the configuration in (42) is in line with the PIC if the phase-level evaluation takes place at the C-phase, or if the complement of v^* is sent to Spell-Out once the head of the next higher phase is merged. It is thus crucial for this proposal that although the head of the next higher phase cannot access the complement of the previous phase, a head located between the two phase heads can. Accordingly, the T-head in (42) is permitted to probe into the complement of the internal phase v^* so that on the basis of the BEA approach the derivation does not violate PIC.

Things are different in the OP approach. Here the Agree feature is only introduced when C is merged so that T cannot probe into the internal complement of the v^*-phase prior to Merge of the next higher phase head. But if Merge of C makes the complement of v^* inaccessible, (42) should violate the PIC. Thus, parallel probing turns out to have consequences different from the assumption that evaluation of "locality" takes place at the next higher phase. Essentially, as noted in section 2.1, the assumption that all operations take place at the phase level, which results in parallel probing, forces us to adopt the first version of the PIC, in which T cannot probe into the complement of v^*, (or at least cannot effect a phonological change there).

Legate (2002) points out several other problems, which essentially stem from the same assumptions. As we have seen, OP assumes that TRANSFER (and thus Spell-Out) applies to the complement of a phase-defining head H at HP, i.e. when H's phase is completed, instead of waiting until merger of the next phase-defining head. This alternative implies that Agreement into the complement of a phase-defining head never affects a phonological change within the complement (Legate 2002). While this alternative will be able to account for long distance agreement in examples such as (39), where the goal of Agree is pronounced in the embedded clause, it is incompatible with Chomsky's analysis of (42).

Things become even more complicated in view of an observation reported in Legate (2002) (originally due to Halldór Sigurðsson) according to which the Icelandic experiencer construction does not permit overt raising of the embedded nominative when the experiencer has undergone topicalization, compare (42) with (45):

(45) *Þer myndi Ólafur þa hafa virst vera gáfaður
them$_{DAT}$ would Olaf$_{NOM}$ then have seemed be intelligent
'To them, Olaf would then have seemed to be intelligent.'
(Legate 2002)

So if the simultaneous application of movement (C-probing) to the experiencer and the embedded nominative is supposed to explain the grammaticality of (42), it should likewise imply that (45) is well-formed.

We have seen that the parallel application of operations permits an explanation of an interesting observation on Long Distance Agreement in Icelandic raising construction. The observation was that *wh*-movement of the matrix dative Experiencer (which is assumed to pass through Spec,T) blocks Long Distance Agreement of matrix T with the embedded object although agreement is permitted when this Experiencer raises to Spec,T and becomes a quirky subject. Again, parallel operation of the Edge probe (from C) and the Agree probe (from T by inheritance) provides the A-movement effect on the one hand (the chain (Spec,T, Spec,*v**)) and on the other hand permits the Goal of the Edge Probe (the singleton A-chain in base position) to act as an intervener for Long Distance Agreement. We have also seen that there still remain some problems that need to be solved.

3.4 Expletives

The treatment of expletives in the minimalist program is not uniform. The standard view of expletives is that they are merged in Spec,*v**, and that they are probed by T and raised to Spec,T in the same way that θ-role bearing subjects are.

In BEA (p. 114), Chomsky proposes that expletives may be *externally* merged into Spec,T, and act as a probe itself, probing T. Suppose we have a φ-defective expletive, like English *there*. Two cases must be distinguished: φ-defective T (raising and ECM infinitives), and φ-complete T (finite clauses):

(46) a. there_i seems [t_i to be a man in the garden]
b. there is a man in the garden

In (46a), the expletive is externally merged with the embedded T. It then probes T, but we know that it is not inactivated, because it cannot remain in Spec,$T_{\varphi\text{-def}}$. It must raise to the specifier position of the higher, finite, T. That is, the unvalued (uninterpretable) feature of T is valued, but the unvalued feature of the expletive (person) is not.

In (46b), the expletive cannot raise further, so we must conclude that its unvalued features are valued. The unvalued features of T, on the other hand, are *not* valued, because we know that T can probe a lower associate and agree with it.

In both cases, the expletive is a defective probe. Why is it then that this defective probe values the unvalued features of defective T, *without* itself being inactivated, but does *not* value the features of non-defective T, while itself being inactivated?

3.5 Parallel probing

The assumption that the probing by C and the (derivative) probing by T takes place in parallel has been made on both empirical and conceptual grounds. Empirically, parallel probing allows us to account for the example in (14), repeated here:

(14) $[\text{of which car}]_i$ was [the driver t_i]$_j$ awarded t_j a prize?

Since extraction out of a subject is prohibited, the *wh*-element in (14) must have been extracted from the base position of the DP *the driver of which car*. Similar considerations apply to the example in (21), repeated here, where probing by C must have taken place from an intermediate position:

(21) of which car_j is [[the driver t_j]$_i$ [t_i likely to [t_i cause a scandal]]]?

Note, however, that such examples, in combination with standard subject islands, merely show that C can attract an element from an intermediate position. In

older models, this observation might have led to the conclusion that C and T must probe in parallel: once T has raised the subject to its specifier, C cannot extract anything from it, so probing by C must take place before or simultaneous with probing by T.

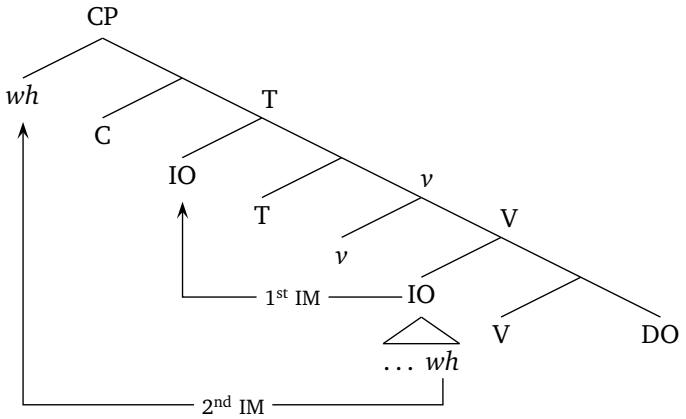
But in the current model, such a conclusion is not inevitable. Chomsky does assume that the Generalised Inactivity Condition makes probing the subject in Spec,T by C impossible, because once the subject has reached Spec,T, all its features are valued. However, probing by C is trigged by EF, which means that C must be able to probe elements that have all their features valued.

Consider *wh*-movement of an object. The *wh*-object has its case feature valued in Spec,V. It is then raised to Spec,*v** and in the next cycle, it is probed and raised by the EF of C. No uninterpretable feature comes into play, and the inevitable conclusion seems to be that once the *wh*-object has moved to Spec,V, all its uninterpretable features (in actual fact just the single CASE feature) are valued, which should render it inactive, and block any further movement.

A similar point can be made for adjuncts, and, in V2-languages such as Dutch or German, for any constintuent that can move to Spec,C. Note, furthermore, that the *wh*-element *of which car* in the subject DP of examples (14) and (21) must have all its features valued inside the DP. Still, C can probe and raise the DP as a whole, or, under certain conditions, just the *wh*-phrase inside it.

A related consequence of the EF-feature approach, one already discussed in section 2.3, is that minimality effects no longer follow. This, together with the copy theory of movement, implies that there is no *a priori* way to exclude the possibility that C probes a *lower* copy of an A-chain. That is, we could in principle assume that the derivation of (14) takes place along the following lines, with C probing *after* T:

(47)

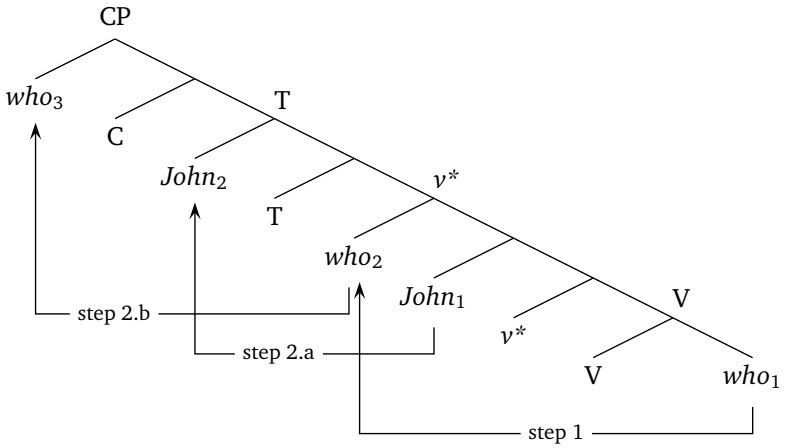


Nothing in the theory of movement assumed in the present framework seems to prohibit this interpretation. Moreover, parallel probing actually raises problems in the *v**-domain. Take again the straightforward *wh*-object question in (19), repeated here:

(19) who did John see?

The derivation of this structure, according to the parallel probing framework, is given in the tree in (20), also repeated:

(20)



The analysis of (47) does not take the merger of the subject into account, however. *John*, the subject of the clause, is initially merged in Spec,v^*, and must then move to Spec,T. But note that in a parallel-probing approach, the merger of *John* must take place simultaneously with the merger of v^* and with the merger of the object *who* in Spec,v^*.

Now, we may simply state that this is not a problem, and claim that "it just works", but we do not believe this issue to be trivial. The most pressing question seems to us to be how we can make sure that when two elements are merged simultaneously into the specifier position of one head, they end up in the correct hierarchical order. Note that the lower specifier of v^* is an A-position, to which the C-I interface assigns the external ϑ-role, while the outer Spec,v^* is an A′-position. That is, we must ensure that the subject is merged in the lowest specifier position, otherwise it ends up in an A′-position.

Furthermore, note that in the C-T domain, all the operations that take place simultaneously, merger of C, inheritance of Agree features, probing by C and T, merger of the subject and some element in Spec,C and TRANSFER, are all in a way related. They are all operations needed to establish the referenial force of the clause. In the v^*-domain, however, the merger of the subject is of a completely different nature than the merger of the object *wh*-element. The former is required in order to satisfy the ϑ-requirements of the predicate, and is therefore an instance of external merge. The latter is a device to ensure that the *wh*-element is available for further movement operations, and is an instance of internal merge.

In short, while it may be possible to argue for simultaneity of operations in the C-T domain on the basis of a certain uniformity of the various operations, such an argument is not easy to maintain for the v^*-domain. The operations are of a different nature and serve completely different purposes. Furthermore, it is not clear how we can obtain the correct order of elements.

The ultimate reason for assuming that probing by C and probing by T proceed in parallel is a conceptual one. Because uninterpretable features must be deleted immediately upon valuation, and because uninterpretable features must remain until TRANSFER (because of the phonological effects they may have), valuation can only take place together with TRANSFER. As a consequence, all operations must be triggered by the phase heads, and valuation of the CASE feature of object and subject can only take place once their respective phase heads have been merged. The inevitable conclusion is that probing by C and probing by T must take place simultaneously.

Note, however, that on this assumption it is not only the probing by C and T that must proceed in parallel. While we may still argue that T is merged with *v*P before

the merger of C, everything that takes place after that must take place in one fell swoop. That is, the merger of C, passing on the Agree features to T, probing by C, probing by T, valuing of the unvalued features on T and on the subject, remerging of the elements being probed, and TRANSFER of the TP.

In an explicitly derivational theory, such a conglomerate of simultaneous operations must be regarded with some suspicion. A derivation in which all operations take place simultaneously is in effect a representational analysis. When conceptual considerations take us in such a direction, it may be worth our while to reexamine the essential points in these considerations.

Ultimately, it is the principle of Full Interpretation and the manner in which the interface to C-I is conceived of that lead to the assumption of parallel probing. First, FI leads to the assumption that features cannot have features, so that once an uninterpretable feature is valued, it cannot be distinguished from an interpretable feature (i.e., it cannot have a feature marking it as an uninterpretable feature). Second, FI also suggests that uninterpretable features must be deleted before the structure arrives at C-I, because the C-I module would not be able to deal with an uninterpretable feature.

Deletion of uninterpretable features is a task of the C-I interface.[27] An interface system is a system that mediates between two modules A and B, passing along information from A to B, or, put in different terms, translating structures from A into structures readable by B.[28] As pointed out by e.g. Jackendoff (1997), an interface generally does not pass on *all* information available in A. Only those pieces of information that are relevant to B are passed on.

That is, the C-I interface does not pass uninterpretable features on to C-I. But does the C-I interface really need to establish which features are uninterpretable on the basis of the syntactic derivation, or could we say that the interface is only sensitive to those features that we commonly call "interpretable"? That is, interpretable features in the syntactic derivation are translated into features that the C-I module operates with, while uninterpretable features are simply ignored.

This does not mean that the C-I interface must be exceptionally 'smart'. C_{HL} is a formal system, and the set of features that it operates with is limited. As a result, the interface can actually be a rather simplistic, 'dumb' system, with a list of all the features it must pass on/translate. Anything not 'on the list' is silently discarded.

This ensures that no feature ever reaches the C-I module that it truly cannot interpret. It also means that we can drop the notions of *uninterpretable feature* and *deletion* of features from our model of C_{HL}. The computational system only needs to know about features and whether they have a value or not. If not, the derivation must ensure that they receive one, the standard mechanism for this being Agree.

With this notion of interface, it is no longer necessary to maintain that probing by C and probing by T, merging of the probed elements and TRANSFER all take place simultaneously. However, to solve the No-Tampering problem discussed in section 3.2, we need to show that T can attract the subject before C has been merged.

We could, of course, return to the older analysis, but we cannot ignore the evidence that suggests the Agree features are inherent to C rather than T. However, when examined more closely, some of this evidence suggests that the situation is somewhat more complex.First, Chomsky says in footnote 21 that "[s]ometimes the φ-features of C are morphologically expressed, as in the famous West Flemish examples" (OP p. 9). Similar effects are observed in other Germanic dialects, e.g.

[27] Note that it is important to see that an interface is not the same as the system it connects to. That is, the C-I interface and the C-I module are two separate entities. Furthermore, deletion cannot take place in syntax proper, because uninterpretable features may be relevant for the SM module, so must survive until the SM interface.

[28] And possibly vice versa, although standard generative theory does not include this option.

Frisian and Bavarian:

(48) a. . . . da-n-k (ik) werk-en *(West-Flemish)*
that-1sg-SubjCl (I) work-1sg
'. . . that I work'
b. Er hot gsagt, dass-st Du a Depp bist *(Bavarian)*
he has said that-2sg you a jerk are.2sg
'he said that you are a jerk'
c. . . . dat-st do jûn kom-st *(Frisian)*
that-2sg you tonight come-2sg
'. . . that you come tonight'

Note, however, that at least in Bavarian, there is no complementiser agreement with sluicing and with comparatives lacking the verb:

(49) a. I bin größer als wie-(*st) du
I am taller than how you
'I am taller than you'
b. Du host des Buch kafft aber i woaß ned wo-(*st)
you have the book bought but I know not where
'you bought the book, but I don't know where'

The same appears to be the case in *Kölsch*, the dialect of Cologne:

(50) a. isch bin jrößer als wie du / als wie-st du bist
I am taller than how you / than how you are
'I am taller than you (are)'
b. du häs dat Buch jekäuf, isch weiß evver nit wo / wo-st du
you have the book bought, I know but not where / where-2sg you
et jekäuf häs
it bought have.2sg
'you bought the book, but I do not know where (you bought it)'
(Amina Hallab, p.c.)

Contrast these facts with the following example from English. Many English varieties use a non-nominative pronoun when there is no overt finite T:

(51) a.I'm taller than him
(51) I'm taller than he/*him is

That is, when no overt T head is present, the case of the pronoun appears to be a default case, not a structural case assigned through Agree.

Taken together, these facts suggest that the Agree features on C are present to establish the Agree relation between T and its subject. Since the earliest minimalist investigations, Agree has been considered the trigger for movement, but in the current framework, we have an alternative. As we have seen, movement into the C-domain is simply triggered by the edge feature of C, without needing an Agree relation as trigger. Suppose that the same is true for movement to Spec,T. That is, T is merged with *v*P, and, by virtue of its edge feature, may attract an element from its domain, which is then merged into Spec,T. In the next step, C is merged, with a set of Agree features. It passes these features on to its complement, where they establish the familiar subject-verb agreement.[29]

If T probes merely by virtue of its edge feature, there is of course no guarantee

[29]For some speculations on the reason why phase heads have a set of Agree features, see Kremers (2007). Note that the assumption that the edge feature of T may trigger internal merge might go some way towards explaining the "residual" EPP effects that Chomsky mentions.

that it will attract the subject. However, in standard cases, the attracted element must be the subject, because it has an unvalued case feature which it can only value through Agree with T. If any other element is attracted, the subject's case feature is not valued, and the derivation will crash.[30]

The approach sketched here obviously raises questions about the exact nature of the Agree relation. It seems that the standard case of Agree must be reinterpreted as a Spec-Head relation.[31] However, the approach has several advantages. All movement is triggered by EF, and only EF. We no longer need to make a distinction between movement triggered by EF and movement triggered by Agree+EF. Note that the A/A′-distinction can still be made: it is still the case that A-positions are licensed by Agree.

There is also no need to assume some unintuitive version of the No-Tampering Condition. We can maintain the assumption that merge is always at the root, and "tucking in" is not required. We account for the facts that led Chomsky to assume that Agree features are inherent to C, not to T, without having to resort to parallel probing, and all the consequences it brings.

References

Barss, A. (2001). Syntactic reconstruction effects, *in* M. Baltin & C. Collins (eds), *The handbook of contemporary syntactic theory*, Malden, Mass./Oxford: Blackwell Publishers, pp. 670–696.

Boeckx, C. (2003). *Islands and Chains*, Amsterdam/Philadelphia: John Benjamins.

Chomsky, N. (1986). *Barriers*, Vol. 13 of *Linguistic Inquiry Monograph*, Cambridge, Mass: The MIT Press.

Chomsky, N. (1995a). Bare phrase structure, *in* G. Webelhuth (ed.), *Government and binding theory and the minimalist program*, Oxford: Blackwell, pp. 383–439.

Chomsky, N. (1995b). *The minimalist program*, Cambridge, Mass: The MIT Press.

Chomsky, N. (2000). Minimalist inquiries: the framework, *in* R. Martin, D. Michaels & J. Uriagereka (eds), *Step by Step*, Cambridge, Mass: The MIT Press, pp. 89–155. MIT Occasional Papers in Linguistics 15.

Chomsky, N. (2001). Derivation by phase, *in* M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, Mass: The MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy, *in* A. Belletti (ed.), *Structures and Beyond*, Vol. 3 of *The cartography of syntactic structures*, Oxford: Oxford University Press.

Chomsky, N. (2006). Approaching UG from below. ms. MIT.

Chomsky, N. (forthcoming). On phases, *in* R. Freidin, C. Otero & M.-L. Zubizaretta (eds), *Foundational Issues in Linguistic Theory*, Cambridge, Mass: The MIT Press.

Fanselow, G. & Çavar, D. (2002). Distributed deletion, *in* A. Alexiadou (ed.), *Theoretical Approaches to Universals*, Amsterdam/Philadelphia: John Benjamins, pp. 65–107.

[30] The Icelandic LD agreement facts discussed in section 3.3 show that in certain circumstances, there are ways to save the derivation. What these circumstances are remains to be clarified.

[31] Cf. Koopman (2006), who argues in favour of Spec-Head, and against Agree.

Fuß, E. (2007). Cyclic spell-out and the domain of post-syntactic operations: evidence from complementizer agreement, *Linguistic Analysis* **33**(3–4): 2–37. (Special issue on Dynamic Interfaces, ed. by K. Grohmann).

Grewendorf, G. (2001). Multiple *Wh*-fronting, *Linguistic Inquiry* **32**(1): 87–122.

Holmberg, A. (1986). *Word Order and Syntactic Features in the Scandinavian Languages and English*, PhD thesis, University of Stockholm.

Holmberg, A. & Hróarsdóttir, T. (2003). Agreement and movement in Icelandic raising constructions, *Lingua* **113**(10): 997–1019.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*, Vol. 28 of *Linguistic Inquiry Monograph*, Cambridge, Mass: The MIT Press.

Koopman, H. (2006). Agreement configurations: in defense of "spec-head", *in* C. Boeckx (ed.), *Agreement Systems*, Amsterdam/Philadelphia: John Benjamins, pp. 159–199.

Kremers, J. (2007). Zu strukturellem Kasus und Phasentheorie. Talk at the Linguistics Study Group, Cologne, 11 July 2007.
URL: *http://joostkremers.fastmail.fm*

Legate, J. A. (2002). Phases in *Beyond Explanatory Adequacy*. ms. MIT.

Nespor, M., Guasti, M. T. & Christophe, A. (1996). Selecting word order: the rythmic activation principle, *in* U. Kleinhenz (ed.), *Interfaces in Phonology*, Vol. 41 of *Studia Grammatica*, Berlin: Akademie Verlag, pp. 1–26.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*, Vol. 28 of *Studies in Generative Grammar*, Dordrecht: Foris.

Nunes, J. (2004). *Linearization of Chains and Sideward Movement*, Cambridge, Mass: The MIT Press.

Nunes, J. & Uriagereka, J. (2000). Cyclicity and extraction domains, *Syntax* **3**(1): 20–43.

Postal, P. (1974). *On Raising*, Cambridge, Mass: The MIT Press.

Richards, M. (2007). On feature inheritance: An argument from the phase impenetrability condition, *Linguistic Inquiry* **38**(3): 563–572.

Richards, N. (2001). *Movement in Language*, Oxford: Oxford University Press.

Rizzi, L. (1997). On the fine structure of the left-periphery, *in* L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer Academic Publishers, pp. 281–337.

Ruys, E. (2000). Weak Crossover as a scope phenomenon, *Linguistic Inquiry* **31**(3): 513–539.

Saito, M. (1994). Improper adjunction, *MIT Working papers in linguistics 24* **24**: 263–293. Formal approaches to Japanese Linguistics 1.

Sigurðsson, H. (1996). Icelandic finite verb agreement, *Working Papers in Scandinavian Syntax* **49**: 1–26.

Stechow, A. v. (2005). LF in einem Phasenmodell: Bemerkungen anhand von Fischers Bindungstheorie. Talk at the GGS Conference Tübingen, May 2005.
URL: *http://www2.sfs.uni-tuebingen.de/~arnim10/Handouts/index.html*

Truckenbrodt, H. (1995). *Phonological phrases: their relation to syntax, focus and prominence*, PhD thesis, Cambridge, Mass: MIT.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases, *Linguistic Inquiry* **30**(2): 219–255.

Truckenbrodt, H. (in press). The syntax-phonology interface, *in* P. de Lacy (ed.), *Cambridge Handbook of Phonology*, Cambridge: Cambridge University Press.