

# Multiple Feature Inheritance and the phase structure of the left periphery

Phil Branigan

April 28, 2016

## 1 *When phase theory and cartographic structures collide*

Two of the most significant developments in syntactic theory over the past two decades have been phase theory (Chomsky 2000) and cartographic syntax (Rizzi 1995). Both have provided significant advances in our understanding of the nature of human language. Both have enabled the identification of important generalizations and developed persuasive models to explain them. But the basic premises of these two approaches have remained inconsistent with each other. In particular, cartographic models rely on the postulation of richer structures than phase theory has managed to accommodate, especially in the clausal "left periphery". The goal of this paper is to effect a partial reconciliation of phase theory and cartographic syntax, by exploiting certain analytic avenues opened up by Chomsky's (2013; 2014) *labeling theory*.

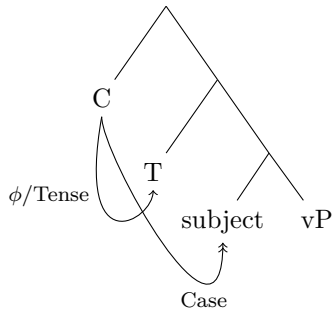
Beyond some specific and important technical incompatibilities which I will examine, the disjunction between phase theory and cartographic approaches is problematic on quite a large scale. Cartographic models have proven to be invaluable tools for investigating how syntactic structures interact with semantic/pragmatic principles which play a large role in how language is used in real life. Phase theory, on the other hand, has been at the heart of recent developments in "biolinguistics", where essential questions about evolution and human cognition are addressed. In a way, the divide between cartography and phase theory implies a gap between the study of the biology of human cognition and the study of language and culture.

The particular technical difficulties posed for phase theory by a richly articulated left periphery can be readily appreciated if we first consider the simpler structures where phase theory offers a successful account of important generalisations. One such set of structures is seen in (1), where different types of complementisers are matched up with different types of verbal inflections and with different Case forms on subjects.

- (1) a. (...that) I shot the sheriff.
- b. ...for me to shoot the sheriff...
- c. ... $\emptyset$  PRO to shoot the sheriff...

Chomsky's (2008) account of this set of dependencies derives both the content of T and the Case form for the subject from the nature of the C phase head. In this analysis, C provides Tense/Finiteness features, and  $\phi$  features, all of which are transferred to underspecified T by the Feature Inheritance operation (FI). C also provides Case features for the subject, which are assigned as a side-effect of the  $\phi$  feature valuation procedure: (2).

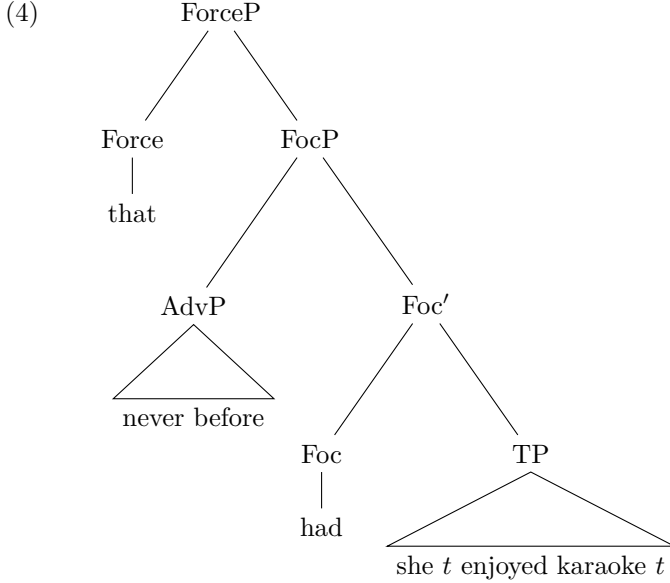
(2) *Feature transfer from the C phase head*



But this account of the C-T-subject dependencies does not readily accommodate structures in which additional phrase structure appears to separate C from T. Rizzi's (1996) cartographic analysis of English negative inversion structures like (3) constitutes an example of this problematic type.

(3) Prudence insisted that never before had she *t* enjoyed karaoke.

Even a streamlined variant of Rizzi's treatment of the left-periphery of the embedded clause in (3) will involve the structure (4).



Such a structure provides a dedicated landing site for the displaced negative focus and the auxiliary verb, and situates them to the right of *that*, as is clearly necessary. But the Force head seems equivalent to phase theoretic C. In phase theoretic terms, C/Force is the most plausible initial repository of unvalued  $\phi$  features, together with whatever interpretable Tense/Finiteness/*etc.* features C must transfer to T. And crucially, in Chomsky's formulation, C transfers all such features to the head of its complement, which is why agreement features end up realised on the T position. But if we start with a structure like (4), then TP is not the complement to Force/C, so it should be impossible for features originating in the phase head to be transferred past FocP to T.

The problem becomes still more challenging as more material is admitted into the clausal zone between C and T. And standard cartographic models actually accomodate quite a large amount of such material. Rizzi's account of the embedded clause in (5) includes two Topic phrases and a Focus phrase between Force and TP, for example.

- (5) *Credo che ieri QUESTO a Gianni avreste dovuto dirgli.*  
 believe-1 that yesterday THIS to Gianni have-2 should say-him  
 'I believe that yesterday you should have said THIS to Gianni.'  
 (Rizzi 2004)

Yet C-T dependencies are manifestly instantiated in these structures, as well. For example, in standard Italian, the selection of subjunctive mood by a matrix verb must evidently be mediated by the complementiser. For simple cases like (6), this is easily accomodated in a phase theoretic analysis; one need only suppose that C is the source of mood features, which are realised on T due to the application of a Feature Inheritance operation.

- (6) *Voglio che tu \*vai/vada.*  
 want.1 that you go/go:SUBJ  
 ‘I want you to go.’ (Manzini 2000)

The transferred features are not subject to morphological realisation on C, but they can still be selected on C by a matrix predicate.<sup>1</sup> Yet Italian subjunctive complements can still contain foci, topics, and adverbs between C and the finite verb, as seen in (7).

- (7) a. *Gianni crede che LA MELA Maria abbia mangiato, no*  
 Gianni believes that the apple Maria has:SUBJ eaten not  
*la pera.*  
 the pear  
 ‘Gianni believes that Maria as eaten THE APPLE, not the pear.’  
 (Giorgi and Pianesi 2004)
- b. *Credo che fortunatamente/evidentemente/probabilmente*  
 think-1 that luckily/evidently/probably  
*l’abbia già fatto.*  
 it has-SUBJ already done  
 ‘I think that luckily/evidently/probably he has already done it.’  
 (Cocchi and Poletto 2002)

But if FI can only transfer features from a phase head to the head of its complement, then C should not be the source for subjunctive features in (7). But then it becomes difficult to explain how the matrix predicate can select for subjunctive features at all in examples like this.

In short, phase theory works best if the left periphery is minimal, with T as the head of the complement of C. But cartographic analyses have shown that structures in which this is not the case are common, and some of the patterns that phase theory must attempt to model are found in these richer structures.

Clearly, something has to give. My intent in this paper is to show that a small change in the phase theory makes it possible to accomodate the richer left peripheral structures, and to increase empirical coverage in the process. The analysis I will defend is based on the following two premises: (8).

- (8) a. A single C may introduce multiple interpretable discourse features into the derivation (Shlonsky 2006), including [TENSE], [WH], [TOPIC], [FOCUS], [SUBORD] and perhaps others.<sup>2</sup>

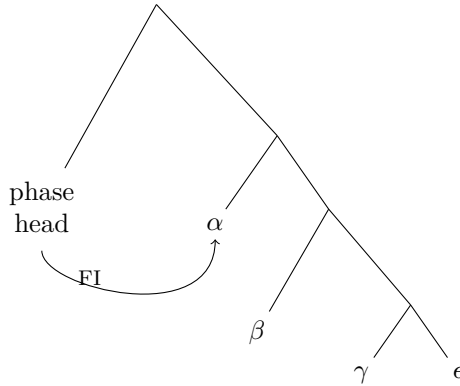
1. I am grateful to Cecilia Poletto for this observation. Shlonsky (2010) also notes that selection of subjunctive mood is a problem for standard phase theoretic notions of locality in selection.

2. Shlonsky (2006) develops a proposal similar in spirit to what I am suggesting, but which is implemented in the opposite direction. While I claim that FI can transfer features downwards to a variety of heads, Shlonsky’s analysis has C raising upwards through a series of “reprojected” positions, and actualising different feature sets in turn.

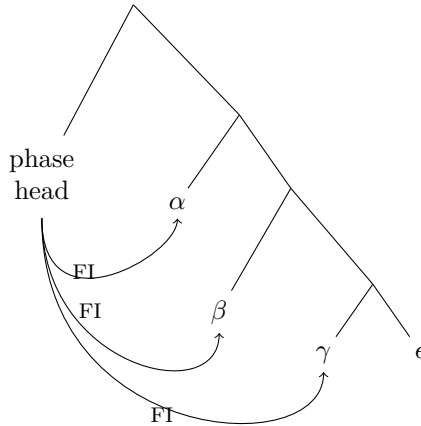
- b. Feature Inheritance can apply multiple times, transferring feature complexes to successively more remote phrasal heads.

The second point (8b) is the crucial innovation. In Chomsky's (2008) account, Feature Inheritance operates only between a phase head and the head of its complement. I will show that this operation can apply across a larger span of structure, as long as it also applies in the shorter spans as well. The canonical context for FI is (9a), where  $\alpha$  is the head of the complement of the phase head, and  $\beta$  and  $\gamma$  are the heads of the more complements to  $\alpha$  and  $\beta$ , respectively; my claim is that FI can occur as illustrated in (9b), as well.

- (9) a. *singular FI*



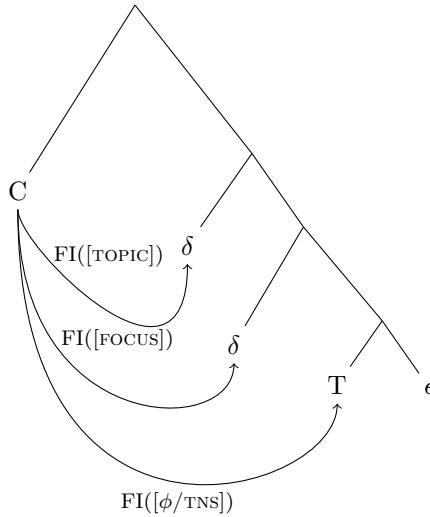
- b. *multiple FI*



The abstract structure in (9b) takes on the more concrete shape in (10) when the clausal left-periphery will eventually include multiple “criterial” specifiers. At the derivational point of FI, T lacks a specifier, because movement of the

subject upwards depends on the presence of those features in T which legitimate one:  $\phi$  features, in the labelling theory. The same will be true of higher left-peripheral heads, which will acquire their own specifiers—topics, foci, *etc.*—in due course. And like T, the other heads which obtain content from C must be underspecified initially. Thus, rather than the derivation merging a cartographic Top or Foc head to build up the left periphery, some number of underspecified heads which can obtain their content from C when FI transfers discourse heads may be merged. I use the symbol  $\delta$  to indicate such feature-neutral heads.<sup>3</sup>

(10)



It may be worth emphasizing at this early point that the absence of any initial categorial distinctions between a  $\delta$  which will end up with a topic specifier and one which merges with a focus does not mean that these distinctions will not be present when the phrase structure is presented to the interfaces. As far as interpretation is concerned, the structures generated when FI transfers discourse features downwards will be indistinguishable from those which cartographic analyses entertain, because the interfaces see only the end result, and not the mechanism which generates it. The end result is still a derivation in which scope-discourse semantics is “syntacticised”, in Rizzi’s (2014) sense.

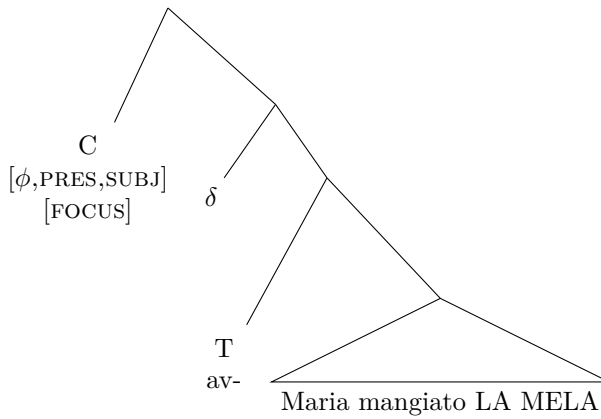
How this proposal resolves the general problem can best be illustrated by examining a specific case. To that end, consider how the Italian subjunctive example (7a) will be characterised in a model where FI applies multiple times to transfer different features of C. The C phase head must be the initial source for all the features which are realised on T in the surface string. Since T in (7) bears

3. This terminological choice is inspired by Lochbuhler and Mathieu’s (2013) analysis of discourse features in Ojibwe conjunct mode structures.

$\phi$  features, Tense and Mood, these all originate in C, and T itself originates as a defective, underspecified (potential) head. But C must also provide features to identify a  $\delta$  as a head with which a contrastive focus may merge, i.e. Foc. So the initial phasal structure of the complement clause in (7a) will be (12).

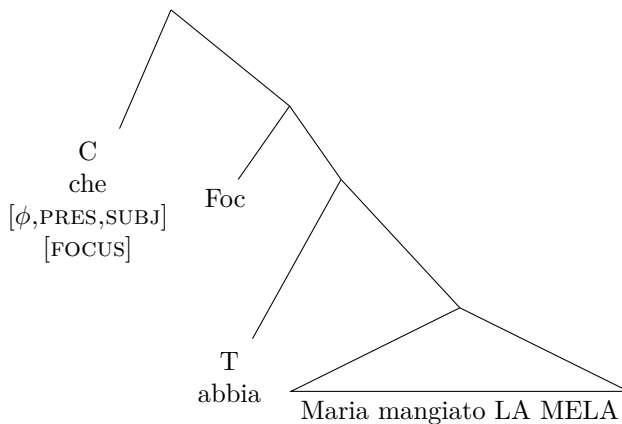
- (7a) *Gianni crede che LA MELA Maria abbia mangiato, no la pera.*  
 Gianni believes that the apple Maria has:SUBJ eaten not the  
 pear

(11)



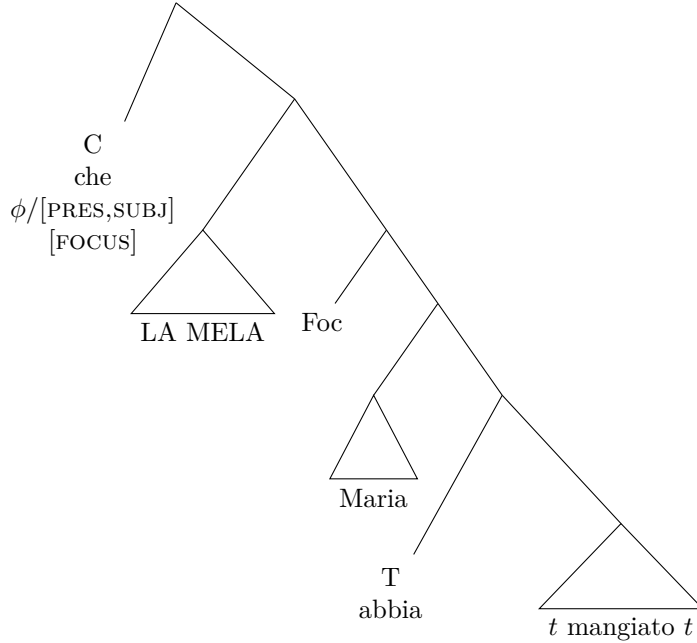
Given this initial structure, multiple applications of FI can transfer the  $[\text{FOCUS}]$  to  $\delta$  and the  $\phi$ /Tense/Mood feature cluster to T, generating (12), where T and  $\delta$  are no longer underspecified.

(12)



And then merge of the subject with T and the focus with Foc derives the complete phase-final structure (13).

(13)



FI provides Foc and T with the features necessary for the internal structure of the phase to develop into a legitimate structure. But C also retains the featural content which will matter for selection at the next phase level.

In some cases, however, C itself can be deleted, if all of its content has been displaced by FI (Chomsky 2014). This occurs in some Italian complement clauses. (Rizzi 1982; Poletto 1995). In subjunctive, future, or conditional complement clauses, C can be deleted optionally, in specific grammatical contexts which mirror those in which embedded verb-second word order is found in German. Poletto (1995) provides the examples in (14).

- (14) a. *Credo (che) abbia già parlato con te.*  
think.1 that have-SUBJ already spoken with you  
‘I think that he has already spoken with you.’  
b. *Credo funzionerebbe meglio, se lo riparassi.*  
think.1 work-COND better if it repaired.2  
‘I think it would work better if you repaired it.’

Given phase theoretic assumptions, the  $\phi$ /Tense/Mood features must originate in the C phase head—which must therefore exist at some point in the derivation within these complement clauses. Once C deletes, however, later selection of mood features by a matrix predicate may take place on the basis of



But such selection can only occur if T is the highest head after C is deleted. Thus, as Poletto (1995) shows, C-deletion cannot take place if TP is preceded in the left periphery by material other than the complementiser. For example, both of the examples in (7) become ungrammatical if the complementiser is deleted: (15).

- This result follows immediately, given how the selectional requirements of the matrix verb must be satisfied in C-deletion contexts. If the sister of deleted C is a left peripheral structure, i.e. a TopP, FocP, ModP, *etc.*, in cartographic terms, then the subjunctive features sought by the matrix predicate will not be available to be selected, since they are present only lower down in the complement clause. In example (15a), then, the selectional requirements of the matrix predicate *credere* are not satisfied by its complement, which is a TopicP phrase, and not a projection of subjunctive T.<sup>4</sup>

While the necessity for multiple Feature Inheritance can be inferred from selectional dilemmata, it can be shown to be necessary on other grounds, independent of these. One argument emerges by examining movement constraints in negative inversion and topicalisation, and how these compare to what is found with *wh*-movement.

4. As Poletto observes, C-deletion is also less acceptable when the subject of the embedded clause precedes the verb. This may follow from the workings of the labelling theory (Chomsky 2013), since the label for the embedded clause in this case will not be T, but rather the features shared by T and the subject. When the subject does not merge with T, as may be the case in (14), T will be the label for the complement of deleted C, and later for the matrix predicate.

phrases to specifier position is activated at the phase level, subject movement and wh-movement must both occur at the same point in the derivation. Supporting evidence comes from the contrasts in (16).

- (16) a. They asked us which visitor we had heard rumours about.  
 b. \*They asked us which visitor rumours about had shocked us.  
 c. ?They asked us which visitor rumours about had been repeated.

As Chomsky notes, extraction of a wh-phrase from an object nominal is permitted, but extraction from within a subject depends on its base position. If the subject originates as a vP specifier, as in (16b), extraction is impossible. Subjects which originate as “objects” pattern like true objects, however, and allow extraction, as in (16c). Chomsky’s point is that all subjects should pattern alike if movement to the specifier position in TP precedes wh-movement, contrary to what the data shows. Therefore wh-movement must occur no later than subject raising, which he interprets as meaning that all movement within a phase is simultaneous.

The same contrast is found with in negative inversion structures: (17).

- (17) a. They insisted that few visitors had they heard worse rumours about.  
 b. \*They insisted that few visitors had worse rumours about shocked them.  
 c. ?They insisted that few visitors had worse rumours about been repeated.

Given Chomsky’s reasoning, movement of the negative focus to its surface position must take place at the same point as movement of the subject from within the verb phrases. In this case, however, both the subject and the focus are situated to the right of the complementiser. One could try to resolve the problem by supposing that Foc—the head to which the auxiliary verb raises—is the phase head in this context, but this gambit fails quickly, since negative inversion does not prevent wh-movement from displacing a wh-phrase to the left of the inverted focus position. The relative clause in (18) provides an example of such wh-movement past a proposed negative focus.<sup>5</sup>

- (18) This was the stage on which few finer performances had they ever seen.

The phase head must therefore be to the left of the inverted focus, or such wh-movement would violate Phase Impenetrability. And in fact, such wh-movement occurs even in the same context where the focus and the subject must be displaced at the same derivational moment.

5. Wh-movement in questions is less compatible with negative inversion. I suppose that the clash in focus structure is a factor in the degraded status of examples like (i).

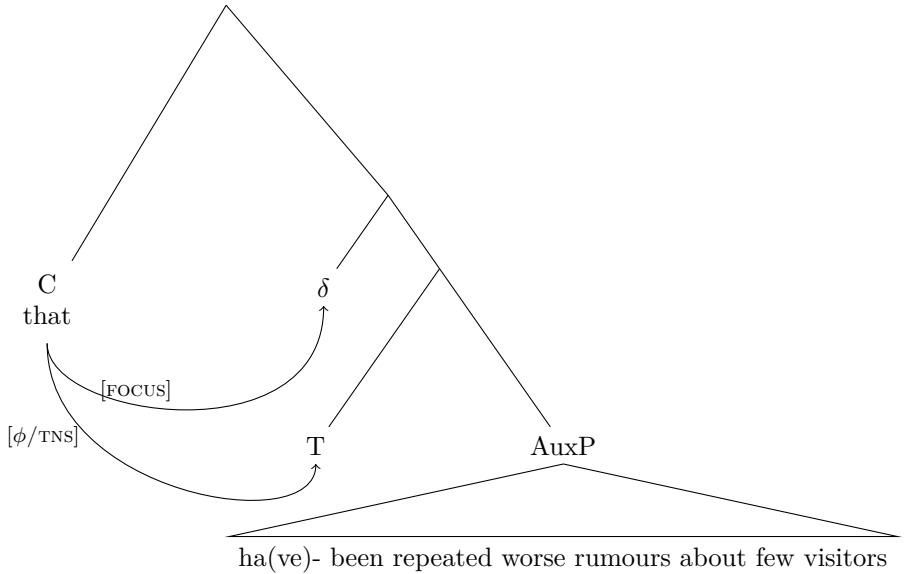
(i)\*I wonder which stage few finer performances had they ever seen on.

- (19) a. ?This was the stage on which few plays had finer performances of ever been presented.  
 b. \*This was the stage on which few plays had finer performances of ever impressed us.

If movement operations take place at the phase level, then the conclusion must be drawn that movement can create specifiers in a sequence of phrasal categories immediately following the phase head.

The contradiction is resolved only if the initial structure of the clausal phase lacks a specifier for both the focus phrase and for TP. For TP, the lack of an initial specifier reflects the defective character of T before FI provides it with content. *Ceteris paribus*, the same should be true of the head with which the negative focus must merge. Then the initial structure for the embedded clause in (18c) when C is merged must be (20).

(20)



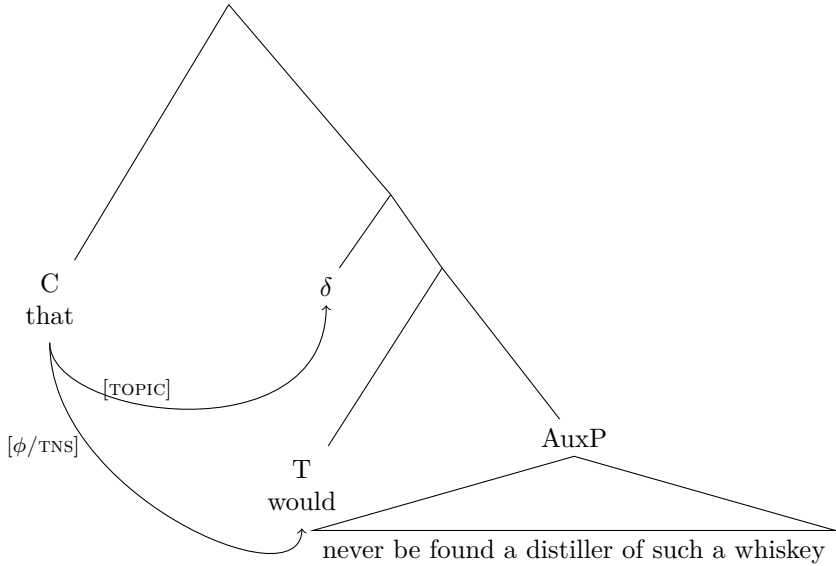
Essentially the same pattern is found in the (21) examples, in which the auxiliary verb is not displaced, but where a topic is.

- (21) a. We feared that such a whiskey, we might never meet a distiller of.  
 b. \*We feared that such a whiskey, a distiller of might never call us.  
 c. ?We feared that such a whiskey, a distiller of might never be found.

Again, successful  $\bar{A}$  extraction of a phrase from within a subject DP depends on the base position of the subject. If the subject originates as the specifier of

vP, extraction is impossible, as in (21b). In contrast, if the subject originates as the complement of the verb, as in (21c), then extraction is allowed. As subject raising obscures this distinction, it must be the case that the topicalisation in (21) takes place no later than subject raising, i.e. at the same time. So fronted topics, like negative foci, are moved only after the C phase head has transferred a discourse feature to the head with which the topic must merge. And the structure which permits the complement clause in (21c) to be derived will then be (22).

(22)



Of course, in embedded clauses like those in (23), where both topics and negative foci are present, the FI operation must be still more active, operating on the topic feature, the focus feature, and the  $\phi$ /Tense features in turn.<sup>6</sup>

- (23) a. ?We feared that such a whiskey, never might a sober distiller of ever be found.  
b. ??We would hardly visit a home in which such a fine whiskey, only rarely would bottles of ever be shared

The use of FI to transfer a discourse feature rather than  $\phi$  features and Tense/Finiteness features accords well with Chomsky's (2013; 2014) approach to *that*-trace effects and subject *wh*-phrases, where subjects must remain in their

6. I take such examples to be essentially grammatical, although clearly the processing load is quite taxing.

“criterial” position—as sister to TP—until the derivation of the clausal phase is complete. For embedded interrogative clauses like the (24), the consequence is that *who* does not raise into the left periphery at all.

(24) David asked who wanted dessert?

In order for such a clause to be interpreted properly as a *wh*-question, some mechanism must permit the *wh*-phrase to serve as a “co-label” with a *wh*-bearing head in this structure. Chomsky proposes that interrogative C can transfer its *wh* feature to T with FI, so that T and its specifier match in *wh* features. This is sufficient to establish [ *who* wanted dessert ] as an interrogative structure for the semantic interface. But one further step is necessary to allow the complement clause to satisfy the selectional requirements of the matrix context. C must be deleted, which is possible in Chomsky’s account after FI has transferred all the formal features originating in C down to T. With no C present in the complement clause, it is the interrogative DP+TP structure which serves as the true complement to the matrix predicate.

Although Chomsky does not discuss it, it is an attractive consequence of his model that the parallelism constraint on Across-the-Board *wh*-movement (Williams 1978) follows immediately. The pattern is seen in (25):

- (25) a. She asked who called Petunia and visited Maurice?
- b. She asked who Petunia called and Maurice visited?
- c. \*She asked who called Petunia and Maurice visited?
- d. \*She asked who Maurice visited and called Petunia?
- e. She asked who Maurice visited and Betty believed had called Petunia?

Here, a *wh*-phrase can be shared if it is the subject of two conjoined clauses, or if it is a non-subject of two clauses (even if it is a subject in an embedded clause), but not if it is a subject of one clause and not of the other. Under Chomsky’s account, *who* in (25a) will remain as a sister to TP, while *who* in (25b,e) will be merged higher, becoming a specifier to CP. Since these are not the same position, (25c,d) are impossible.

The claim put forward here is then essentially the same as Chomsky’s treatment of *wh*-features in subject questions. Like *wh*-features, focus and topic feature originate in the phase head, and like *wh*-features, they are prone to being transferred by the FI operation. Unlike *wh*-features, however, focus and topic features do not signify the clause type (Cheng 1991), so they cannot remain in the phase head at the edge of the clause, but must always be subject to the influence of FI.

Independant confirmation of the idea that FI transfers a discourse feature like Focus or Topic from C downward emerges from a contrast observed by Progovac (1992). She notes that that English *only*-phrases in object position cannot license polarity items, while *only*-phrases which undergo negative inversion do.

- (26) a. Only pennies did we share with a soul.  
 b. \*We shared only pennies with a soul.

But *only*-phrases license a polarity item from the subject position without needing to raise further.<sup>7</sup>

- (27) (We understand that) only Francis gave pennies to a soul.

The acceptable (26a) and (27) examples pattern together if the Focus feature always originates in C, and FI then transfers it to the head of its complement. In negative inversion structures, the complement is a phrase higher than TP, but C can transfer Focus to T if TP is its complement, just as C transfers *wh* to T in subject *wh*-questions (Chomsky 2013). *only*-phrases then license polarity items only as specifiers to heads with a Focus feature.

### 3 *English wh-questions and subject-auxiliary inversion*

One incidental advantage of this way of thinking about the role of FI in rich left peripheral structures is that it provides a simple account of the conditions under which subject-auxiliary inversion takes place in English questions. Consider the paradigm (28).

- (28) a. Matilde asked who wanted pie.  
 b. Who wants pie?  
 c. Matilde asked which dessert we would like.  
 d. Which dessert would you like?

In both of (28a,b), the *wh*-phrase must remain below the left periphery, and C must delete, as discussed above. In (28c), the *wh*-phrase can merge directly as the specifier for CP; the matching *wh*-features on specifier and head establish this structures as interrogative. But something different occurs in (28d), where the auxiliary verb must raise up. Other differences between embedded and root questions are found when left-peripheral topics are added to the clause, as in (29). In a root clause, adverbial topics are acceptable to the left of the *wh*-phrase. In embedded questions, they are impossible altogether.

- (29) a. On hot days, which dessert do you prefer?  
 b. Matilde asked (\*on hot days) which dessert (\*on hot days) you prefer.

In root questions like (29a), the word order indicates that the *wh*-phrase does not actually become a specifier for the CP phase. Instead, it must merge

7. Branigan (2010) concludes from the contrast that the subject actually raises to an  $\bar{A}$ -position in all cases. But if Chomsky's (2014) approach to *that*-trace effects is adopted, then the subject must remain as a specifier to TP throughout the derivation of the clausal phase.

in a lower position, like the focus in a negative inversion structure (Rizzi 1996). This result is readily accommodated in a model in which C can transfer discourse features to a  $\delta$  head. Like the focus and topic features, a *wh* feature can be freely transferred either to T or to  $\delta$ . For (29a), the fronted auxiliary verb occupies the same position ( $\delta$ ) as it does in a negative inversion structure. In fact, since *wh*-phrases and negative foci share the semantic property of being downward entailing quantifiers, the conditions under which T raises to  $\delta$  can be stated as (30).<sup>8</sup>

(30) *Residual verb second rule*

$\delta$  must contain a root when  $\delta$  has a monotone-decreasing specifier.

But unlike negative inversion, *wh*-phrases cannot merge with  $\delta$  in embedded clauses, at least in many (standard) English variants: (31).<sup>9</sup>

- (31) a. \*Matilde asked which dessert would you like.  
 b. \*Which dessert you would like?

This means there are two questions to answer concerning the interaction of C and  $\delta$  in questions: first, why can a *wh*-phrase not merge with C in a root clause, and why must it do so in an embedded clause? The latter can presumably be reduced to selection, given that embedded questions are interpreted as such. Root clauses are not selected, so the possibility of a *wh*-phrase merging lower than the phasal root can be derived. What still requires explanation is the unacceptability of root questions where the a *wh*-phrase merges too high.

One obvious explanation for the unacceptability of (31) would exploit Chomsky's (2014) idea that C can be deleted if FI transfers all of its formal features downwards. If the [*wh*] feature is transferred from C to  $\delta$  or T in a root clause, then the conditions for C-deletion may be satisfied. If now C-deletion is not only possible, but necessary, in English root interrogatives, then there will be no root CP for a *wh*-phrase to merge with. The ungrammaticality of (31b) follows immediately.

It bears noting at this point that the interaction of C-deletion with movement and FI implies a specific order of operations. Deletion must obviously be a late operation, since it must follow FI, and since FI triggers movement. In fact,

8. The rule (i) constitutes a nano-parametric setting, in the sense of Biberauer et al. (2013) and Biberauer and Roberts (2015).

9. Henry (1995) and McCloskey (2006, and others) have documented varieties of English in which subject-auxiliary inversion is possible in embedded questions. For these structures, the text analysis implies that even in embedded clauses, interrogative C may sometimes be deleted, leaving *wh*-enriched  $\delta$  as the head which attracts a *wh*-phrases. For some speakers, the interpretation of such embedded questions is different than what is found with their non-inverted counterparts (Woods 2016); this might be a function of the C-deletion procedure.

Similarly, in Afrikaans, embedded questions can also employ verb-second word order (Biberauer 2012, 2016). Again, this would suggest that C-deletion can apply in a greater range of contexts than is found in many other Germanic languages, given the analysis of verb-second word order developed below.

since the only effect of C-deletion is to diminish the size of the portion of the phase which undergoes Transfer, it seems reasonable to conclude that C-deletion occurs immediately before Transfer takes place. This detail will be significant when verb-second islands are examined below.

(There are further specific technical issues associated with this explanation, which I examine below; they bear on the understanding of how verb-second islands can be understood in this model.)

In Chomsky's account of the *that*-trace effect, deletion of C leads to an automatic reorganisation of the phasal structure of a clause, because the head to which C transfers its features becomes a new, derived phase head. The idea is that FI makes heads in the domain of C *potential* phase heads, but only the highest potential phase head serves as the actual phase head. If C is not deleted, then it serves as the phase head, but if C transfers all of its features to T and then deletes, then T becomes the new phase head. This will be the case in (28b). When C transfers a [WH] feature to  $\delta$  before deleting, as in (28d), the effect should be that  $\delta$  becomes the derived phase head instead, as  $\delta$  is higher than T.

The parallelism constraint on ATB wh-movement applies in root questions, as in embedded ones, as seen in (32).

- (32) a. Who called Petunia and visited Maurice?  
 b. Who did Petunia call and Maurice visit?  
 c. \*Who called Petunia and did Maurice visit?  
 d. \*Who did Maurice visit and called Petunia?  
 e. Who did Maurice visit and Betty believe had called Petunia?

This result is ensured if interrogative root C always deletes, because again the positions occupied by *who* are different for wh-phrases which are local subjects and wh-phrases which are not. The former are specifiers for T; the latter, specifiers for  $\delta$ . And the ungrammaticality of (32c,d) follows since these two specifier positions are structurally distinct.

This interaction of multiple FI with C-deletion offers an account of a number of respects in which wh-subjects and other subjects do not behave alike. One of the simpler cases involves the use of epithets inside wh-phrases. As Pesetsky (1987) observes, such epithets are acceptable only with wh-phrases at the periphery, and not with wh-phrases which remain *in situ*.

- (33) a. Who the hell did they pay for this work?  
 b. Who the hell did they pay for what?  
 c. \*Who did they pay for what the hell?

But subject wh-phrases can contain such epithets, unless they are the secondary wh-phrase in a more complex multiple wh structure.



- (34) a. Who the hell paid for this work?  
 b. Who the hell said that who had done what?  
 c. \*Who said that who the hell had done what?

If one concludes from such data that only wh-phrases outside of TP may contain epithets, then this would constitute an argument that subject wh-phrases do raise to a left-peripheral position. But if the position of the phase head is taken into consideration, another account of the same data is available. In all cases in (33)–(34) except (34a), the derived phase head for the root question is a  $\delta$ , after C deletes. In (34a), the derived phase head is T. And the proper characterisation of where epithets can appear inside a wh-phrase is that they must appear to the left of the derived phase head in the clause in which they appear. In other words, such epithets cannot be a part of the Transfer domain in their clause.

Bošković (2015) offers another argument against the idea that subject wh-phrases occupy the usual subject position, based on McCloskey’s (2000) West Ulster English data in (35).

- (35) a. Who was arrested all in Duke Street?  
 b. \*They were arrested all last night.

In this variant of English, *all* can be stranded inside the verb phrase by movement of a wh-phrase subject, although normal subject movement to Merge with TP does not allow the same quantifier to be stranded. Again, if the explanation is that wh-movement alone can strand the *all* quantifier, then the contrast in (35) might be said to follow, although this raises further questions about how EPP/labelling requirements are satisfied in (35a). But a simpler account of this contrast is actually readily available in the current model. In (35b), the effect of *all*-stranding is that *they* and *all* are interpreted in different position within the same Transfer operation, which affects the full TP. But in (35a), *who* is interpreted at a different derivational stage than *all*. If the rule is that *all* and its associated DP must not be pronounced in the same Transfer operation, then the contrast in (35) obtains.

More positive support for the idea of C-deletion in root interrogatives comes from considering the scopal asymmetries with different types of wh-questions noted by May (1977). The pattern can be seen in (36).

- (36) a. What dessert does everybody prefer?  
 b. Where will everyone sit?  
 c. Who likes every type of pie?

In (36a,b), the questions are ambiguous. Each has a pair-list reading, in which the universal quantifier (*everybody*, *everyone*) scopes over the wh-phrase, as well as a single question reading, with the wh-phrase scoping higher. But in (36c), the pair-list reading is absent, and the wh-phrase must have wider scope than *every*.

The literature on this problem is quite extensive, but any solution must recognize that subjecthood makes a difference. May’s original solution was that the *Scope Principle* allows two quantificational expressions to permute if they c-command each other. For (36a), the LF structure (37a) ensured this result, with the IP-adjoined quantifier c-commanding the wh-phrase. But in (36c), the ECP prevented movement of the quantifier to a position between the wh-phrase and its trace, so no scopal permutation was allowed, and the only reading available obtains of QR adjoins the quantifier to the verb phrase, where it is c-commanded by the wh-phrase.

- (37) [CP what dessert [IP everybody [IP *t* prefer *t* ]]]

May’s precise account is obviously difficult to maintain in current models, including this one. But similar results can be derived in the model proposed here, given the phase structural difference between questions and statements. Consider the structures which are generated under the current approach.

Let us start with the assumption that QR of objects adjoins them to the edge of a vP phase (Legate 2003). For a sentence like ambiguous (38a), this ensures the structure (38b) will be available at the clausal phase level.

- (38) a. Everybody likes some dessert.  
 b. [ C [ everybody T [vP some dessert [vP <everybody> likes (···)]]]]

When the complement of C undergoes Transfer to the C-I interface, this structure will be interpreted with wide scope for *everybody*, if nothing else occurs. But since there are two copies of *everybody* available, *some dessert* can take wide scope if *everybody* “reconstructs” to its base position (Hornstein n.d.; Bruening 2001). In other words, if the content of the upper copy is deleted in the Transfer operation, then the subject quantifier is interpreted in its lower position and will take lower scope.

For wh-phrases, scopal interactions with other quantifiers are based on the position of the highest copy within the clause proper. The wh-phrase merged with  $\delta$  (or C) provides an interrogative label for the clause but does not appear to affect how relative scope is determined.<sup>10</sup> With that in mind, we may compare the structures of (36a) and (36c) after C-deletion occurs, and where the derived phase head is the highest head to which C has transferred features: (39). In (39a), the wh-phrase will have escaped the vP phase *via* its edge.

- (39) a. [  $\emptyset$  [ what dessert  $\delta$  [ everybody T [vP <what dessert> [vP <everbody> prefer (···)]]]]]  
 b. [  $\emptyset$  [ who T [vP every type of pie [vP <who> likes (···)]]]]

10. The (scopal) interpretation of the wh-phrase is essentially that of an indefinite quantifier (Karttunen 1969).

In (39a), the derived phase head is  $\delta$ , which means that everything to the right of  $\delta$  belongs to the portion of this clause which undergoes Transfer to the C-I and S-M interfaces at the end of this phase. In particular, the C-I interface will be presented with the structure (40).

(40) [ everybody T [<sub>vP</sub> <what dessert> [<sub>vP</sub> <everybody> prefer (···)]]]

Just as with the ambiguous declarative sentence above, this structure accommodates both wide and narrow scope for *everybody*, since the subject quantifier can be interpreted in its surface position, or it can be “reconstructed” below the wh-phrase by deleting the content of the top copy.

But now consider what gets sent to the C-I interface in the (39b) case. Since T is the derived phase head, it will only be the vP which undergoes Transfer: (41).

(41) [<sub>vP</sub> every type of pie [<sub>vP</sub> [ <who> likes (···)]]]

But the existential quantificational content of *who* is not part of what is provided to the C-I interface because the phase includes a higher copy of the same item, even though the higher copy is not part of the Transfer domain. And the higher copy cannot itself be deleted in this case because it is not part of the portion of the phase which undergoes Transfer. So the quantifier adjoined to vP cannot scope over the wh-phrase, and the pair-list reading is impossible.<sup>11</sup>

May’s Scope Principle allows quantifiers to permute when they are close enough to each other, where close enough is identified as mutual c-command. The alternative approach suggested here also derives ambiguity from two quantifiers being close to each other, but closeness is a matter of belonging to the same Transfer domain, a concept which is already built into phase theory on other grounds.

More generally, it seems that the claim that C-deletion is an obligatory part of the derivation of English root questions provides a more principled way to understand a number of recalcitrant facts about such questions. But C-deletion can be an option only if FI transfers features downwards, and in any structure

11. The explanation appears to extend as well to more complex structures in which an universal quantifier combines with multiple wh structures, as in (i).

(i) a. Which book did everyone write a review of for what publisher?

b. Who wrote a review of every book for what publisher?

The (ia) example is ambiguous (Pesetsky 2000). Under the pair-list reading, this requests an answer which specifies for every person which book they wrote a review of for what publisher. Under the other reading, it asks for pairs of people and publishers, such that each person wrote a review of every book for the corresponding publisher. But (i-a) lacks a pair-list interpretation. As in the simpler examples, this can be derived from the phasal structure, if wh-tuples in multiple wh questions derive their scope relative to other quantifiers from the topmost position of the (upper) wh-phrase within TP. In (i-a) the subject quantifier *everyone* may be interpreted above the vP-adjoined *which book*, or it may be reconstructed to its base position inside vP. But in (i-b), it is vP alone which undergoes Transfer to the C-I interface, and so the higher copy of *who* cannot be eliminated from the scope calculation.

where C and T are not contiguous, this can only mean that FI transfers features more than once.

#### 4 Phase head transfer and islands in verb-second structures

In Chomsky’s original formulation, phase theory is designed to provide an account for far more than just the dependencies between C and T. In fact, the principal role of phase theory is to explain a range of island constraints which regulate syntactic phrasal movement. But if C is responsible for the features of both left peripheral heads and T, then the effects of this relationship should also be seen in this other type of phase theoretic effect. This turns out to be the case. But in order to see how the interaction of C,  $\delta$ , and T plays out in island effects, it is necessary first to examine how multiple FI plays out in the analysis of word order in Germanic verb-second languages.<sup>12</sup>

##### 4.1 Verb-second word order

Let us consider the German case, first, where the data is more transparent in several respects than in Dutch, Norwegian, *etc.* There are four general contexts to consider: root clauses with the subject in initial position, root clauses with a proposed non-subject (topic, focus, wh-phrase), embedded non-v2 clauses, and embedded v2 clauses. German examples of the first three appear in (42).

- (42) a. *Er hat dich gestern nicht angerufen, weil er dich nicht*  
he has you yesterday not called because he you not  
*stören wollte.*  
disturb would  
‘He didn’t call you yesterday because he didn’t want to disturb you.’
- b. *Im Winter wohne ich in Vancouver.*  
in-the winter live I in Vancouver  
‘In the winter I live in Vancouver.’
- c. *Wo wohnt ihr Sohn?*  
Where lives your son  
‘Where does your son live?’

A typical cartographic analysis of such structures is based on the idea that preposed topics, *etc.* are ‘specifiers’ for associated functional categories, i.e. Top, Foc, *etc.* The (42b) and (42c) examples then should differ only in the interpretive character of the specifier-head pairings to the left of the subject—(42b) is a Topic phrase; (42c), a wh-question. The status of the verb second order in subject-initial sentences like (42a) has been more controversial. Given

12. For brevity’s sake, I put to the side the thorny and contentious questions raised by the “symmetric” verb second languages, Icelandic, Faroese, and Yiddish.

the analysis of the English analogues already presented, however, a simple unification of these word order cases is readily attainable. It has already been argued that C provides some range of discourse features which can be transferred to a sequence of  $\delta$  heads in the left periphery, and these heads in turn serve as the targets for internal merge of topics, foci, wh-phrases, adverbs, *etc.* For the (42b,c) examples, the implication is that they should have the structures on the left in (43), and they should be transformed into the structures on the right.

- (43) a. [ C [  $\delta$  [ T [ ich im Winter in Vancouver wohne ] ] ] ]  
 $\rightarrow$  [  $\emptyset$  [ im Winter Top [ ich T [ *t t* in Vancouver wohne ] ] ] ]
- b. [ C [  $\delta$  [ T [ ihr Sohn wo wohnt ] ] ] ]  
 $\rightarrow$  [  $\emptyset$  [ wo Q [ ihr Sohn T [ *t t* wohnt ] ] ] ]

Three details in how these structures are presented in (43) are important to keep in mind. First, German is head-final when the verb-second word order is not involved (Koster 1975; Den Besten 1983), so the initial position for the finite verb is shown at the right edge of the phase.<sup>13</sup> Second, the initial complementiser is shown as deleted on the right side of the arrow in (43). This is a matter of parametric choice, and simply extends the analysis of English “residual” verb second presented already. In other words, German—and the other regular, assymetric verb-second languages—must delete a root complementiser, presumably for semantic reasons. Third, the position of the verb is not yet correct in the derived structures in (43).

For a subject-initial clause, there are actually several sub-issues to consider in the analysis of German word order. One is the criterial status of German subjects. In English, subjects must raise to T, and this ensures the surface word order, and it also ensures that English will be subject to the *that*-trace effect (Chomsky 2014). In German, the status of the subject is less clear, but it appears that subjects can sometimes be found lower in their clause (Haider 1993), and the *that*-trace effect is arguably not a feature of German grammar (Haider 1983). The implication is that T need not always have a specifier in German, unlike in English.

One consequence of the greater positional freedom of German subjects is that “criterial freezing” of subject wh-phrases is not to be expected. In other words, German wh-subjects may well appear in a higher position than English wh-subjects, so that the equivalent embedded questions in (44) may differ in their structures, as indicated.

- (44) a. I wonder [  $\emptyset$  [ who T read this book ] ]
- b. *Ich möchte wissen, warum er nicht geschrieben hat.*  
 I would.like know why he not written has  
 ‘I would like to know why he has not written.’

13. For the purposes of this paper, I can remain happily agnostic on the question whether the best analysis of German verb-finality involves roll-up derivations.

On the other hand, subject-initial word order in German root clauses is found with weak subject pronouns and with expletive subjects, as Travis (1984) notes. Since weak pronouns and expletives make very poor topics or foci, at least these “subjects” must be analysed as specifiers in TP. And since there is no pre-subject criterial phrase in (45), the structure should be no richer than C+TP.

- (45) a. *Es hat das Gras gefressen.*  
           it has the grass eaten  
           ‘It has eaten the grass.’ (Schwarz and Tomaselli 1990)
- b. *Er werd gestern getanzt.*  
           it was yesterday danced  
           ‘There was dancing yesterday.’
- c. *Es steht ein Mann vor der Tür.*  
           it stands a man at the door  
           ‘There stands a man at the door.’ (Haider 1990)

In this case, of course, C provides no extra discourse feature, and FI will transfer only the  $\phi$ /Tense features to T. The derivation for (45a) will then be (46), *modulo* the verb movement.

- (46) [ C [ T [ es das Gras gefressen hat ] ] ]  
       → [  $\emptyset$  [ es T [ t das Gras gefressen hat ] ] ]

The verb-second “problem” is then how to characterise the conditions under which the finite verb moves from final position in a clause if the position to which it moves is sometimes  $\delta$  and sometimes T. T is evidently not a category which needs a verb attached to it, since the verb remains distant from it in normal embedded clauses. So something other than simply categorial identity is involved in determining when the verb moves.

The multiple FI model takes us a long ways towards an answer immediately.  $\delta$  and T share the property of being targets for FI from the C phase head. This in itself is not sufficient to distinguish T which requires an adjoined verb from T which does not, since C will always provide  $\phi$ /Tense features to T. But C-deletion makes a difference which can be exploited in characterising the verb-second context. What  $\delta$  and T have in common in verb-second structures is not just that FI enriches them, but also that they become clausal phase heads in the process, when C deletes. Since C must delete in a root clause, the phase head of a root clause will always be the next head to its right.<sup>14</sup> The proper description of the verb second requirement is then (47).

14. This approach makes a prediction about the availability of inverse scope in German verb-second clauses. If T becomes the derived phase head, the subject will appear outside the Transfer domain for its clause, and no reconstruction of the subject should be possible.

(47) *Verb-second rule*

A derived clausal phase head must contain a root.

To a large degree, the other assymmetric verb-second languages behave like German, at least in root clauses. This implies that the two grammatical factors which give rise to German word order should be active in Dutch, Danish, Frisian, *etc.* In other words, root C is always deleted, and the verb-second rule (47) is active.

In all the languages of this family, verb-second word order is also found in embedded clauses, although the contexts in which this occurs are often restricted, and the relevant contexts are nuanced, variable, and contested. The analyses developed here will consider only a subset of those that the literature has documented for embedded verb-second, although I believe this subset to be generally representative. Again, we start with the German case.

Den Besten (1983) showed that German permits verb-second word-order in "subjunctive" embedded contexts, as in (48).

- (48) a. *Anna glaubt, Hans-Peter habe das Buch gelesen.*  
Anna believes Hans-Peter have the book read  
'Anna believes Hans-Peter has read the book.'
- b. *Anna glaubt, das Buch habe Hans-Peter gelesen.*  
Anna believes the book have Hans-Peter read
- c. \**Anna glaubt, dass das Buch Hans-Peter gelesen hat.*  
Anna believes that the book Hans-Peter read have

Since subjunctive mood must be a feature transferred from C to T, these embedded clauses must be generated as full CP phases. But in the surface string, the complementiser is always omitted, and it seems safe to conclude that C-deletion is involved, as in root clauses. It can hardly be an accident that C-deletion in German mirrors the Italian complementiser deletion construction discussed already (Poletto 1995). In both cases, it is limited to subjunctive complements serving as the complement to a specific class of predicates (Bentzen

Inverse scope readings should therefore not be available for subject-initial verb-second clauses. This is generally true, as seen in (i-a), but since German is a "rigid scope" language, the same pattern is found even in non-verb-second clauses, as in (i-b).

- (i) a. *Mindestens ein Student hat jeden Roman gelesen.*  
at.least one student has every novel read  
'At least one student read every novel. ( $\forall > \exists$ )' (Krifka 1998)
- b. *...weil mindestens ein Student jeden Roman gelesen hat*  
because at.least one student every novel read  
'...because at least one students has read every novel ( $\forall > \exists$ )'

So the prediction made by the multiple Feature Inheritance model is accurate, it seems, but it may be so for independant reasons.

et al. 2007; Heycock 2006, and many others). It seems that subjunctive complementisers must be less richly endowed with formal features than indicative complementisers are, and so FI is able to empty them out more readily.

As C-deletion occurs only after FI has transferred features from C to the next head down, the next head down becomes the new phase head in (48) when C deletes. What is more, C-deletion is evidently necessary whenever C itself contains other discourse features in an embedded context; witness the unacceptability of (48c). The implication would appear to be that C can only contain a limited set of features in German and languages of its type: C must accomodate  $\phi$ /Tense/Mood features, and it *may* contain one other discourse feature as well. But this property of the language must already be recognised in some form, because Germanic languages must be distinguished from Romance languages, where the opportunities for a complex left periphery are much richer. As German can accomodate only a single discourse feature in C, it follows that when FI transfers that feature downwards, C will automatically delete, because it retains nothing which can be interpreted at the C-I interface.

German is rather atypical in how it structures embedded verb-second clauses. In the other Germanic languages, embedded verb-second clauses may or must be introduced by an overt complementiser which precedes the initial component, as in (49).

- (49) a. *Jan zei dat hij kende dat boek niet.* (Dutch)  
 John said that he know that book not  
 ‘John said that he didn’t know that book.’ (Zwart 1997)
- b. *Jan zei dat Marie kuste hij niet.*  
 John said that Mary kissed he not  
 ‘John said the he didn’t kiss Mary.’
- c. *Eva säger att hon ser aldrig på TV.* (Swedish)  
 Eva says that she watches never at TV  
 ‘Eva says that she never watches TV.’ (Holmberg 2010)
- d. *Jar tror att i det fallet har du rätt.*  
 I think that in that respect har you right  
 ‘I think that you are right in that respect.’
- e. *Der waard ferteld, (dat) it skip wie juster fergien* (Frisian)  
 there was told that the ship was yesterday wrecked  
 (De Haan 2001)

Given the premise that the the finite verb raises to the phase head position in verb-second clauses, embedded verb-second in all of the (49) examples must involve a complementiser located to the left of the phase head (T or Top). As T or Top only become phase heads by virtue of C-deletion, it follows that



these data must reflect CP-recursion structures, in which the lower C is deleted and the upper C remains intact. The Dutch example (49a) will then have the structure in (50) for the full complement clause.

- (50) [ dat [  $\emptyset$  [  $\phi$ ,  $\phi$  > hij kende-T dat boek niet  $t$  ] ] ]

The upper, recursive C *dat* in (50) must evidently have properties distinct from those of the usual declarative complementiser. For one thing, it does not supply Case or  $\phi$  features into TP, since that is the responsibility of the lower, deleted C phase head. As such, the higher C cannot express complementiser agreement features, unlike its regular counterpart (Meer 1991, cited by Zwart 1997). For another, it must match [ $\pm$ FINITE] features with the following clause through selection, rather than transferring any features downwards. This is possible in languages of this type because the verb which bears such features appears in  $\delta$ , the head of the complement to the upper C.

#### 4.2 Embedded verb-second and island effects

C-deletion gives rise to structures in which the phase boundary is moved rightwards. One would therefore expect that wh-extraction of the subject into a higher clause should be possible in German, just as in English, and this expectation is justified: (51).

- (51) a. *Wer sagte sie, werde ihm seine Arbeit hier bezahlen?*  
 who said she will him.DAT his work here pay  
 ‘Who did she say will pay him for his work here?’  
 b. *Was sagte sie, werde Hans-Peter lesen?*  
 what said she will Hans-Peter read  
 ‘What did she say Hans-Peter would read?’

Since non-subjects can be the initial constituent in embedded verb-second clauses, with  $\delta$  inheriting phase head status, non-subjects can also be extracted into a higher clause, as in (51b).

In such cases, of course, the wh-phrase will serve as the sister/specifier of TopP before raising into the higher clause. Since Top is the derived head of the clausal phase, only its TP complement is transferred to the interfaces, and *was* remains available for internal merge operations during the next phase.

(As Henry (1995) shows, Belfast English patterns with German in allowing non-subjects to undergo wh-movement from a verb-second complement. In this language, it is never the subject of the embedded clause which is extracted in this way (Pesetsky and Torrego 2000). (Focal stress on the embedded T makes *do*-support possible in this context, but this is irrelevant for the matter at hand.)

(52) \*Who did John say [(*\*did*) go to school]?

Since English—even in Belfast—does not require that a T phase head contain a verb, *do*-support would not be expected in this case anyways.)

On the other hand, C-deletion actually removes an escape hatch which enables successive-cyclic wh-movement of non-subjects. This change in the structure make it possible to derive the well-documented verb-second island effect, as in (53).<sup>15</sup>

- (53) a. \**Womit glaubte sie, das Kind hatte dieses Brot gegessen?*  
 what-with thought she the child had this bread eaten  
 ‘What did she think the child had eaten this bread with?’  
 (Schwartz and Vikner 1996)
- b. \**Was glaubst du, gestern hat Fritz repariert?*  
 what believe you yesterday has Fritz fixed  
 ‘What do you believe Fritz fixed yesterday?’ (Müller 1995)

Consider the structure of the complement clause in (53a) prior to C-deletion, where *dass* is the phase head: (54).

- (54) [ dass [ das Kind T womit dieses Brot gegessen hatte ]

Movement of *womit* to the edge of this structure will leave it accessible for successive cyclic movement at the next phase, so an account of the island effect requires that wh-movement not take place in (54) *if C will be deleted*. Nothing bars wh-movement *per se* in this structure (for some speakers), as long as it does not coincide with C-deletion. We can derive the v2 island effect, therefore, by limiting C-deletion to contexts in which no wh-movement occurs. In other words, structures like (55) must block C-deletion.

- (55) [ womit dass [ das Kind T <womit> dieses Brot gegessen hatte ]

But notice that the effect of wh-movement in (55) is that *dass* is no longer positioned at the edge of its phase. In fact, *dass* is no longer the unambiguous head of the phasal constituent, since there are two equally prominent phasal sisters immediately below the root. If C-deletion is an operation which targets the

15. Hans-Martin Gärtner has reminded me that the problem of V2 islands is a real one only if extraction from a complement clause is the correct analysis of long-distance wh-movement in German. Reis (1995) has argued that most such cases are actually structures in which a simpler wh-question is broken up by a verb-first parenthetical phrase, so that they merely mimic successive cyclic wh-movement structures superficially. As Wiltschko (1998) observes, however, Reis actually accepts that some long wh-movement does occur in German, so the problem remains, even if there is some question about the status of individual cases.

head of the root constituent, then Minimal Search considerations are sufficient to ensure that it cannot take place in (55).

Verb-second island effects for embedded clauses with initial topics will follow on the same grounds. The structure of the complement clause in (53b) prior to topicalisation, wh-movement, and C-deletion will be (56).

(56) [ dass [  $\delta$  [ Fritz T gestern was repariert hat ]]]

FI transfers the requisite Topic feature from C to  $\delta$ , making  $\delta$  a potential phase head in (56). But in order for  $\delta$  to become the phase head, the phase must end with C in a position to delete, which means C must be at the edge. If wh-movement occurs, C cannot delete, thereby violating Full Interpretation at the C-I interface. But if wh-movement does not occur, then *was* remains within the complement of Top; *was* therefore undergoes Transfer within the phase, and it becomes inaccessible for operations at the next phase level.

Turning to the island effects in the non-German verb-second languages, we find essentially the same verb-second island pattern. The presence of the extra complementiser does not alter the internal structure of the first clausal phase. As such, embedded V2 in most of these languages produces island effects which block extraction of any of the non-initial components of the clause.

- (57) a. \* *Welke film zei je dat Jan had al gezien?* (Dutch)  
 which film said you that Jan had already seen  
 ‘Which film did you say that Jan had already seen?’ (Zwart 1997)
- b. \* *Hvilken film sagde hun at Peter havde allerede set?* (Danish)  
 which film said she that Peter had already seen
- c. \* *Wa hie Pyt my in boadskip stjoerd, dat koe moarn komme?* (Frisian)  
 who had Pyt me a message sent that could tomorrow not  
 come  
 ‘Who had Pyt sent me a message that could not come tomorrow?’

No new explanation is needed here. The structure of the first clausal phase constructed within the complement clause in the Dutch (57a) will be (58).

(58) [ dat [  $\emptyset$  [ $_{<\phi,\phi>}$  Jan T al welke film gezien had ]]]

If the wh-phrase undergoes movement to the edge of this phase, then C cannot delete, and T will never become the phase head. This structure will not be an island, but it will also not exhibit verb-second word order. If the wh-phrase does not move within this phase, then C may delete, and T will then attract the verb, but then *welke film* will be affected when the Transfer operation processes the complement of T. By the time the V2 clause merges with a higher head, C or the verb root, the island status of the first phase is already determined.

Although the presence of the extra complementiser does not alter the extractability of material from low in a verb-second clause, it does affect movement of the initial constituent. The initial constituent in German can undergo  $\bar{A}$ -movement into a higher clause. The initial constituent in languages which add a complementiser to embedded verb-second clauses cannot, which ensures the apparent extended *that*-trace effect observed by Vikner (1991).

- (59) \**Hvilket æble siger de sagkundige at smager ikke bedst?* (Danish)  
 which apple say the experts that tastes not best  
 ‘Which apple do the experts say doesn’t taste good.’

Vikner attributes this particular island effect to the *that*-trace filter. But Chomsky’s account of the *that*-trace effect in English does not apply to the verb-second situation. Consider the Danish example in ([v2that.danish]), where the full structure of the first clausal phase in the complement clause will be (60).

- (60) [ at [ $\phi$ ,  $\phi$ ] hvilket æble T ikke smager bedst ]]

T will become the phase head when C deletes, so only the complement of T is subject to Transfer. And while the subject *hvilket æble* cannot be displaced within this phase, it nevertheless appears at the eventual phase edge, so it remains accessible at a later point in the derivation. As Chomsky shows, the role of the subject in labelling the sentence is complete once the phase itself is finished, so later operations can displace the subject freely.

Fortunately, an explanation of this extended *that*-trace effect presents itself at the next phase level. After C-deletion and merge of the lower clausal phase with an impoverished *at* complementiser, the structure will be (61), with *at* now the new phase head.

- (61) [ at [  $\emptyset$  [ $\phi$ ,  $\phi$ ] hvilket æble smager-T ... ]]]

In order for long extraction to occur, the *wh*-phrase must reach the edge of this new phase so that it remains accessible to later operations.

While embedded V2 clauses in German allow  $\bar{A}$ -movement from the edge into a higher clause, such movement is constrained by the target position in the next phase (Staudacher 1990; Haider 1989; Müller and Sternefeld 1993). The pattern is seen in (62).

- (62) a. *Wen meinst du, hat sie getroffen?* (German)  
 whom think you has she met  
 ‘Who do you think she has met?’ (Müller 2011)  
 b. \**Ich weiß nicht wen (dass) du meinst, hat sie getroffen.*  
 I know not whom that you think has she met

‘I don’t know who you think that she has met.’

- c. \* *Wen glaubt er dass du meinst, hat sie getroffen?*  
 whom believes he that you think has she met  
 ‘Who does he believe that you think she has met?’

Wh-movement from initial position in the complement clause is acceptable if the new merge site is also at the edge of a verb-second clause, but not if the new merge site is at the edge of an embedded non-V2 clause, whether or not the wh-phrase remains there or moves on.

This constraint is not specific to German, either. It operates in Belfast English, too, which indicates that it must reflect the workings of some fairly deep grammatical machinery. (Poverty-of-the-stimulus considerations would enforce the same conclusion.)

The labelling theory offers an explanation for why this pattern obtains. Consider the structure of the complement clause in (62a) just before C-deletion occurs at the conclusion of the phase: (63).

- (63) a. [  $\emptyset$  [ $\langle_{wh,wh}\rangle$  wen hat- $\delta$  [ $\langle_{wh,wh}\rangle$  sie T *t* getroffen *t* ] ] ]  
 b. [  $\emptyset$  [ $\langle_{wh,wh}\rangle$  wen meinst- $\delta$  [ $\langle_{wh,wh}\rangle$ ] du *t* [  $\emptyset$  [ $\langle_{Topic,Topic}\rangle$  *t* hat- $\delta$  ... ] ] ] ]  
 c. [ wen (dass) [ $\langle_{wh,wh}\rangle$  du meinst [  $\emptyset$  [ $\langle_{wh,wh}\rangle$  *t* hat- $\delta$  ... ] ] ] ] ]

Movement of the wh-phrase internal to each clausal phase generates a structure in which  $\delta$  matches features with its wh-phrase sister. (The [WH] feature on  $\delta$  originates in C, which deletes afterwards.) Before FI takes place,  $\delta$  lacks sufficient content to serve as a label, but after FI takes place, it may still only serve as a label in tandem with the wh-phrase.

But features like [TOPIC], [FOCUS], and [WH] have semantic import, even if the label itself need not, and so the presence of these features on  $\delta$  is potentially problematic at the C-I interface. If a  $\delta$  head has an appropriate specifier, we can assume the C-I interface will be able to interpret the structure. But if the specifier moves on, the Top feature will be stranded. The ungrammaticality of (62b) and (62c) then represents the uninterpretability of a [WH] feature on the embedded  $\delta$  when there is no accompanying specifier.

The question then becomes how to account for the acceptability of (62a), in which the specifier for a Top head in the complement clause has also been displaced. In this case, we must consider the structure of the result of the successive cyclic movement (64).

- (64) [ C [ $\langle_{wh,wh}\rangle$  wen  $\delta$  [ $\langle_{\phi,\phi}\rangle$  du meinst [  $\emptyset$  [ $\langle_{wh,wh}\rangle$  *t*  $\delta$  ... ] ] ] ] ]

The pertinent difference in this case is that there are two heads which bear the Focus feature. Just as the “specifier” of the lower clause is “deleted” under identity with the higher instance of *wen*, we may conclude that the problematic

[WH] feature in the lower  $\delta$  is deleted under identity with the features of the upper  $\delta$ . No such option is available in the ungrammatical (62b) and (62c).

If this is correct, then the unacceptability of sentences in which a phrase is extracted from the "first" position in an embedded verb-second complement is one more piece of evidence which supports the approach taken here. These structures fail because they provide an uninterpretable  $\delta$  category to the C-I interface, and the problematic  $\delta$ s obtain their fatal flaw by Feature Inheritance from a C phase head. And as the same C must also provide  $\phi$ /Tense/Mood features to a following T, FI must apply more than once within a single phase.

## 5 Conclusions

To summarise, a disparate range of considerations support the following set of specific conclusions.

- (65) a. Under the right circumstances, FI transfers features from C to some sequence of  $\delta$  heads, then T.
- b. When C deletes, the derived phase head can be T or  $\delta$ , depending on the context.
- c. Phase theory *complements* cartographic theory, rather than conflicting with it.

As always, numerous issues remain open. I inherit from cartographic approaches a number of questions concerning the raising of possible left-peripheral structures. Such questions include: how many discourse functional heads are there, how are they hierarchically organized, and how do they relate to the semantic/pragmatic interfaces? In my terms, some of these questions translate to questions about the initial featural content of the C phase head, and how FI distributes content to heads in the domain of C. It is actually not immediately evident that the nature of these general questions is altered much in translation.

Other questions raised by the approach promoted here concern the role of multiple Feature Inheritance in other syntactic domains, such as vP, DP, and any other phasal categories. At first glance, the left periphery of the vP phase seems conservative, but DP is less so. One might therefore expect to find that multiple Feature Inheritance should allow D to shower features down upon a sequence of underspecified  $\delta$ -like heads. This might then provide a generative source for the rich hierarchical structures posited in the nominal domain by Harley and Ritter (1999) and many others. But this would require work to go beyond wild speculation.

## References

- Bentzen, Kristine, Thorbjörg Hróarsdóttir, Gunnar Hrafn Hrafnbjargarson, and Anna-Lena Wiklund. 2007. "Embedded V2 in Scandinavian: Empirical observations." Unpublished handout from NORMS workshop on Verb Movement, Reykjavik, January 26, 2007, January.

- Biberauer, Theresa. 2012. "One Peculiarity Leads to Another: On Embedded WH-V2 in Afrikaans." Paper presented at the 2012 LAGB conference, University of Salford. September.
- . 2016. "Afrikaans V2 and its implications for our understanding of V2 parameterisation." Rethinking Verb Second workshop, University of Cambridge, March 22.
- Biberauer, Theresa, Anders Holmberg, Ian Roberts, and Michelle Sheehan. 2013. "Complexity in Comparative Syntax: The View from Modern Parametric Theory." *Manuscript*.
- Biberauer, Theresa, and Ian Roberts. 2015. "Rethinking Formal Hierarchies: A Proposed Unification." *Cambridge Occasional Papers in Linguistics* 7:1–31.
- Bošković, Željko. 2015. "On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling." *The Linguistic Review*.
- Branigan, Phil. 2010. *Provocative Syntax*. Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.
- Bruening, Benjamin. 2001. "QR obeys Superiority: Frozen scope and ACD." *Linguistic inquiry* 32 (2): 233–273.
- Cheng, Lisa. 1991. "On the Typology of Wh-Questions," MIT.
- Chomsky, Noam. 2000. "Minimalist Inquiries: the Framework." In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, edited by R. Martin, D. Micheals, and J. Uriagareka. MIT Press.
- . 2008. "On Phases." In *Foundational Issues in Linguistic Theory Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta. Cambridge, MA.: MIT Press.
- . 2013. "Problems of projection." *Lingua* 130 (June): 33–49.
- . 2014. "Problems of Projection: Extensions." In *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Amsterdam and Philadelphia: Benjamins.
- Cocchi, Gloria, and Cecilia Poletto. 2002. "Complementizer Deletion in Florentine: The Interaction between Merge and Move." In *Romance Languages and Linguistic Theory 2000: Selected Papers from "Going Romance 2000, Utrecht, 30 November 2 December"*. John Benjamins Publishing.
- De Haan, Germen J. 2001. "More is Going on Upstairs than Downstairs: Embedded Root Phenomena in West Frisian." *Journal of Comparative Germanic Linguistics* 4:3–38.
- Den Besten, Hans. 1983. "On the interaction of root transformations and lexical deletive rules." In *On the Formal Syntax of the Westgermania*, edited by Werner Abraham, 47–131. John Benjamins.

- Giorgi, Alessandra, and Fabio Pianesi. 2004. "Complementiser Deletion in Italian." In *The Structure of CP and IP*, vol. 2. Oxford Studies in Comparative Syntax.
- Haider, Hubert. 1983. "Connectedness effects in German." *Groninger Arbeiten zur germanistischen Linguistik* 23 (2).
- . 1989. *Parameter der deutschen Syntax*. Tübingen: Narr.
- . 1990. "Null subjects and expletives in Romance and Germanic languages." In *Issues in Germanic Syntax*, 49–66. De Gruyter.
- . 1993. *Deutsche syntax, generativ: Vorstudien zur Theorie einer projektiven Grammatik*. Gunter Narr Verlag.
- Harley, Heidi, and Elisabeth Ritter. 1999. "Meaning in morphology: Motivating a feature-geometric analysis of person and number." 00086 Unpublished ms.
- Henry, Alison. 1995. *Belfast English and standard English*. Oxford University Press.
- Heycock, Caroline. 2006. "Embedded root phenomena." In *The Blackwell Companion to Syntax*, edited by Martin Everaert and Henk van Riemsdijk, 2:174–209. Blackwell Publishing.
- Holmberg, Anders. 2010. "Verb second." In *Syntax: an International Handbook of Contemporary Syntactic Research*, 2nd edition. Berlin: De Gruyter.
- Hornstein, Norbert. n.d. "Minimalism and quantifier raising." In *Working minimalism*. Cambridge, Mass.: MIT Press.
- Karttunen, Lauri. 1969. "Discourse referents." In *Proceedings of the 1969 conference on Computational linguistics*. Association for Computational Linguistics.
- Koster, Jan. 1975. "Dutch as an SOV Language." 00000, *Linguistic Analysis* 1:111–136.
- Krifka, Manfred. 1998. "Scope inversion under the rise-fall contour in German." *Linguistic Inquiry* 29 (1): 75–112.
- Legate, Julie. 2003. "Some Interface Properties of the Phase." *Linguistic Inquiry* 34:506–515.
- Lochbihler, Bethany, and Eric Mathieu. 2013. "Clause typing and feature inheritance of discourse features." Ms., *University of Edinburgh and University of Ottawa*.
- Manzini, M. Rita. 2000. "Sentential Complementation: The Subjunctive." In *Lexical Specification and Insertion*, edited by Peter Coopmans, Martin Everaert, and Jane Grimshaw. John Benjamins.
- May, Robert. 1977. "The Grammar of Quantification," MIT.



- McCloskey, James. 2000. "Quantifier Float and Wh-Movement in an Irish English." *Linguistic Inquiry* 31:57–84.
- . 2006. "Questions and questioning in a local English." *Crosslinguistic research in syntax and semantics: Negation, tense, and clausal architecture*: 87–126.
- Meer, G. van der. 1991. "The "Conjugation" of Subclause Introducers: Frisian *-st*." *NOWELE* 17:63–84.
- Müller, Gereon. 1995. *A-Bar Syntax: a Study in Movement Types*. Berlin: Mouton de Gruyter.
- . 2011. *Constraints on displacement: A phase-based approach*. Vol. 7. John Benjamins Publishing Company.
- Müller, Gereon, and Wolfgang Sternefeld. 1993. "Improper Movement and Unambiguous Binding." *Linguistic Inquiry* 24:461–507.
- Pesetsky, David. 1987. "Wh-in-situ: Movement and Unselective Binding." In *The Representation of (in)definiteness*, edited by Eric Reuland and Alice ter Meulen, 98–129. Cambridge, Mass.: MIT Press.
- . 2000. *Phrasal Movement and Its Kin*. MIT Press.
- Pesetsky, David, and Esther Torrego. 2000. "T-to-C movement: causes and consequences." In *Ken Hale, A Life in Language*, edited by Micheal Kenstowicz. Cambridge, Mass.: MIT Press.
- Poletto, Cecilia. 1995. "Complementizer Deletion and Verb Movement in Italian." *University of Venice Working Papers in Linguistics* 5 (2): 49–79.
- Progovac, Ljiljana. 1992. "Nonnegative Polarity Licensing Must Involve Comp." *Linguistic Inquiry* 23:341–347.
- Reis, Marga. 1995. "Extractions from Verb-Second Clauses in German?" In *On Extraction and Extraposition in German*. John Benjamins.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- . 1995. "The Fine Structure of the Left Periphery." Ms. University of Genève.
- . 1996. "Residual Verb Second and the *Wh-Criterion*." In *Parameters and Functional Heads: Essays in Comparative Syntax*, edited by Adriana Belletti and Luigi Rizzi, 63–90. Oxford University Press.
- . 2004. "Locality and the Left Periphery." In *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, edited by Adriana Belletti, 223–251. Oxford University Press.
- . 2014. "Syntactic Cartography and the Syntacticisation of Scope-Discourse Semantics." In *Mind, Values, and Metaphysics*, edited by Anne Reboul, 517–533. Cham: Springer International Publishing. ISBN: 978-3-319-05145-1 978-3-319-05146-8.

- Schwartz, Bonnie D., and Sten Vikner. 1996. "The Verb Always Leaves IP in V2 Clauses." In *Parameters and Functional Heads: Essays in Comparative Syntax*, edited by Adriana Belletti and Luigi Rizzi, 11–62. Oxford University Press.
- Schwarz, Bonnie D., and Alessandra Tomaselli. 1990. "Some implications from an analysis of German word order." *Issues in Germanic Syntax* 44:251–274.
- Shlonsky, Ur. 2006. "Extended projection and CP cartography." *Nouveaux cahiers de linguistique française* 27:87–93.
- . 2010. "The Cartographic Enterprise in Syntax." *Language and Linguistics Compass* 4 (6): 417–429.
- Staudacher, Peter. 1990. "Long Movement from Verb-Second-Complements in German." In *Scrambling and Barriers*, edited by Gunther Grewendorf and Wolfgang Sternefeld, 319–339. Amsterdam: Benjamins.
- Travis, Lisa. 1984. "Parameters and Effects of Word Order Variation," MIT.
- Vikner, Sten. 1991. "Verb Movement and the Licensing of NP-Positions in the Germanic Languages." Doctoral dissertation, Université de Genève.
- Williams, Edwin. 1978. "Across-the-board extraction." *Linguistic Inquiry* 9:31–42.
- Wiltschko, Martina. 1998. "Review of: "On extraction and extraposition in German"." *The Journal of Comparative Germanic Linguistics* 2 (3).
- Woods, Rebecca. 2016. "A Different Perspective on Embedded V2: Unifying Embedded Root Phenomena." Rethinking Verb Second workshop, University of Cambridge, March 23.
- Zwart, C. Jan-Wouter. 1997. *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.