

## The Spec-head vs head-Spec asymmetry: Toward a theory of post-syntactic linearization and an account of the embedded-topicalization paradox\*

---

Mark de Vos

Department of Linguistics, Rhodes University, Grahamstown, South Africa  
Email: m.devos@ru.ac.za

### Abstract

This paper is a preliminary exploration of how English specifiers and heads are linearized at the PF interface. In particular, it argues that linearization may reflect underlying grammatical relationships of agreement and selection. To evaluate this proposal, it uses the theoretical lens of Relational Theory and empirically investigates topicalization constructions in English which have somewhat paradoxical properties with respect to the perceived relative positioning of the head and specifier of TopicP. These paradoxes can be explained by the linearization proposal in this paper. However, before accepting the proposed solution it is necessary to ensure that the proposal can be consistently applied to other Spec-head relationships in English; the paper therefore includes a discussion of double-object constructions as well as the relative Spec-head orders of CP, TP,  $\nu$ P, VP and PP. The paper ends with a discussion of how the proposal relates to movement within the Minimalist Programme more generally.

**Keywords:** Linearization, linearization paradoxes, LCA, topicalization, focus, movement, Relativized Minimality, left periphery

### 1. Constraints on linearization

Within the Principles and Parameters framework, the Head Parameter allows, in principle, for either heads or specifiers to occur on the left or right. However, there appears to be an asymmetry between heads and specifiers: while heads can indeed occur on the left or right, it seems that “no clear case of a generally Specifier-final language has been discovered” (Roberts 1997:26). Kayne (1994) also notes that even within single languages, head-complement orders are much more stable and harmonic than head-specifier (head-Spec) orders.

One response to this was the Antisymmetry framework, mapping asymmetric c-command to linear order in a one-to-one way (Kayne 1994:6). According to Kayne (1994), the specifier of a head will always precede the head in linear order. This is because a specifier asymmetrically

---

\* I would like to thank the anonymous reviewers for their insightful and thorough comments. An early version of this paper, entitled “Heads and specifiers: On Spec-head word orders in English”, was published in working papers format in *Groninger Arbeiten zur germanistischen Linguistik* (De Vos 2012:53-78) and has since been reworked substantially.

c-commands its head and consequently, assuming the Linear Correspondence Axiom (LCA), there is a one-to-one mapping from asymmetric c-command to linear precedence. The Antisymmetry framework thus responds to the asymmetry between complements and specifiers by making both equally rigid in terms of word order: heads precede complements; specifiers precede heads. The fact that specifier-head (Spec-head) orders are less stable than head-complement orders is not directly encoded in the LCA. However, this does not reflect the reality where specifiers appear to be much more variable in their ordering than complements. Yet, as Kayne (1994:36) points out, alternative mappings are a priori possible. Notwithstanding the success of the Antisymmetric program, in what follows I will explore one such alternative which encodes the asymmetry directly into a grammatical system. It does so by the following suggestion, to be explored in more detail in the rest of the paper: that syntactic relationships are directionally asymmetric and that the directionality of a syntactic relation is reflected in word order. Thus, complements are always selected and, observationally, tend to follow the selecting head in a head-initial language. In contrast, specifiers tend to be in AGREE relationships with the corresponding head: just as either a head or the XP in the specifier may carry an uninterpretable feature, so the directionality of AGREE varies<sup>1</sup>.

There are two valid, non-mutually-exclusive ways of approaching a theoretical proposal such as this. The first stresses the continuity between the proposal and existing theory, demonstrating how the proposal draws in existing concepts and extends them. However, if the continuity or theoretical development is denied then there remains the second approach from a more mathematical modelling tradition, namely to explore the properties of the system resulting from its axioms, e.g. that instead of linearization being a function of asymmetric c-command, I will explore the consequences of taking linearization to be a function of the asymmetric relations instantiated by AGREE and MERGE. To the extent that the properties of the resulting system look “interesting” or are congruent with data or shed light on an existing problem, the system may be explored further by future research. In some ways, the Minimalist Programme (Chomsky 1995b) is an exercise of this nature. I shall adopt the second approach, exploring some of the “interesting” properties of a certain type of grammatical system.

### 1.1. The structure of this paper

In exploring the Spec-head asymmetry, this paper will also end up being a preliminary exploration of linearization strategies. Linearization is the process whereby a two-dimensional “tree” is converted into a one-dimensional linear string of words at the PF interface. Informally speaking, if one takes a tree and “squashes” it down, the terminals in the tree will ultimately end up being arranged as a string of terminals. In order that such “squashing” be done in a well-defined manner, a suitable linearization algorithm is needed to map the tree to linear order. There are probably many possible mappings, but the null hypothesis is that dependencies are mapped directly to linear ordering in a one-to-one and meaning-preserving manner.

Section 2 outlines the theoretical basis of the proposal. Having pointed out that a priori there may be more than one linearization strategy, I would like to focus on the possibility that linearization may reflect underlying grammatical relationships of agreement and selection. With the theoretical foundation in place, I use topicalization constructions in English to evaluate the proposal in section 3. Topicalization constructions have somewhat paradoxical

---

<sup>1</sup> An anonymous reviewer points out that there are some instances where the occupant of a specifier may not necessarily have any agreement relation with the head: escape-hatch movement and EPP may be examples of this. The framework explored in this paper predicts that for non-agreement relations, there may be more flexibility for linearization options; I will return to these in section 7.

properties with respect to the perceived relative positioning of the head and specifier of TopicP, elaborated in section 4. These paradoxes are explained by the linearization proposal in section 5. However, before accepting the proposed solution, it is necessary to ensure that the proposal can be consistently applied to other Spec-head relationships in English; section 6 develops the theoretical proposal with respect to other Spec-head relationships in English and demonstrates that they are a good fit with the proposal. The paper ends with a discussion of how the proposal relates to the theory of movement more broadly.

I acknowledge that this paper is programmatic in nature and is largely focused on English whereas the proposal has universal import. However, for the sake of space and the need to publicize the proposal to the scientific community, thereby opening it to wider critique, it is necessary to narrow the focus of this particular paper. Subsequent research by myself and others can validate the framework and broaden the scope to additional languages and dialects<sup>2</sup>.

## 2. Theoretical outline of linearization by functional dependency

This section outlines the theoretical underpinnings of the proposed theory of linearization. Central to the proposal is that (i) syntactic relationships can be expressed as mathematical relations (section 2.1), and (ii) these relations are the input to a function – the Relational Precedence Axiom – which maps relations to linear precedence (section 2.2) in accordance with a locality principle.

### 2.1. Directionality of syntactic relationships

It is a trivial observation that dependency relations exist in syntax. I take it as unproblematic that syntax includes reference to various dependencies in general since these are implicitly assumed in Chomsky (1995a) and Lasnik et al. (2005), amongst others. These may be expressed by c-command, selection, agreement, etc. For instance, structural relations of selection, instantiated by MERGE, yield a phrase structure marker with the formal representation  $\{p, \{p, q\}\}$ , graphically realized as a “tree” (Chomsky 1995a):  $p$  and  $q$  are merged and  $p$  projects as the label of the resultant constituent. In this case, the notation is not trivial insofar as, mathematically speaking,  $\{p, \{p, q\}\}$  instantiates a partially ordered set which is also a functional dependency (Zwart 2011, Fortuny 2008, Langendoen 2003, Kracht 2003, Uriagereka 1999, Halmos 1960)<sup>3</sup>. A functional dependency is simply a deterministic, one-to-one relationship, formally defined. Thus, for two linguistic entities,  $p$  and  $q$ , each value for  $p$  is uniquely associated with a value for  $q$ . Given this explicit status of functional dependencies in existing theory, I will use the formal construct of ‘functional dependency’ as a meta-language for expressing linguistic relationships<sup>4</sup>.

<sup>2</sup> The theoretical framework for a theory of representation can be found in De Vos (2008). An early implementation addressing EPP (De Vos 2009b) can be found online at <http://ling.auf.net/lingbuzz/000971>. The proposal has also been extended to languages other than English: Afrikaans adpositions (De Vos 2009a), object shift phenomena in Germanic languages (De Vos 2014a), V-T phenomena in English and French (De Vos 2013), and to V2 in Germanic (De Vos 2013).

<sup>3</sup> A partial ordering does not necessarily entail a linear ordering, but is rather a directional dependency within the set. In contrast, a linear order is a total order (Kayne 1994:4). Therefore, the function of the LCA is to map hierarchical orders to total orders.

<sup>4</sup> An alternative view is to take it as axiomatic that a linguistic dependency can be expressed as a functional dependency. Adopting this view, a linguistic system instantiating functional dependencies is just one of many possible systems and the object of its study is to identify its properties. So readers may either take functional dependencies as axiomatic or as intrinsic to the system (e.g. as entailed by Bare Phrase Structure (Chomsky 1995a)), although it makes no difference to the results of this paper.

For instance, if a verb subcategorizes for a DP object, then the verb's SUBCAT/selection feature uniquely determines a DP; every time the SUBCAT feature is present in a verb's lexical feature matrix, it will determine the nature of the DP object. Thus, there is a functional dependency between the SUBCAT/selection feature and the DP it selects. The directionality of these relationships can be tested by individually examining each constituent in the dependency: generally, in each case the nature of the dependent constituent can be determined by examining the features of the determining/selecting/controlling constituent. But the inverse is not true. For instance, examining the features of P,  $V^{\text{Transitive}}$  and  $v$ , it is possible to infer from their SUBCAT features alone that they will all select a DP. However, if one were to examine the features of a DP, it would not be possible to infer by which category (i.e. either  $V^{\text{Transitive}}$ ,  $v$  or P) they will be selected. Using this argumentation, it is possible to say that a selector will always functionally determine a dependent/selectee: selection (and therefore MERGE) instantiates functional dependencies. For ease of reference, I will follow the notation used in Relational Theory and indicate a functional dependency using an arrow, in this case,  $V \rightarrow \text{DP}$ .

Functional dependencies can also be argued to be instantiated by agreement. Consider, for example, subject-verb agreement in French (1).

- (1)
- a. *Je mange-Ø*  
I.1sg eat-1sg  
'I eat'
  - b. *Tu mange-s*  
you.2sg eat-2sg  
'You eat'
  - c. *Elle mange*  
she.3sg.fem eat-3sg  
'She eats'
  - d. *Nous mange-ons*  
we.1pl eat-1pl  
'We eat'
  - e. *Vous mange-z*  
you.2pl eat-2pl  
'You eat'
  - f. *Elles mange-nt*  
they.fem.3pl eat-3sg  
'They eat'

In each of these examples in the paradigm, the interpretable  $\phi$  features on the noun provide a "value" that is ultimately expressed on the verb through agreement. The dependency between them is implicit in the terminology: the noun is the controller of agreement on the verb. In the case of SpecTP, examining the  $i\phi$  features on DP allows an inference about the final form of morphological agreement spelled out on T (e.g. 3sg.masc, etc.). If one were to look at the set of interpretable features of a noun, say *elle* in (1c), in isolation, then one could predict that the morphology on the verb would be 3sg. However, if one were to look only at the uninterpretable  $\phi$  features of the verb prior to agreement, then it would not be possible to predict the nature of the agreement that is ultimately expressed on the verb post-agreement. Thus there is a functional dependency between the subject and the verb in  $T^0$  such that  $\text{DP}[i\phi] \rightarrow \text{T}[u\phi]$ . Thus,

AGREE instantiates a functional dependency where for any interpretable/uninterpretable feature pair under AGREE,  $iF \rightarrow uF$ .

- (2) a. Selection (and therefore MERGE) instantiates functional dependencies such that the selector  $\rightarrow$  the selectee;
- b. AGREE instantiates functional dependencies such that  $iF \rightarrow uF$ , i.e. controller  $\rightarrow$  target.

Having pointed out that selection and AGREE instantiate functional dependencies, I would like to propose that they be utilized for the purposes of linearization.

## 2.2 Linearization and PF legibility conditions

Given the strong Minimalist Hypothesis (Chomsky 2000), narrow syntax is determined by its interfaces and the need to provide representations that are well-formed at that particular interface<sup>5</sup>. Assuming that linearization of hierarchical structures occurs at PF, the first constraint imposed by PF is that it is necessary that syntactic representations be linearizable (but not necessarily linearized). For Kayne (1994), this means that syntax creates representations that already have a total order. For the current proposal, I would like to explore the possibility that syntactic representations have a partial order. To the extent that the framework is successful, it constitutes evidence for the strong Minimalist Hypothesis. Since functional dependencies are encoded in syntactic representations in terms of MERGE and AGREE, I will adopt the null hypothesis that linearization is a one-to-one mapping between normalized functional dependency pairs and linear order<sup>6</sup>.

- (3) Relational Precedence Axiom (RPA):
  - a. Let  $p$  and  $q$  be two syntactic objects (e.g. a head and its complement or a head and its specifier, etc.);
  - b. For any syntactic relation (indicated by  $\rightarrow$ ), if  $p \rightarrow q$  then  $p$  precedes  $q$  in linear order.

The RPA (3) ensures that all complements follow their selecting heads. However, specifiers can either precede or follow their heads in principled and predictable ways, depending on their feature composition and agreement.

### 2.2.1 Schematic representations

As a derivation proceeds, MERGE, MOVE and AGREE create representations which encode functional dependencies. When a phase-head is merged, these representations are transferred to the interfaces, which is to say that the sets of functional dependencies are transferred to the interfaces. For an interpretable/uninterpretable feature pair, the category containing the interpretable feature will precede the category containing the uninterpretable feature (all things

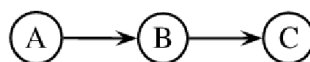
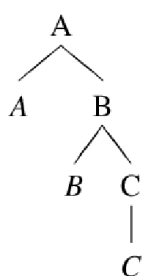
<sup>5</sup> See De Vos (2008) for a discussion of how functional dependencies are a response to the LF interface.

<sup>6</sup> It has been suggested by an anonymous reviewer that the LCA “allows one to get rid of special linearization rules like those proposed by the author”. This perception may lead to a misconstrual of the proposal. The LCA allows a one-to-one mapping between asymmetric c-command and linear order but it does not eliminate the need for linearization per se. Indeed, the LCA itself is such a linearization principle: “d(A) is a linear ordering of T” (Kayne 1994:6). Seen in this way, this theoretical characteristic is shared by the current proposal and the LCA: they both propose a single, universal, one-to-one mapping between syntactic structures – whether seen as asymmetric c-command (LCA) or as the hierarchy denoted by partially ordered sets underlying phrase structure (the current proposal).

being equal, it will be immediately adjacent in linear order unless another dependency intervenes)<sup>7</sup>.

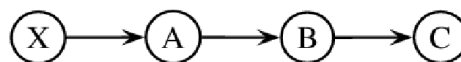
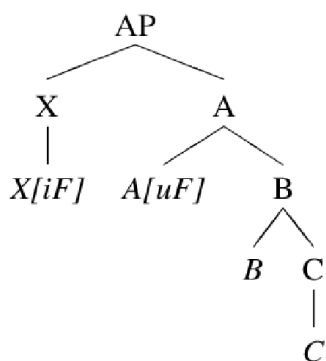
To illustrate, consider the structure in (4) where A selects B and B selects C. For the sake of clarity, the phrase structure diagram is also represented graphically as a graph below. Since selection instantiates a functional dependency,  $A \rightarrow B$  and  $B \rightarrow C$  which, by the RPA (3), yields the linear order  $A > B > C$ .

(4)



Now consider the more complex example in (5), where XP is located in the specifier of AP. Let us assume that X has interpretable F features which value uninterpretable F features on A, and that A does not select XP. Further, let us assume that A selects B and B selects C. Since agreement and selection instantiate functional dependencies, we obtain  $X \rightarrow A$  and  $A \rightarrow B$  and  $B \rightarrow C$ , yielding the linear order  $X > A > B > C$ .

(5)

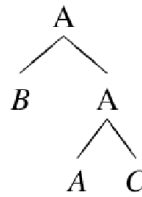
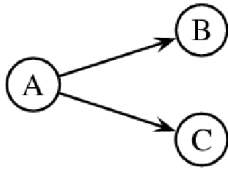


Note that the RPA (3) alone does not guarantee immediate precedence, which is a function of the locality requirement in (7). Situations may arise where a single constituent determines two or more dependents and underdetermines linear order. This is informally, graphically represented in (6), where a constituent A determines B and C. This kind of situation might arise, for instance, if a category c-selected a constituent as its complement and s-selected

<sup>7</sup> One might query how the PF interface “knows” for a given (A,B) pair, which one has interpretable features and which one has uninterpretable features. This concern is especially germane given that AGREE must check/delete uninterpretable features before the representation reaches PF. The answer lies in the nature of the representation that is passed to the interfaces. I am not suggesting that uninterpretable features are passed to the interface; I only say that AGREE, by virtue of its application, instantiates a dependency between A and B, yielding an (A,B) pair. Once the functional dependency has been established, the interfaces are blind to the means by which it was achieved – whether by AGREE or MERGE. Thus, even once the features are deleted, the dependency remains and it is this dependency that is passed to the interface.

another category in its specifier, as illustrated below. For instance, little *v* selects its complement *V* and also selects a DP subject in its specifier and assigns a theta role to it.

(6)



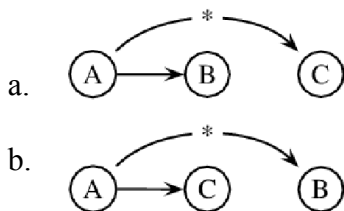
In this situation, the RPA (3) seemingly predicts that there is no total order between B and C, yielding either  $A > B > C$  or  $A > C > B$ . It is also reasonable to acknowledge that PF may include some type of constraint on linear locality, so I will assume the caveat on the RPA (3) that the PF object (A,B) must be spelled out as locally as possible.

(7) Relational Locality Condition (RLC):

- a. *p* should precede *q* as “closely” as possible;
- b. *p* is 0-close to *q* if *p* is immediately left-adjacent to *q*; *p* is 1-close to *q* if there is one category, *r*, between *p* and *q*, etc.;
- c. if *p* is not 0-close to *q*, then *q* incurs a violation of the RLC.

I prefer not to consider this locality an additional stipulation given that locality constraints are endemic in natural language, including phonology. So, regardless of whether I acknowledge them here, they will be independently necessary anyway.

(8)



In both the orders in (8), there is one violation of strict adjacency between A and that which it functionally determines. Thus for the order  $A > B > C$ , the constituent C is not strictly adjacent to A, and for the order  $A > C > B$ , the constituent B is not strictly adjacent to A either. It is reasonable to assume that under these conditions, which order is preferred is subject to PF parameterization<sup>8</sup>. Note, however, that this is not a violation of the RPA (3) per se because in each case the orders  $A > B$  and  $A > C$  are represented. Rather, these orders represent a violation

<sup>8</sup> It is an open question as to how linearization can be parameterized. According to Hauser et al. (2002), it may be that the only universal property of language is recursion. This leaves open the possibility that the interfaces, and SM/PF in particular, might lie outside the universal domain. Consequently, it is a (in my view, dispreferred) possibility that languages may be parameterized according to which linearization algorithm they use, e.g. some languages might utilize the LCA, others may utilize the system outlined in this paper. In my view, the more elegant solution is to have a single, universal linearization system yielding optimal outputs which are affected by language-specific variation in features.

of the assumed locality caveat that the constituents are as close as possible. Therefore, I propose that the locality caveat be treated along the lines of a violable constraint<sup>9</sup>.

### 3. Topicalization

This section explores some of the properties of English topicalization constructions as a tool to evaluate the linearization proposal made in the previous section. It will be demonstrated that topicalization constructions have some paradoxical properties that can be accounted for using the framework suggested here.

In English embedded topicalization constructions, an XP can be fronted into the field immediately to the right of the complementizer *that* (9a, 10a). In both cases, it is impossible for the phrasal constituent to precede the complementizer *that* (9b, 10b).

- (9) a. I said that John I like but Peter I don't  
 b. \*I said John that I like but Peter I don't
- (10) a. Al Capone said that whatever happened he'd never be convicted  
 b. \*Al Capone said whatever happened that he'd never be convicted

Embedded topicalization contrasts with WH-movement where it is known that the WH-word moves to SpecCP which is the field to the left of the complementizer – notwithstanding the fact that in English, the complementizer in these constructions is null as a result of what is traditionally known as the Doubly Filled Comp filter.

- (11) I wonder who (\*that) arrived?

Evidence that the WH-item is in SpecCP to the immediate left of C<sup>0</sup> comes from do-support constructions, where the dummy verb *do* is in C<sup>0</sup>.

- (12) What do you want?

There is also evidence from closely-related languages that a WH-item in a specifier position precedes the head.

- (13) a. For who **that** wolde him wel advise, What hath befalle in this matiere  
 'Whoever realizes well what in this thing some men befell' (Brodie 2007, own emphasis)
- b. Ik weet niet wie dat Jan gezien heft  
 I know not who that Jan prt-seen has  
 'I don't know who Jan has seen' (Haegeman 1991:382)

Example (13a) from Middle English shows a WH-item in SpecCP preceding the complementizer. Flemish also demonstrates the possibility of the WH-item preceding the complementizer (13b). To the extent that these examples from closely-related languages are generalizable to English, they demonstrate that WH-items in SpecCP precede the head of C<sup>0</sup>.

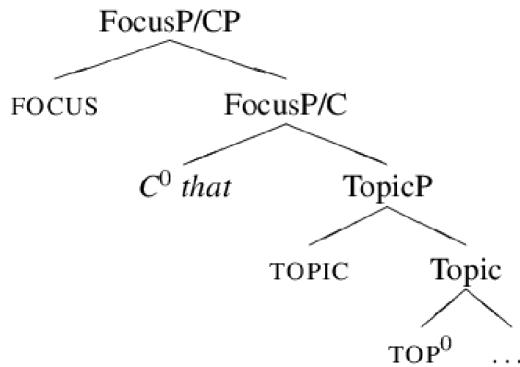
---

<sup>9</sup> See De Vos (2009a) for an exploration of how locality violations can account for the various orders in the adpositional domain in Afrikaans.



The generalization from the preceding examples is that embedded topicalization targets a specifier in the field to the right of the complementizer while WH-movement targets a specifier to the left of it. Assuming an articulated CP layer (Rizzi 1997) in conjunction with the LCA, one arrives at the conclusion that WH-items and topicalized XPs are hosted by different projections within CP, as exemplified by the structure below. This double-layer analysis is forced by the requirement of the LCA that specifiers precede heads.

(14) The double-CP analysis



This structure accounts for the relative linear orders of the complementizer, the WH-moved and the topicalized XPs. However, since there are two distinct positions for WH-moved and topicalized XPs, this type of analysis also makes fairly strong predictions about the independence of WH-movement and embedded topicalization. It is predicted that the two positions are independent and that movement to one will not necessarily block movement to the other. However, I will demonstrate that these predictions conflict with the empirical facts and that, indeed, the two positions show close symmetries.

First, there are some troubling aspects to the structure in (14): FocusP/CP and TopicP interact in curious ways: the generalization seems to be that for any sentence, there can be a maximum of one CP-layer head and one CP-layer specifier which can be realized phonetically, for example, the head of FocusP, *that*, interacts with the specifier of TopicP (9a, 10a). The nature of this “communication” between the specifier of one XP and the head of another remains mysterious: it is odd that topicalization of discourse-old information to SpecTopicP would necessarily require the overt lexicalization of FocusP<sup>10</sup>.

Second, this type of analysis requires that there be independent justification for why the head of TopicP is never lexicalized in English (this justification comes in the form of the Doubly Filled Comp filter, which is simply a stipulation)<sup>11</sup>.

<sup>10</sup> An anonymous reviewer queries whether this might not be a semantic restriction on the co-occurrence of topic and focus within a single construction. This seems unlikely because sentences typically manage discourse-old and discourse-new information within single constructions quite generally, as in the following example.

Q: Did you know Mary wore her socks in the shower?

A: Yes, she does that quite often.

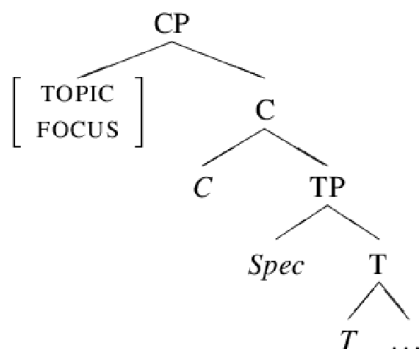
a. Which pair<sub>Focus</sub> did she shrink this time?<sub>Topic</sub>

Furthermore, even if such a restriction were to exist, the semantics still would not easily explain why fronting of old information to SpecTopicP would ensure the lexicalization of the head of FocusP; generally one would expect that the head of TopicP would be overt in these instances.

<sup>11</sup> Furthermore, to the extent that this structure is generalizable to verb-second (V2) languages (e.g. all Germanic languages except English), a problem arises for traditional accounts of V2 (Den Besten 1989). V2 languages are

The paradoxical conclusion I will reach is that, according to various tests, these two positions are actually one and the same position. Naturally, such a conclusion is incompatible with current theory because it is theoretically inconceivable that they should be the same position: the double-layer analysis in (14) is forced by the requirement of the LCA that specifiers precede heads. In other words, the tests to be presented in this paper suggest the structure in (15) which has only a single specifier position as a target for both WH-movement and embedded topicalization.

(15) The single-CP analysis



The structure in (15) also makes fairly strong predictions. If WH-movement and embedded topicalization indeed target the same position, then it is predicted that the two types of movement should be highly correlated and that one will block application of the other. On the other hand, the structure makes a prediction that the complementizer should always follow the topicalized XP, which obviously runs against the facts (9,10). Therefore, the nature of the paradox is that different empirical effects suggest incompatible analyses: evidence from the interdependence of the two types of construction suggests a structure which does not account for the relative word orders (15), while the structure that accounts for the relative word orders (14) does not account for the interdependence of the two types of constructions.

#### 4. Evidence for a single CP in English

In this section, I will evaluate the evidence for a single CP in English. It will be noted that the same evidence does not seem to apply to languages like Italian so those languages may indeed have a double CP. The double-layer and single-layer analyses in (14) and (15) respectively make distinct, testable predictions. It has been known, at least since Chomsky (1977), that WH-movement and topicalization have the same diagnostics in English. Additional evidence that SpecCP and SpecTopicP are, in fact, one and the same position is that:

---

typically characterized by an asymmetry where the presence of an overt complementizer blocks movement of the verb, thus resulting in the verb remaining in situ, which is clause-final in the Afrikaans examples below.

- (i) Ek weet dat ek van Jan hou  
I know that I of Jan like  
'I know that I like Jan'
- (ii) Ek weet (\*dat) ek hou van Jan

This is because the absence of V2 in embedded contexts is crucially dependent on the complementizer that is filling the only possible head to which a verb might move, thus precluding V2 in embedded clauses (14a). Since the analysis in the double-CP analysis posits the presence of a further head below the complementizer, the traditional account is unworkable; a verb could always raise to Topic<sup>0</sup> thus giving rise to V2 in embedded contexts.

- (i) Ek weet [CP [C<sup>0</sup> dat [TP ek [T<sup>0</sup> [AgrO van Jan [vP [v hou . . . ]]]]]]]
- (ii) Ek weet [CP ek [C<sup>0</sup> hou [TP [T<sup>0</sup> [AgrO van Jan [vP [v . . . ]]]]]]]

- (i) Movement to SpecCP blocks movement to SpecTopicP and vice versa (section 4.1),
- (ii) They both disallow stacking in English (section 4.2)<sup>12</sup>, and
- (iii) They both evidence that-trace effects (section 4.3).

Thus the filling of the two specifiers occurs in complementary distribution in a number of different contexts which leads to the, at first glance paradoxical, conclusion that they are one and the same specifier.

#### 4.1. Blocking effects

The double-CP structure in (14) has two specifier positions, one for WH-items (SpecCP) and one for topicalized elements (SpecTopicP), leading to a prediction that WH-extraction should be possible even in the presence of a topicalized element.

- (16) a. Jake thinks that Peter you gave a book, but Sarah you gave a CD  
 b. What did you think that Jake gave Peter?  
 c. \*What did you think that Peter Jake gave?

Example (16a) shows a topicalized DP following the complementizer, whereas (16b) demonstrates that WH-extraction is possible in the absence of embedded topicalization, presumably using SpecCP as an escape hatch. Interestingly, (16c) is ungrammatical, suggesting that the embedded SpecCP cannot be used as an escape hatch when TopicP is filled. The inverse also applies: it is not possible to apply long-distance topicalization over an embedded WH-item, as illustrated in (17b).

- (17) a. I wonder when I gave Peter that book?  
 b. \*A book I wonder when I gave?

The structure in (14) also predicts that topicalization and WH-movement should be possible within the same clause, and again this is shown to be inconsistent with the facts (18b). The ungrammaticality of (18b) follows from a structure with only a single SpecCP.

- (18) a. I wonder when I gave Peter that book?  
 b. \*I wonder when Peter I gave that book?

The double-CP analysis derives these facts with the additional assumption that Relativized Minimality blocks moving a WH-item across a topic. There are a number of concerns with this approach. First, note that since this is an additional assumption, Occam's razor rules it out. Second, blocking effects are language specific (e.g. Spanish (Rivero 1978) and Italian (Delfitto 2002) do not have blocking effects) which suggests that the features responsible for the intervention effects are language specific; but Topic and Focus features are universal and rooted in semantics, so it is not easy to see how these could be made language specific. Third, it is not clear why Topic and WH-features should cause intervention effects since they have nothing in common: a WH-item bears information focus whereas a topicalized element could be a non-focused topic. Furthermore, when a constituent in a sentence has a topic feature but has not moved (e.g. when topichood is indicated by stress or by pronominal status), then no

<sup>12</sup> Note that it is a particular fact of English that SpecCP and SpecTopicP cannot be stacked, yet Italian, for example, allows stacked topics (Cinque 1999; Rizzi 1991, 1997). The fact of cross-linguistic variation in this regard does not necessarily preclude this fact being used as a diagnostic in a language in which it holds, namely English.

intervention effect is caused (19b), indicating that it is not simply the presence of an intervening TOPIC feature which causes ungrammaticality.

- (19) A: John came around to visit me last night.  
B: Oh yeah? What did HE want?

Furthermore, if one assumes that the WH-word *why* can be base generated in CP without movement (Ko 2006, Rizzi 1991), then Relativized Minimality should, in principle, be bypassed, leading to grammaticality. However, the data do not bear this out: in (20), a topic DP has moved to SpecTopicP, while *why* is base generated in SpecCP without having to cross over the embedded topic, but the result is still ungrammatical. These data strongly suggest that Relativized Minimality is not at play. Culicover (1996) cites evidence to show that blocking effects can be circumvented by various means including the use of stress. As Culicover (1996:458) notes, these exceptions are similar to those noted by Pesetsky (1987) where Relativized Minimality can be violated in WH-contexts. Pesetsky's D-linking effects are not taken generally as evidence against Relativized Minimality. As such, Culicover's exceptions prove the rule and his arguments against the double-CP model are, in spirit, similar to those made here.

- (20) \*I wonder why John I like, but Peter I don't?

In contrast, assuming only a single SpecCP (15) predicts that WH-extraction should not be possible in the context of topicalization as the data demonstrate.

#### 4.2. Stacking

Under the standard assumption that SpecCP cannot be recursively stacked in English, the single-CP structure (15) predicts that both multiple WH-movement (21b) and multiple topicalization (21c) are ungrammatical in English.

- (21) a. I gave Peter a book and Sarah a CD  
b. \*Who what did I give and Sarah a CD  
c. \*Peter a book I gave, but I gave Sarah a CD

These data strongly support a view that SpecCP serves both as a location for WH-items as well as for topicalized DPs, and that the double-CP analysis (14) does not fit the data without additional assumptions being made.

Rizzi (1997) argues that TopicP can be stacked, at least in Italian. This, again, predicts the independence of topicalization and WH-movement which is not borne out in English. In English, multiple embedded topics are allowed if they are adjuncts, but the multiple topicalization of arguments is disallowed. This suggests that there is only one specifier position for TopicP, the remaining topicalized constituents being adjuncts<sup>13</sup>.

<sup>13</sup> Culicover (1996) argues that multiple topicalization in English is possible and shows evidence to this effect. I do not agree with several of his grammaticality judgments. Additionally, many of his examples require a special intonation; Pesetsky (1987) also noted the effects of intonation and D-linking in evading WH-islands, but that is not to say that WH-islands do not exist. Moreover, he notes that multiple topicalization cannot apply to two NPs (i.e. arguments which cannot adjoin to IP) but seems to be limited to XPs such as PPs. He suggests that topicalized XPs "require identification" (Pesetsky 1987:454). My interpretation of these facts is that this is consistent with (i) a single landing site for topicalization, from which an A-bar-bound XP can be "identified"

### 4.3. That-trace effects

Another direct prediction of the double-CP approach is that that-trace effects need not apply to both constructions. Since there are two distinct specifier positions in the CP layer, each complementizer is theoretically independent. Consequently, it might be expected that that-trace effects apply to only  $C^0$  and not necessarily to  $\text{Topic}^0$ . In contrast, for a single-CP analysis, if  $C^0$  and  $\text{Topic}^0$  fill the same head position then it is expected that both constructions will evidence the that-trace effect.

- (22) a. Who did I say gave Peter a book?  
b. \*Who did I say that gave Peter a book?

The pair in (22) illustrate well-known that-trace effects: an embedded subject can only be extracted in the absence of the complementizer. I do not have an explanation for this effect (but see section 6); I am simply using it as a diagnostic here. Similarly, under long-distance topicalization, a that-trace effect is visible (23b).

- (23) a. You, I said gave Peter a book  
b. \*You, I said that gave Peter a book

- (24) John, I know (that) Sarah likes *t*, Peter, I'm not so sure about

To complete the paradigm, when an object undergoes long-distance topicalization, the that-trace effect is not present (24). This demonstrates that this effect is indeed a that-trace effect specific to subject extraction and is not simply a product of long topicalization per se.

Note that a double-CP analysis could accommodate these data with an additional, reasonable assumption, namely that the effect is due to the property of being a subject which is the same for both positions. However, in a double-CP analysis the complementizer that heads FocusP/CP which does not immediately dominate TP means that there cannot be any direct interaction but must be "action at a distance". If one were to attempt to argue that in these constructions  $\text{TopicP}$  is optional and that FocusP/CP is merged directly to TP, then one would be arguing for, in essence, a single-CP model anyway. Consequently, although a double-CP analysis may be able to accommodate the data, it can only do so at some expense.

It appears that, for at least some examples and speakers, such action at a distance is possible insofar as long topicalization of any constituent across an overt complementizer is degraded, as illustrated in (25) (cf. Maki et al. 1999).

- (25) a. Herself, Mary says (\*that) she would never endanger  
(adapted from Culicover 1996:452)  
b. Mary says that John, she doesn't know but (\*that) she'd like to see drunk  
(Rochemont 1989, cited by Authier 1991)

---

through reconstruction to its trace, and (ii) the possibility of adjunction of additional adverbial or PP material to IP (cf. Ernst 2002). This correctly rules out topicalization of multiple NPs while allowing a single NP to be topicalized in combination with various adjuncts.

Yet, although these examples seem quite clear, there are others which seem considerably less ungrammatical (26), thereby suggesting that action at a distance is not available in at least some contexts.

- (26) a. John, I know (?that) Sarah likes, but Bill I'm not so sure about  
 b. John, I think (?that) Sarah likes, but Bill I'm not so sure about

Additional empirical work is needed before there is consensus on the facts. Nevertheless, two generalizations can be made despite the variation: (i) the extraction of subjects is more marked than the extraction of objects (27); (ii) additionally, long topicalization of a non-subject is more marked than WH-extraction of a non-subject across an overt complementizer (28), an issue that will be discussed in section 6.

- (27) a. This book, I think (\*that) impressed John  
 b. This book, I think (?that) you would like
- (28) a. ??This book, Bill thinks that I like  
 b. Which book does Bill think that I like?

## 5. An explanation for the embedded topicalization paradox

**Table 1:** Summary of evidence for and against the single- and double-CP analyses of topicalization, respectively

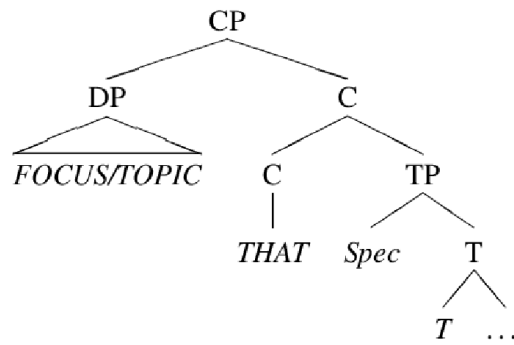
Test	Test explained by the double-CP analysis?	Explained by the single-CP analysis?
Word order	Yes	No
Blocking effects	No	Yes
Stacking	No	Yes
That-trace effects	No	Yes

The evidence from the preceding sections is summarized in Table 1. From this it is clear that both the double- and the single-CP analyses fall short in some respects. What makes the double-CP analysis more palatable is that (i) it is the only theoretical possibility allowed by the LCA and (ii) the blocking, stacking and that-trace effects can presumably be handled by ancillary assumptions and stipulations. On the other hand, the previous sections showed that there is ample evidence to suggest that there is only a single CP layer. Yet, the assumption of the LCA makes a single-CP analysis simply untenable from a theoretical perspective. Consequently, this type of analysis has been spurned for reasons of word order alone, regardless of the other empirical facts. However, if an alternative linearization algorithm could be used to account for the word-order facts in a constrained manner, then the single-CP analysis might be preferable because it accounts for the other data without additional stipulations. In this section, I will explain how the linearization algorithm (3) outlined in section 2 resolves the paradox.

### 5.1. Deriving Spec-head order in a WH-construction

In order to derive a WH-construction using the proposed framework, let us assume the single-CP structure (29), where a single CP has a DP in its specifier.

(29)



In order to derive the relative order of head and specifier in an embedded WH-construction, it is necessary to have phonetically null THAT in the numeration or the derivation will ultimately crash. The derivation proceeds normally until  $C^0$  is merged (30a)<sup>14</sup>. There must be some kind of relationship between the WH-moved DP and the head of  $C^0$ . It is commonly assumed that a WH-word has an interpretable *iWH* feature which is checked against the corresponding *uWH* of  $C^0$ . The WH-item moves to SpecCP and checks its *iWH* feature against its uninterpretable counterpart on  $C^0$  (30b).

- (30) a. [<sub>C</sub> THAT(*uWH*) [<sub>TP</sub> I [<sub>T</sub> will [<sub>VP</sub> invite [<sub>VP</sub> who(*iWH*) ]]]]]  
 b. [<sub>CP</sub> who(*iWH*) [<sub>C</sub> THAT(*uWH*) [<sub>TP</sub> I [<sub>T</sub> will [<sub>VP</sub> invite [<sub>VP</sub> t ]]]]]]]  
 c. Spelled out as: I wonder who  $\emptyset$  I will invite

According to (2b), there is a functional dependency such as *who*[*iWH*] → *C*[*uWH*]. When AGREE takes place, this functional dependency is effectively instantiated by AGREE (even though the *uWH* feature is ultimately “valued” or “deleted”) and is eventually transferred to PF. The linearization of this *who*[*iWH*] → *C*[*uWH*] relation is done by the RPA (3), and the result is *who* > THAT. Thus, the functional dependency results in the linearization of the WH-word preceding the null complementizer<sup>15</sup>.

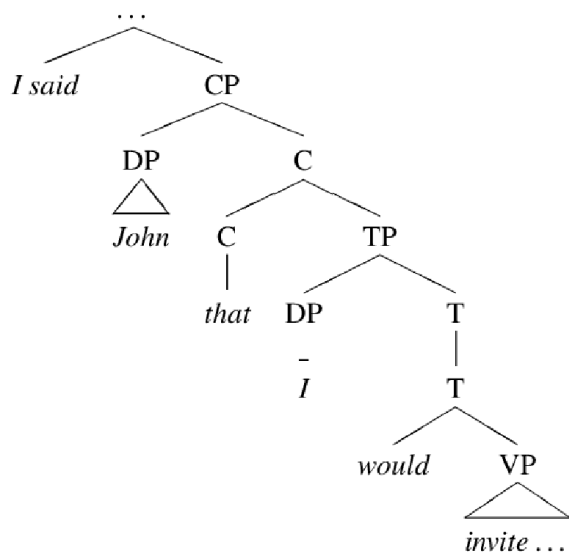
## 5.2. Deriving head-Spec order in an embedded topicalization construction

In order to derive an example of embedded topicalization, assume a structure with a single CP with an XP merged in its specifier.

<sup>14</sup> There are additional steps in the derivation involving movement to phases, etc. These steps have been omitted in the interests of simple exemplification. Their inclusion or exclusion does not affect the analysis.

<sup>15</sup> For conciseness, I will not outline the derivation of the remainder of the sentence in detail here.

(31) a.



b. I said that John I would invite but Peter I wouldn't

It is immediately clear that under the LCA, the complementizer should follow the topicalized DP, predicting an ungrammatical word order. The derivation of an embedded topicalization construction using the proposed framework requires the overt complementizer *that* in the numeration or the derivation will crash. For the sake of argument, assume that the complementizer *that* bears *iTopic* features and that the DP topic has uninterpretable *Topic* features. The derivation proceeds normally until  $C^0$  is merged (32a). The DP bearing a *uTopic* feature moves to SpecCP and checks its features against those of the complementizer by means of AGREE (32b). This operation instantiates a functional dependency such that  $C[iTopic] \rightarrow DP[uTopic]$  (2). In turn, by (3) this is linearized such that  $that > John$  and is spelled out in that order (32c).

- (32) a.  $[C\ that(iTopic)\ [TP\ I\ [T\ would\ [vP\ invite\ [VP\ John(uTopic)\ ]]]]]$   
 b.  $[CP\ John(uTopic)\ [C\ that(iTopic)\ [TP\ I\ [T\ would\ [vP\ invite\ [VP\ t\ ]]]]]]$   
 c. Spelled out as: I said that John I would invite . . .

These derivations demonstrate that it is, in principle, possible to use a single-CP analysis for both WH-movement and embedded topicalization constructions. Whether the XP in SpecCP is linearized to the left or the right of the complementizer depends on its particular feature configurations. In addition, this linearization must occur in predictable and constrained ways.

### 5.3. Tentative evidence for *uTopic* on DP

Before continuing, recall that in my argument above, I simply made the assumption that a topicalized DP has *uTopic* features while the overt complementizer *that* has *iTopic* features. This particular feature configuration is essential in order for the analysis to function. I would like to offer the following tentative evidence for this feature configuration.

First, the uninterpretable/interpretable distinction is primarily a semantic one. Features which are interpretable on a constituent do some semantic work, whereas uninterpretable features do not seem to contribute to meaning in obvious ways. Unfortunately, this argument is not necessarily straightforward for topicalization, which means that the onus for arguing for a feature being



interpretable or uninterpretable falls on other arguments. For example, singular and plural nouns are denotationally distinct: a plural noun refers to something very different in the world in comparison with a singular noun. The same is not necessarily true of topics: whether a constituent is a topic or not does not affect which referent is being denoted. Thus, in (33a), the constituent *John* is focused while in (33b) it is a topic. But both propositions are identical in all possible worlds, that is, there is no possible world where (a) could be true and (b) could be false. This may suggest that, given the ambiguous nature of the evidence, different languages may sometimes parameterize Topic as interpretable on the DP and uninterpretable on the complementizer or vice versa, giving rise to parametric variation. As far as I can tell, there is no intrinsic semantic reason for preferring uTopic on the Topic DP as opposed to Topic<sup>0</sup>.

- (33) a. Who came to your office for a consultation?  
           Answer: John came  
       b. Did John come to your office for a consultation?  
           Answer: Yes, he<sub>John</sub> came

Second, Bošković (2007) argues that movement is always conditioned by an uninterpretable feature on a c-commanded probe. In a configuration like (34), a constituent with an uninterpretable feature is located within a phase. Since the PROBE has not been merged by the time the phase must be spelled out, it follows that the uninterpretable feature, if it remains within the phase, will cause the derivation to crash. From this, Bošković (2007) deduces that an uninterpretable feature must obligatorily move to the phase edge.

- (34)   iF           [PHASE           uF        ]  
         PROBE                   GOAL

If one accepts the argumentation offered by Bošković (2007), then movement is conditioned by uninterpretable features on the moved constituent. Applying this to topicalization, it follows that there must be an uninterpretable feature on the moved topic.

A third argument revolves around overt reflexes of uninterpretable feature checking. It happens to be a fact about English that uninterpretable features often have PF manifestations, whereas interpretable features tend not to.  $u\phi$  on T exhibits a PF reflex, namely morphological marking; uNumber on N is spelled out with the plural morpheme */-S/* (Van Koppen et al. 2009); uT on D is spelled out as nominative case (Pesetsky and Torrego 2001) and, in non-tensed environments, uCase on DPs is spelled out as accusative, genitive, dative, etc. The PF reflexes of uninterpretable features are not limited to affixes: if one looks at question formation, then the question complementizer which has uWH features is expressed at PF as a null complementizer (see discussion above). WH-items themselves carry both interpretable and uninterpretable features (iWH and uQ (Bošković 2002, Hagstrom 1998)) and WH-morphology

could be seen as a reflex of  $uQ$ <sup>16</sup>. Thus, it is a fairly robust generalization that, in English, uninterpretable features often result in PF manifestations on the categories that carry them<sup>17</sup>.

Given this, it is arguable that embedded topics also carry a PF manifestation of an uninterpretable feature in the form of stress. For many speakers, embedded topicalization requires an obligatory special intonation in order to license it (35a) whereas the same sentence with neutral intonation is marked (35b). To the extent that this is true, it may be evidence for  $uTopic$  features on the DP topic<sup>18</sup>.

- (35) a. I said that JOHN I like  
b. \*I said that John I like

A fourth and related argument is that languages which do mark topichood with a morphological reflex tend to do so on the topic constituent and not on the complementizer. An example of such a language is Japanese which displays a PF reflex on the topic in the form of a topic marker e.g. *wa*.

Finally, it is worth contrasting English topicalization with that of Italian. Delfitto (2002) argues that in Italian, the topicalized DP bears an interpretable  $iTopic$  feature, i.e. he argues for the opposite configuration in Italian to what I am arguing for English here. Interestingly, Italian has very different topicalization options to English. Unlike English, it allows for multiple topics and does not exhibit blocking effects. Since these options are excluded for English, it suggests that the two languages may differ according to where the uninterpretable feature is located. Therefore, it is reasonable to suggest that English has an uninterpretable topic feature on the moved constituent whereas Italian locates the uninterpretable topic feature on the complementizer.

#### 5.4. Speculations on complementizer effects

This section is more speculative in character and attempts to shed light on the Doubly Filled Comp effect, that-trace effects, and the ways in which adjuncts can inhibit that-trace effects.

<sup>16</sup> It might be considered whether these features undermine the current analysis or not. If WH-items carry  $iWH$  and  $uQ$  features and Topics carry  $uTopic$  features, then both Topics and WH-items carry at least some uninterpretable features which, at first glance, makes them identical for purposes of linearization: both must follow their heads at some level of representation. If they were indeed identical in this way, then the current analysis would fail. Note, however, that a WH-item contains the pair  $uQ$  and  $iWH$  and the complementizer contains  $iQ$  and  $uWH$  which sets up a linearization paradox, namely that the WH-item must both precede the complementizer (because  $WH[iWH] \rightarrow C[uWH]$ ) and follow it (because  $C[iQ] \rightarrow WH[uQ]$ ). This paradox is resolved by the following linearization schema where the WH-item forms a chain:  $WH[iWH, uQ] > C[uWH, iQ] > WH[iWH, uQ]$ . In contrast, for topicalization, no such pairing of features is postulated and so Topics and WH-items remain non-equivalent in terms of their feature configurations and consequently are linearized differently.

<sup>17</sup> The inverse may not be true because  $iT$  on  $T^0$  appears to be spelled out as tense morphology. However,  $iT$  may itself be a composite feature: within the framework of Giorgi and Pianesi (1997), developing work by Reichenbach (1947), the “upper” Tense node instantiates a relationship between S(peech Time) and R(eference Time), where both S and R are variables to be bound by appropriate “times”. R in particular can be bound by a syntactic constituent (e.g. a temporal adjunct) and if one theorises binding of this type as feature checking, then it follows that  $T^0$  includes (at least) a  $uR$  feature. Consequently, the generalization stated above would hold even for  $T^0$  but I leave this to future research.

<sup>18</sup> The argument that English topics bear an uninterpretable feature is contra Delfitto (2002) who argues the opposite for Italian. Note, however, that Italian embedded topicalization is different to that of English because it allows for multiple topics and does not exhibit blocking effects.

### 5.4.1. The Doubly Filled Comp effect

The proposed framework may shed light on the Doubly Filled Comp effect. In embedded topicalization constructions, the double-CP analysis gives the impression that the upper complementizer is obligatorily present, while the head of TopicP is obligatorily absent.

- (36) a. I said \*(that) John I would invite, but Bill I wouldn't  
 b. Julia thinks \*(that) in all likelihood, David will invite Elizabeth (Delfitto 2002)

This creates a parallel with Subject-WH constructions where the head of CP[WH] is also obligatorily absent. The proposal in this paper simplifies the matter slightly: the head of CP[WH] is indeed phonetically null, however the head of CP[TOPIC] is spelled out overtly as *that*, albeit in a position preceding the topicalized constituent. Consequently, topicalization is not subject to the Doubly Filled Comp filter. This phenomenon thus reduces to a particular lexical fact about English, namely that C[uWH] happens to be spelled out as phonetically null whereas C[TOPIC] is spelled out overtly<sup>19</sup>.

- (37) a. *that* is specified for i $\phi$  (cf. Van Craenenbroeck and Van Koppen 2002) and iTopic features (cf. Delfitto 2002 for Italian) and  
 b. THAT, spelled out as phonetically null, is specified for uWH (and perhaps u $\phi$ ) features.  
 c. These features may vary parametrically.

### 5.4.2. That-trace effects

The proposed analysis also provides some interesting insights about that-trace effects (examples (22), (23), (24)). In (37) it was proposed that *that* is specified at iTopic, i.e. topichood, if present in a clause, is interpretable on this particular complementizer. The RPA (3) ensures that *that* will precede a constituent bearing uTopic features. From this, it is a short step to suppose that the complementizer *that* may mark its immediate complement (in the instance of its specifier being empty) as being a topic. In other words, there is a field immediately following the complementizer that is filled by a discourse topic<sup>20</sup>. Taking this for granted for the moment allows an explanation of the that-trace effect.

An overt complementizer *that* will mark the subject in the field immediately following it as being a topic. Extraction of a focused WH-item from a non-subject position is consequently unproblematic (38a) since the focused WH-item does not originate in the topic field. WH-movement of the object uses the specifier of the complementizer as an escape hatch in a purely

<sup>19</sup> The current analysis offers the delightful possibility of being able to explain why the question complementizer THAT happens to be null. If one assumes that the complementizer has a uWH feature as well as an interpretable iQ feature (i.e. THAT[uWH,iQ]), and that the WH-word has a uQ feature in addition to the iWH feature (i.e. WH[uQ,iWH]), it follows that the question word will both precede and follow the complementizer, leading to a linearization paradox. One way to resolve the paradox is to spell the complementizer out as phonetically null if the language has the morphological resources to do so, therefore obviating the paradox. Thus, the null character of the question complementizer may not have to be lexically stipulated but may follow from the postulated output conditions. On the other hand, in topicalization constructions, there are no uninterpretable features on C which means that there is no linearization paradox and the complementizer can be spelled out overtly. An anonymous reviewer notes that this type of analysis could not carry over to the Middle English data discussed earlier. This possibility is left for future research.

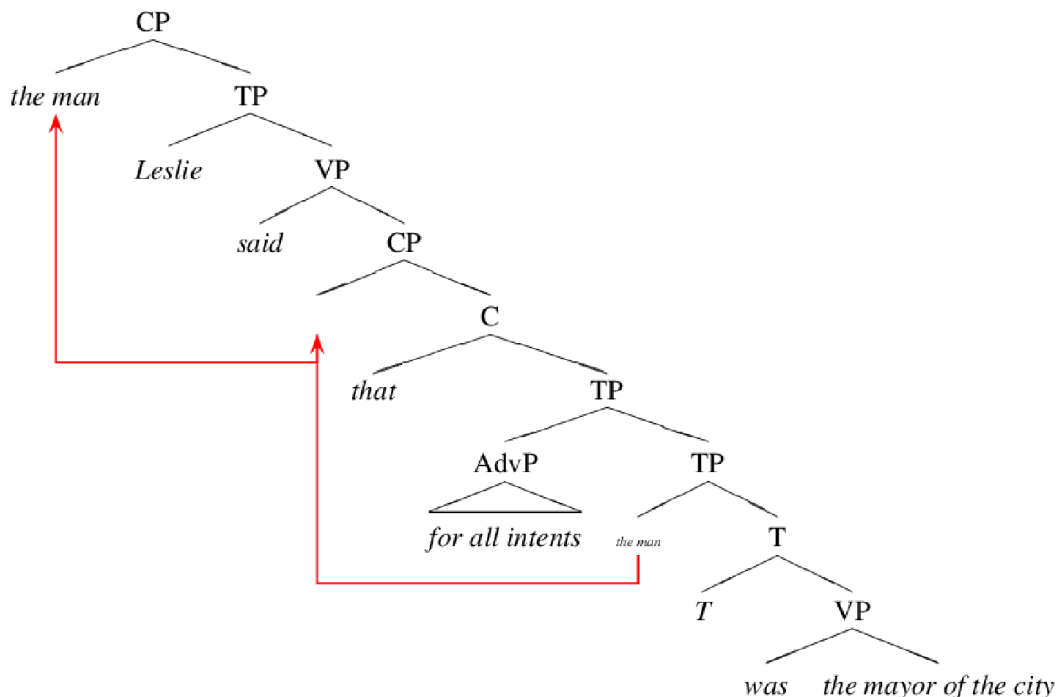
<sup>20</sup> Note that in terms of the analysis proposed in this paper, a complement and specifier are structurally distinguished, but that in terms of linear order, both the complement and the specifier of a *that* head would follow the head, making them superficially indistinguishable.



This line of reasoning makes a prediction that if a constituent is adjoined to IP in the field to the right of *that*, it will be marked as a topic, thereby checking the iTopic feature on *that*, leaving SpecCP open as an escape hatch for subsequent extraction (42). In other words, adjunction of an XP to IP will void that that-trace effect. This is what happens with the so-called “adverb effect”.

- (41) a. Who did Leslie say that, for all intents and purposes, *t* was the mayor of the city?  
 b. Robin met the man Leslie said that, for all intents and purposes, *t* was the mayor of the city (Browning 1996, cited in Delfitto 2002:57)

(42)



## 6. Broader implications for English Spec-head relations

While the initial lines of investigation look promising, before one can reasonably accept an analysis with such wide-reaching implications, it is necessary to see if it can be applied consistently to other constructions – at least within the language in question, namely English. While I acknowledge that, ideally, one might like to evaluate the RPA (3) with respect to as many languages, dialects and constructions as possible, it is not possible in this paper for reasons of space and will be left to future research. The following subsections attempt to demonstrate that the RPA (3) is consistent with other Spec-head relations in English.

### 6.1. The linear orders of specifiers and heads in English

The proposed framework makes predictions that must apply more generally to English. In particular, it predicts the possibility of some specifiers being linearized either to the left or right of their heads, depending on whether a feature of the specifier selects the head or vice versa. This section explores these possibilities with respect to a number of categories in English. Here, I will show that it is not entirely clear that specifiers do in fact precede their heads in linear

---

head of SpecCP, into the matrix clause. At PF, the moved DP is prosodically marked, and at LF the moved DP is given an interpretation corresponding to topichood.

order in all situations, as their underlying configuration may be obscured by subsequent operations.

The relative order of heads and specifiers in the CP domain has already been identified in section 3: SpecFocusP(WH) precedes its head and SpecTopicP follows its head. This section evaluates the orders between English heads and specifiers in four other major domains, namely TP/IP, VP,  $\nu$ P and PP. Of these, the only clear instances of a specifier preceding the head are in the TP and the CP Focus domains. All others seemingly display head-Spec orders, as predicted by the proposed analysis.

### 6.1.1. SpecTP

The canonical example of a specifier which precedes its head is SpecTP.

- (43) a. John will probably come  
c. John has probably come

The subject precedes auxiliaries and modals (43a,b). On the understanding that the adverb in the previous sentences marks the left edge of the  $\nu$ P, the fact that the modal and auxiliary are themselves to the left of the adverb shows that these verbs are in  $T^0$ , with the resulting implication that the subject is indeed in SpecTP. This follows from the RPA (3) because  $u\phi$  on T is checked by  $i\phi$  on DP. Consequently,  $D[i\phi] > T[u\phi]$ .

There is also a case dependency running in the opposite direction where  $iT$  on T checks  $uT$  on D (Pesetsky and Torrego 2001). Thus, according to the RPA (3),  $T[iT] > D[uT]$ . Putting these two linearizations together, we derive  $D[i\phi] > T[u\phi, iT] > D[uT]$ . The two copies of D constitute a chain and the highest one is spelled out in line with standard assumptions about chains<sup>22</sup>. Thus, D is spelled out before T.

### 6.1.2. SpecVP

Turning our attention to the VP, when V selects a clausal complement, the direct object is in a specifier of V in examples where a verb selects a clausal complement (44) (cf. Broekhuis 2008, Barbiers 2005)<sup>23</sup>.

- (44) a. I told John **t** that I don't like Peter  
b. \*I John told that I don't like Peter

However, "big" V appears to precede its specifier, resulting in V-O word order (44a), with the O-V word order that represents Spec-head order being ungrammatical (44b). The same applies in passive contexts (45a,b).

The V-O order is arguably due to short  $\nu$ -V raising which moves V to a position preceding the object. So the underlying Spec-head, O-V pattern is obscured by subsequent movements

<sup>22</sup> This derives EPP in agreement contexts (an early implementation addressing EPP (De Vos 2009b) can be found online at <http://ling.auf.net/lingbuzz/000971>). It also predicts that Case is checked either in situ in Spec $\nu$ P or in a Case-assignment position *below* T. This derives the notion that the TP layer is actually a composite where the lower projection checks Case while the higher projection checks Agreement.

<sup>23</sup> It follows that if a verb selects a clausal complement, then the direct object cannot be a complement in those constructions. For this reason, these constructions are useful in ensuring that the object is in a specifier position.

(Barbiers 2000)<sup>24</sup>. Consequently, there is no direct evidence that SpecVP precedes its head; derived surface order shows that the V head ends up preceding the specifier of VP.

- (45) a. I have beaten John  
b. \*I have John beaten

It might be countered that in pseudo-causative contexts, the DP object precedes the most deeply embedded verb, as in (46a). However, in this example, it is not at all clear that the DP object is in the specifier of VP since an adverbial can intervene between the DP and the participle verb (46c).

- (46) a. I got/had John beaten  
b. \*I got beaten John  
c. I got John soundly/well-and-truly beaten

Additional evidence that the DP object in (46a,c) is not in SpecVP comes from extraction facts. Example (47a) demonstrates that argument *picture* DPs are not islands when they occur in their base position to the right of the verb. However, in the pseudo-causative construction, when the *picture* DP occurs to the left of the participle it is an island (47b). This demonstrates that for the pseudo-causative, O-V order, the DP is in a derived position and is thus very unlikely to be the specifier of the most deeply embedded VP.

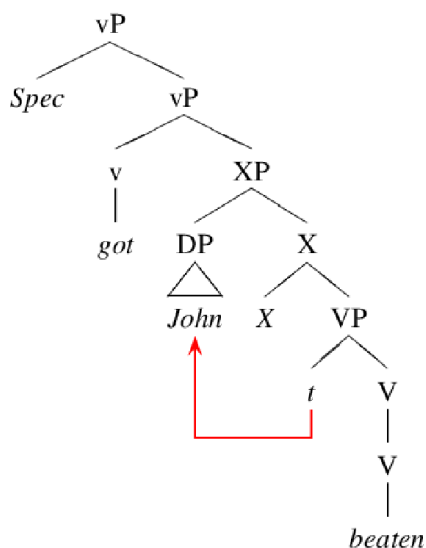
- (47) a. Who did you take a picture of *t* last week?  
b. \*Who did you get a picture of *t* taken last week?

This raises an interesting problem: the adverbial intervention and extraction facts show that the DP is not in SpecVP, yet since the DP follows the pseudo-causative verb, it cannot (according to the LCA) be in the specifier of the pseudo-causative verb itself, but must be in the specifier of a null (and unidentified) head. This is illustrated in (48).

---

<sup>24</sup> However, it is worth noting that, to date, there is no widely accepted reason for why *v*-V raising should occur at all in English and it appears to be used simply to derive the correct word order. That *v*-V raising is language-specific is highlighted by the fact that it putatively does not occur in languages like Dutch (Barbiers 2000) and Afrikaans (De Vos 2005).

(48)



The upshot of all of this is that there is little evidence to suggest that the specifier of VP precedes V; the O-V order is obscured by subsequent movement<sup>25</sup>. This is consistent with the RPA (3) because V selects its object and therefore determines it, thus  $V > O$ .

### 6.1.3. SpecvP

In English, the lexical verb undergoes short head movement to little *v*. Little *v* also seems to precede its specifier in existential constructions which are usually analysed as having the subject remain in its in situ position in SpecvP.

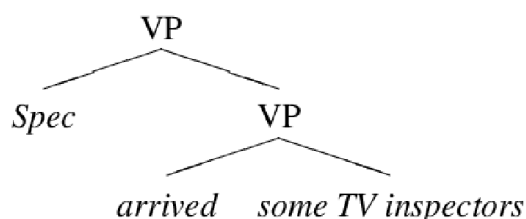
(49) There arrived **some TV inspectors** at the door

This word order must be explained in one of two ways. First, it could be postulated that V-raising occurs to some position between T and *v* (cf. Jonas and Bobaljik (1993) for Icelandic data to this effect). However, this is less than convincing because English is notable for lacking V-raising out of vP (unlike French for example (cf. Emonds 1978)). The second option would be to argue that the class of verbs allowing existential constructions (i.e. unaccusatives) does not have an articulated vP shell but rather consists simply of VP with the DP as the complement (Alexiadou et al. 2004). If the unaccusative verb lexicalizes the head, then the head would precede the DP.

<sup>25</sup> Furthermore, it is interesting to note that in both declarative and pseudo-causative contexts, in order to maintain the idea that specifiers precede heads in English, movement to null heads is necessary. I hasten to add that this observation does not amount to a denial of the possibility of null categories in contexts where they are motivated.



## (50) The structure of unaccusatives



However, this raises issues of its own. First, note that this analysis leaves the specifier of VP obligatorily empty and thus provides no evidence that the specifier precedes the head. Second, if the DP remains in situ, then there remains the question of why this is necessarily the case, especially when SpecVP remains a viable landing site. In summary, there is little evidence to suggest that SpecVP precedes its head; once again, the underlying S-V order is obscured by subsequent movements. Nevertheless the surface order is consistent with the RPA (3) because  $v$  selects and  $\phi$ -marks its subject, and therefore  $v$  determines its subject. Thus, by the RPA (3),  $v > \text{Subject}$ , which is indeed the case as the evidence above shows<sup>26</sup>.

## 6.1.4. SpecPP

Prepositional phrases occupy an awkward position among the typology of English categories, apparently being the only specifier-less phrase in what is otherwise a language that consistently requires them. The only phrasal material available in a PP is the complement of P which follows P in linear order<sup>27</sup>.

## (51) PP

- a. John is **in the kitchen**
- b. \* John is **the kitchen in**

One type of analysis is that P selects a DP complement, although this does not explain why SpecPP is systematically and obligatorily empty in English<sup>28</sup>. The other type of analysis holds that there are substantial movements within an articulated structure that eventually yield the correct word order but which obscure the underlying Spec-head order in the process (cf. Den Dikken 2008; Kayne 2005, 2001; Koopman 2000). Thus, it appears that the PP domain also does not provide evidence for Spec-head linear orderings in English. However, once again, the word-order facts are consistent with the RPA (3): P selects a DP argument and iCase on P<sup>0</sup> checks uCase on the DP which means that P determines its DP complement<sup>29</sup>. Therefore, by the RPA (3),  $P > \text{DP}$  which is what the data show.

<sup>26</sup> It has been suggested by a reviewer that transitive expletive constructions in Belfast English may be problematic for a Kaynian proposal (cf. Henry and Cottell 2007, and the response from Borsley 2009). Transitive expletives (or indeed, expletives of any kind) would be problematic if the verb and the associate were in situ in  $v^0$  and SpecVP respectively, with the expletive in SpecTP, yielding constructions like [<sub>TP</sub>There [<sub>VP</sub> many people [<sub>v^0</sub> read the book]]]. The current proposal predicts that should such constructions exist, they would be linearized as *There read many people the book*.

<sup>27</sup> The DP object of a preposition is usually assumed to be a complement. However, it is not clear why it could not be a specifier of P as, by analogy with VP, the object of V is in SpecVP when it is not a complement. Thus, PPs could, under certain theoretical assumptions, be taken as evidence that the P head precedes its own specifier, just as V precedes its specifier.

<sup>28</sup> In Principles and Parameters theory, the specifier of PP was taken to be occupied by modifiers such as *just*.

<sup>29</sup> Interestingly, it does not matter whether P selects a DP complement or a DP in its specifier: the word order remains  $P > \text{DP}$ .

### 6.1.5. Intermediate summary

In this section, I have reviewed some of the main categories in English. The results are listed in the left-hand column of Table 2.

**Table 2:** Linear orders of specifiers and heads match direction of syntactic dependency

Orderings	Nature of Spec-head relationship	Direction of dependency
[SpecTP DP] > T <sup>0</sup>	AGREE: TP checks $u\phi$ on T $\phi$	→ ←
[SpecCP WH] > C <sup>0</sup>	AGREE: DP checks $uWH$ on C	→
C <sup>0</sup> > [SpecCP DP <sub>Topic</sub> ]	AGREE: that checks $uTOPIC$ on DP	→
V <sup>0</sup> > [Spec $\nu$ P DP]	Selection: V selects object	→
$\nu^0$ > [Spec $\nu$ P DP]	Selection: $\nu$ selects subject	→
P <sup>0</sup> > DP	Selection: P selects object of preposition	→

The data show that SpecTP precedes T<sup>0</sup> (section 6.1.1) and that SpecCP[WH] also precedes C<sup>0</sup> (section 3). I demonstrated that each of these orders follows from the RPA (3) because in each of these cases an uninterpretable feature on the head is checked by a corresponding interpretable feature on the phrase in the specifier.

With the exception of SpecTP and SpecCP[WH], there is little empirical evidence that specifiers always precede their respective heads, i.e. for SpecPP, SpecVP and Spec $\nu$ P, specifiers either follow a head and/or the data show that the postulated Spec-head order is obscured by subsequent movements rendering the postulated order unprovable. On the other hand, the linear orders for SpecPP, SpecVP, Spec $\nu$ P and SpecCP[TOPIC] all follow from the RPA (3).

## 7. Toward a theory of movement

The framework developed in this and other papers (cf. De Vos 2014a and b, 2013, 2009 a and b, 2008) has implications for the theory of movement and displacement in Minimalist syntactic theory in terms of what moves, what triggers movement and the positions which movement targets. Although a full description is beyond the scope of this paper, an outline is in order.

In standard Minimalist theory, feature checking takes place through AGREE and movement is understood as a combination of AGREE and (internal) merge. Consequently, movement is not necessary in order to check features. Syntactic movement is understood as a costly operation which must be triggered. It may be triggered by AGREE in conjunction with an EPP feature or, as a last resort, escape-hatch movement to exit from a phase, perhaps triggered by a generalized EPP feature. The presence or absence of generalized EPP features may be a language-specific parameter. In addition, movement may occur for semantic reasons (e.g. quantifier raising) and at PF.

In the framework developed in this paper, movement is a function of linearization which, in turn, is a function of the syntactic relations encoded by MERGE and AGREE and constrained by locality. The universal character of these generalizations is tempered by the language-

specific feature specifications of particular lexemes in particular languages: the presence/absence of a feature and/or the morphological configurations available by which to spell out syntactic relations. In what follows, I will explain a number of configurations and the types of movement that follow from them.

### 7.1. Y[uF] . . . X[iF]

In the configuration where an uninterpretable feature probes for its interpretable counterpart within its domain (52a), AGREE establishes a relation (X[iF],Y[uF]), where the interpretable feature on X checks the uninterpretable feature on Y. When this relation is passed to the linearization component at the PF interface, it will be linearized such that X precedes Y, as in (52b).

- (52) a. Underlying configuration: Y[uF] . . . X[iF]  
 b. Surface linearization: X Y . . . t<sub>x</sub> OR  
 c. Surface linearization: X Y . . . X

This is very similar to the standard Minimalist conception of movement. The relation is fundamentally syntactic since it is established by AGREE within narrow syntax. Also, since the uninterpretable feature probes as soon as it is merged at the root of the structure, X will be linearized left-adjacent to the root, Y<sup>30</sup>.

The question arises as to whether such movement is syntactic or a PF phenomenon. More precisely, by syntactic movement, we usually intend that the feature bundle X c-commands Y and the trace of X – does this apply in this configuration? At some level, the movement is syntactic since it depends on AGREE, a syntactic operation. Whether or not X c-commands Y largely depends on how AGREE works. In De Vos (2008), I argued that the output of AGREE (X,Y) was mathematically an ordered pair equivalent to {X,{X,Y}}, which is in itself equivalent to phrase structure. If this argument holds, then AGREE always creates a phrase structure representation where at least the interpretable feature iX c-commands Y: {X[iF],{X[iF],Y[uF]}}. This presumably corresponds to a “feature movement” analysis of agreement. However, we can also consider X to be a set (i.e. a feature bundle) containing iF and a number of other features (Z), and if we consider the feature bundle X to c-command Y as a result of AGREE, then it follows that if we pied pipe those features we obtain the phrase structure {X[iF+Z],{X[iF+Z],Y[uF]}} which would correspond to phrasal movement. Whether phrasal or feature movement is considered in a particular representation will presumably depend on factors intrinsic to the derivation, lexical choice and the morphological resources available (e.g. whether a particular language can spell out a single feature, say, as an expletive.). I see no reason, therefore, why such movement might not be syntactic even though the linearization occurs at PF. Consequently, nothing in the current framework prevents syntactic movement of the standard, feature-triggered type.

<sup>30</sup> This applies to the configuration represented here. However, if there are other feature-checking relationships, these may also impact on the final linearization schema.

## 7.2. Y[iF] . . . X[uF]

Where the current framework differs from the standard model is in its predictions about configurations where an interpretable feature c-commands its uninterpretable counterpart<sup>31</sup>. Let us assume that AGREE operates and creates a pair (Y,X) which implies that the ordering between Y and X remains the same with Y preceding X. There are two possible results, the first illustrated in (53b), where X remains in situ since the RPA (3) is satisfied there; only the locality constraint is flouted. Alternatively, (53c) could also arise, where X moves to a position right-adjacent to Y, leaving a trace and optimally satisfying the locality constraint.

- (53) a. Underlying configuration: Y[iF] . . . X[uF]  
 b. Surface linearization: Y . . . X OR  
 c. Surface linearization: Y X . . . t<sub>x</sub>

Importantly, this type of movement also has representational effects. Since the original relation has not changed, no new phrase structure is created by this application of AGREE (unlike the configuration in the previous section). Consequently, other than word order, one might expect that this type of movement would not yield semantic scope effects, c-command binding, polarity effects, etc. Hence, this would be a type of vacuous movement, commonly associated with PF movement.

A full exploration of this prediction is beyond the scope of this paper, but it is instructive to consider the following scenario. Direct objects have uninterpretable case features which are checked by little *v*. In the absence of  $\phi$  feature checking between the direct object and little *v*, this conforms to the configuration shown above.

- (54) *v*[iCase] . . . D[uCase]

It is therefore interesting that object shift in Scandinavian languages can be optional, with the DP (depending on whether it is a pronoun or a full DP) either remaining in situ or moving to a position right-adjacent to the verb (which has itself undergone V-T movement). It is also interesting that such optional movement has characteristics of PF movement, not yielding scope or binding effects as predicted by the approach proposed here (Holmberg 1999, cited in De Vos 2014a). These types of effects need to be fully explored in further research.

## 7.3. Y[iF,iG] X[uF] . . . Z[uG]

The current framework also has implications for contexts where a category has syntactic relationships with two other categories, for instance, where Y selects X in its specifier and selects Z in its complement, etc.

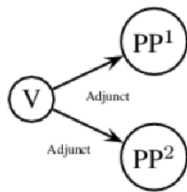
- (55) a. Underlying configuration: Y[iF,iG] X[uF] . . . Z[uG]  
 b. Surface linearization: Y > X > Z OR  
 c. Surface linearization: Y > Z > X

<sup>31</sup> Note that this is the same type of configuration that Bošković (2007) argued always involves overt movement of the probe to a position c-commanding the goal. Interestingly, the current framework as well as that of Bošković both predict that this is a problematic configuration, but they differ on the nature of their predictions.

### 7.3.1. Adjuncts

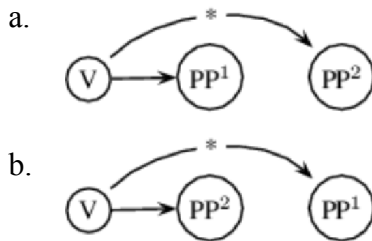
One context where a configuration like (6) may occur is in constructions where two adjuncts are adjoined to a single node. In examples (56a and b), the two PP adverbials of place are plausibly adjoined to the same node, probably at VP or *v*P level<sup>32</sup>. The exact nature of the dependency relationship between the adjuncts and their hosts remains an open question (i.e. which selects which) but for the sake of convenience I will assume that these too are functional dependencies. The relationships within the VP are illustrated graphically in the schema in (56c).

- (56) a. I worked in the field, by the haystack  
 b. I worked by the haystack, in the field  
 c.



There are two different ways to linearize these relations: V PP1 PP2 or V PP2 PP1. In both (57a and b), the relations  $V \rightarrow PP1$  and  $V \rightarrow PP2$  are preserved in linear order and consequently neither violates the RPA (3). However, for both possibilities, one PP will not be immediately adjacent to V. If one treats the locality condition as a violable constraint, then both linearization possibilities are equally optimal. Thus, from the RPA (3) it follows that PP1 is not ordered with respect to PP2 and that, all things being equal, this results in optional word order ((56a) vs (56b)).

- (57)



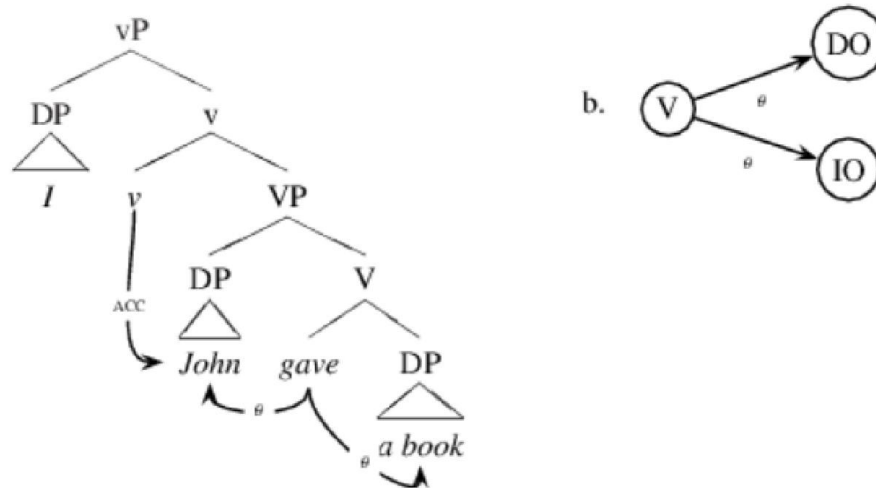
### 7.3.2. Double-object constructions

Another context where it might initially be thought that a configuration like (6) could arise is in double-object constructions, on the basis that in a tree structure like (58a), the verb selects an indirect object (IO) in its specifier and a direct object (DO) as a complement (Larson 1990, 1988)<sup>33</sup>.

<sup>32</sup> Note that if the adverbials are adjoined in different places, then the configuration no longer holds.

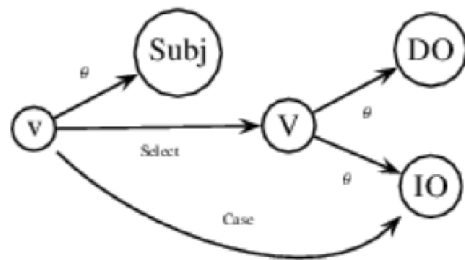
<sup>33</sup> This was pointed out by an anonymous reviewer.

(58)



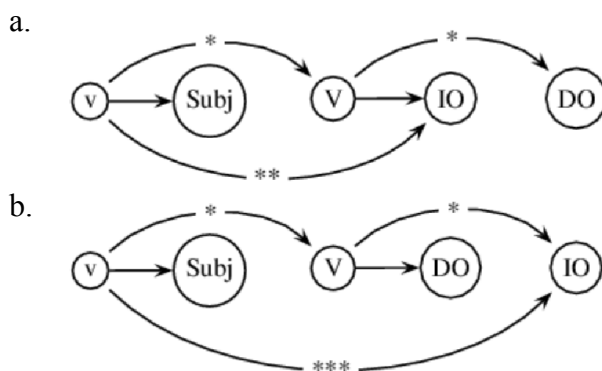
However, when one examines the relations in this structure it is considerably more complex: there is also Case-assignment by little *v* to its closest DP (the indirect object)<sup>34</sup>. Note that the subject would also agree with T and is assigned Case by T, although those relations are omitted for the sake of simplicity. The relations in this structure can be represented schematically as in (59).

(59)



Consequently, the configuration for double-object constructions does not coincide with the structure in (6). Some possible linearization patterns for (59) are represented below.

(60)



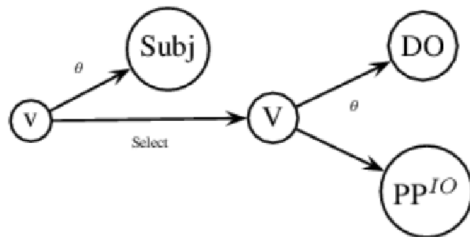
<sup>34</sup> Presumably assignment of Case to the direct object is done by V, although if Case were to be assigned by third V in the VP shell, the outcome of the derivation proposed in (61) would not be affected.

Example (60a) contrasts with (60b) in the length of little *v*'s Case-assignment dependency. The configuration in (60a) is the optimal linearization pattern, thus yielding the asymmetry in (61).

- (61) a. I gave John a book  
 b. \*I gave a book John

This analysis makes the prediction that if Case-assignment from *v* to the IO *John* were to be disrupted, relative word order between the IO and the DO would be optional. This configuration is met when the IO is a PP and where the case of the IO DP is assigned by P, thus conforming to the configuration in (6). The prediction is thus confirmed as both word orders are licit.

(62)

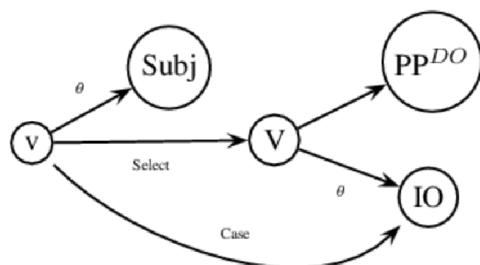


- (63) a. I gave a book to John [DO > IO]  
 b. I gave to John a book [IO > DO]

I acknowledge that (63b) may be improved by appropriate intonation and that constructions of this type may be considered slightly archaic. In addition, issues of heaviness of the constituent undoubtedly play a role in the relative ordering of IO and DO. I have attempted to balance both IO and DO to reduce the effect of heaviness. However, it is important to note that if one compares the relative markedness of (61b) and (63b), it is very clear that the latter is far less marked than the former.

The analysis also predicts that if the direct object is a PP, only IO > DO word order will be allowed since the case of the IO must still be checked by *v*. This situation is evident with predicates such as *spray* (65) which takes an IO DP followed by a DO PP. This contrasts with ditransitives such as *give* (66) where the IO > DO order can only be used if both arguments are DPs (61b) or if the IO is a PP (63b). The prediction is again confirmed.

(64)



In (65a), *paint* is the direct object insofar as it is transferred from a paint-can to a wall and thereby undergoes a change of state (compare with *I gave to John a book* in (66a) where the direct object *a book* is transferred from myself to *John*, and where it undergoes a change of

ownership in the process). *The wall* is the IO insofar as it is the patient and is affected as a result of the spraying event<sup>35</sup>.

- (65) a. Jack sprayed the wall with paint  
           IO                  DO      (Levin 1993:51) [IO > DO]  
 b. \*Jack sprayed with paint the wall  
                           DO IO      (Levin 1993:51) [\*DO > IO]

Thus, one needs to compare the relative statuses of (63b) and (65b), repeated here in (66). The example in (66b) is more marked, as predicted.

- (66) a. I gave to John a book  
 b. \*Jack sprayed with paint the wall

In this section, I have demonstrated that the proposed framework can, in principle, account for non-adjacency effects although much remains for future research. One issue remains at a conceptual (as opposed to a formalized) level, namely how to quantify the “length” of a non-adjacent dependency. A full investigation of this issue is beyond the scope of this paper, but it is interesting to note that PF effects are implicated in exactly the configuration implied by (6), namely heavy-NP effects in English. With heavy-NP effects, the precise characterization of “heaviness” is still unclear in ways that seem to be similar to the notion of ‘length’ used here<sup>36</sup>.

#### 7.4. Concluding remarks on a theory of movement

In this final section, I have attempted to outline how the proposed framework contributes to a larger theory of movement and to explain how it relates to movement within the Minimalist Programme more generally. I have demonstrated that it does not preclude traditional types of syntactic movement, i.e. not all movement is necessarily PF movement even if the fact of overt displacement is an artifact of linearization. I have also demonstrated that the current framework does allow for non-syntactic “movement” under particular constrained feature configurations; this is not so much “movement” as it is displacement as a function of linearization constraints. Regrettably, owing to space considerations, not to mention that it is probably unrealistic to require that any new theory necessarily covers every phenomenon encapsulated by the old, the data come largely from English, although I have attempted to point to areas where research in other languages may be fruitful.

<sup>35</sup> The other side of the spray/load alternation can pattern with the instrumental construction where the PP is presumably an adjunct.

a. Jack sprayed paint on the wall (Levin 1993:51) [IO > DO]  
 b. \*Jack sprayed on the wall paint [\*DO > IO]

An anonymous reviewer notes that these data are specific to English and that languages like Dutch and Afrikaans may have less rigid word order requirements in this regard. An analysis of these constructions in Afrikaans and Dutch is beyond the scope of this paper, but it can be noted that an analysis of the PPs within the West-Germanic middle field might take into consideration the fact that these constituents move to AgrO which precedes vP in these languages. This may imply the existence of an additional agreement relation between AgrO and these constituents.

<sup>36</sup> It is also worth noting that so-called “symmetrical” Bantu languages exist where there is no preferred order between DOs and IOs. It could be the case that the approach suggested here may be able to explain these effects in terms of other demands on structure made at PF. Also note that this approach derives certain word-order effects in the adpositional domain, such as P doubling (De Vos 2009a). Finally, note that the present analysis of double-object constructions has some similarities in spirit with that of Jackendoff (1990) who also proposes an approach to constructions such as heavy-NP shift in terms of optional ordering.



## 8. Conclusion

This paper began with the problem that although topicalization and WH-movement appear to target different positions, there is a variety of evidence to suggest that the positions are identical. This paradox was resolved by reconceptualising linearization as a function of syntactic dependencies; the Relational Precedence Axiom (3) allows a slightly looser, but still constrained, approach to word order. The implications of (3) were then explored for a number of other specifier positions in English, and the patterns between English specifiers and their heads turned out to be revealing: they follow from directional dependencies once these dependencies are mapped to linear order. In addition to being consistent with the attested specifier orders of major categories such as TP,  $\nu$ P, VP, etc., the Relational Precedence Axiom (3) also made subtle predictions about the relative orders of DOs and IOs, accounting for why optional orderings are available in some contexts but not in others. The nature of the explanation suggests that something like the Relational Precedence Axiom (3) could be universal. Whether this is indeed the case, or whether the Relational Precedence Axiom (3) and the Linear Correspondence Axiom are parametric options (given that both are arguably equally “primitive”) remains a question for further study<sup>37</sup>.

## References

- Alexiadou, A., E. Anagnostopoulou and M. Everaert. 2004. Introduction. In A. Anagnostopoulou and M. Everaert (eds.) *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford University Press. pp. 1-21.
- Authier, J.-M. 1991. V-governed expletives, case theory and the projection principle. *Linguistic Inquiry* 22: 721-740.
- Barbiers, S. 2000. The right periphery in SOV languages: English and Dutch. In P. Svenonius (ed.) *The derivation of VO and OV*. Amsterdam: John Benjamins. pp. 181-218.
- Barbiers, S. 2005. Word order variation in three-verb clusters and the division of labour between generative linguistics and sociolinguistics. In L. Cornips and K. Corrigan (eds.) *Syntax and variation: Reconciling the biological and the social*. Amsterdam: John Benjamins. pp. 233-264.
- Borsley, R. 2009. On so-called transitive expletives in Belfast English. *English Language and Linguistics* 13(3): 409-431.

---

<sup>37</sup> Certainly, while these results in no way detract from the general validity of the LCA, they do problematize it. I am not arguing that the LCA is generally incorrect (since it is axiomatic) or otherwise unfruitful, or that specifiers never precede their heads. In fact the Antisymmetry paradigm has raised fundamental questions leading to interesting proposals over the years. I will, however, propose an alternative mapping which can account for the asymmetries in Spec-head orders. And to the extent that it succeeds, it raises questions about some of the ancillary assumptions that are required under the LCA. I also acknowledge that the proposal in this paper cuts to the core of some linguistic “holy cows” and has implications for a range of constructions. Unfortunately, it is simply not possible to cover all of these within a single paper, perhaps not even a book. Just as the LCA, when it was first published, seemed to contradict established facts – e.g. the head-final status of Dutch – it took subsequent research to establish a head-initial syntax for Dutch (Zwart 1994). Hence this paper restricts its aim to “pushing a precise but inadequate formulation to an unacceptable conclusion” (Chomsky 1957:5, cited in Franks 2000) in the interests of opening up debate on these issues.

- Bošković, Ž. 2002. On multiple WH-fronting. *Linguistic Inquiry* 33: 351-383.
- Bošković, Ž. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38: 589-644.
- Brodie, R. 2007. *John Gower's confession amantis: Modern English version*. Available online: <http://www.richardbrodie.com/Book1.html> (Accessed 19 January 2011).
- Broekhuis, H. 2008. *Derivations and evaluations: Object shift in the Germanic languages*. Berlin: Mouton de Gruyter.
- Browning, M. 1996. CP recursion and that-t effects. *Linguistic Inquiry* 27: 237-255.
- Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. 1977. On wh-movement. In P. Culicover, T. Wasow and A. Akmajian (eds.) *Formal syntax*. New York: Academic Press. pp. 71-132.
- Chomsky, N. 1995a. Bare phrase structure. In G. Webelhuth (ed.) *Government and binding theory and the Minimalist program*. Cambridge, MA: Basil Blackwell. pp. 383-439.
- Chomsky, N. 1995b. *The Minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by step: Essays in Minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press. pp. 89-155.
- Cinque, G. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.
- Culicover, P. 1996. On distinguishing A'-movements. *Linguistic Inquiry* 27: 445-463.
- De Vos, M. 2005. *The Syntax of Verbal Pseudo-Coordination in English and Afrikaans*. PhD thesis, Leiden University Centre for Linguistics.
- De Vos, M. 2008. Deriving Narrow Syntax from principles of lexical organization. *Lingua* 118: 1864-1899.
- De Vos, M. 2009a. *Afrikaans mixed adposition orders as a PF-linearization effect: Disharmony is a superficial side effect of deeper harmony*. Paper delivered at the Theoretical Approaches to Disharmonic Word Orders Conference, 30 May–1 June 2009, Newcastle University, Newcastle.
- De Vos, M. 2009b. *Expletives on the interface: Linearization and PF output conditions*. Available online: <http://ling.auf.net/lingbuzz/000971> (Accessed 20 November 2014).
- De Vos, M. 2012. Heads and specifiers: On Spec-head word orders in English. *Groninger Arbeiten zur germanistischen Linguistik* 53(1): 53-78.

De Vos, M. 2013. *Head movement without movement: Head movement is an artifact of optimal solutions to linearization paradoxes*. Presentation delivered at the South African Microlinguistics Workshop, 30 October–1 November 2013, Potchefstroom.

De Vos, M. 2014a. *Deriving Holmberg's generalization without head movement – A linearization driven approach*. Presentation delivered at the annual meeting of the LSSA/SAALA/SAALT, 24–27 June 2014, University of the Witwatersrand, Johannesburg.

De Vos, M. 2014b. The evolutionary origins of syntax: Optimization of the mental lexicon yields syntax for free. *Lingua* 150: 25-44.

Delfitto, D. 2002. On the semantics of pronominal clitics and some of its consequences. *Catalan Journal of Linguistics* 1: 41-69.

Den Besten, H. 1989. *Studies in West Germanic Syntax*. PhD thesis, Katholieke Universiteit Brabant, Tilburg, Amsterdam.

Den Dikken, M. 2008. *On the functional structure of locative and directional PPs*. Available online: <http://web.gc.cuny.edu/dept/lingu/dendikken/papers.html> (Accessed 14 January 2008).

Emonds, J. 1978. The verbal complex V' V in French. *Linguistic Inquiry* 9: 151-175.

Ernst, T. 2002. *The syntax of adjuncts*. *Cambridge studies in Linguistics* (No. 96). Cambridge, UK: Cambridge University Press.

Fortuny, J. 2008. *The emergence of structure in syntax*. Amsterdam: John Benjamins.

Franks, S. 2000. A PF-insertion analysis of *that*. *Syntaxis* 3: 1-27.

Giorgi, P. and P. Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. New York: Oxford University Press.

Haegeman, L. 1991. *Introduction to Government and Binding Theory*. Oxford, UK and Cambridge, USA: Blackwell.

Hagstrom, P. 1998. *Decomposing Questions*. PhD thesis, Massachusetts Institute of Technology.

Halmos, R. 1960. *Naive set theory*. Princeton: Van Nostrand.

Hauser, M. D., N. Chomsky and T. Fitch. 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298: 1569-1579.

Henry, A. and S. Cottell. 2007. A new approach to transitive expletives: Evidence from Belfast English. *English Language and Linguistics* 11(2): 279-299.

Holmberg, A. 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53: 1-39.

Jackendoff, R. 1990. On Larson's treatment of the double object construction. *Linguistic Inquiry* 21: 427-456.

- Jonas, D. and J. Bobaljik. 1993. Specs for subjects. In J. Bobaljik and C. Phillips (eds.) *Papers on case and agreement: The role of TP in Icelandic*. Cambridge, MA: MIT Press. pp. 59-98.
- Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kayne, R. 2001. *Prepositions as probes*. Oxford: Oxford University Press.
- Kayne, R. 2005. *Movement and silence*. Oxford: Oxford University Press.
- Ko, H. 2006. On the structural height of reason wh-adverbials: Acquisition and consequences. In L. Cheng and N. Corver (eds.) *WH-movement: Moving on*. Cambridge, MA: MIT Press. pp. 319-350.
- Koopman, H. 2000. Prepositions, postpositions, circumpositions, and particles. In H. Koopman (ed.) *The syntax of specifiers and heads*. London: Routledge. pp. 204-260.
- Kracht, M. 2003. *The mathematics of language*. Berlin: Mouton de Gruyter.
- Langendoen, D. 2003. Merge. In A. Carnie, H. Harley and M. Willie (eds.) *Formal approaches to function in grammar: In honor of Eloise Jelinek*. Amsterdam: John Benjamins. pp. 307-318.
- Larson, R.K. 1988. On the double object construction. *Linguistic Inquiry* 19: 335-391.
- Larson, R.K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21: 589-632.
- Lasnik, H., J. Uriagereka and C. Boeckx. 2005. *A course in Minimalist syntax: Foundations and prospects*. Malden, MA: Blackwell Publishing.
- Levin, B. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- MacCaull, W. 2000. A proof system for dependencies for information relations. *Fundamenta Informaticae* 42: 1-27.
- Maki, H., L. Kaiser and M. Ochi. 1999. Embedded topicalization in English and Japanese. *Lingua* 107: 1-14.
- Pesetsky, D. 1987. Wh-in-situ: Movement and unselective binding. In E. Reuland and A. ter Meulen (eds.) *The representation of (in)definiteness*. Cambridge, MA: MIT Press. pp. 98-129.
- Pesetsky, D. and E. Torrego. 2001. T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.) *Ken Hale: A life in language*. Cambridge, MA: MIT Press. pp. 355-426.
- Reichenbach, H. 1947. *Elements of symbolic logic*. New York/London: The Macmillan Company.
- Rivero, M.-L. 1978. Topicalization and wh movement in Spanish. *Linguistic Inquiry* 9: 513-517.
- Rizzi, L. 1991. *Relativized minimality*. Cambridge, MA: MIT Press.

- Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.) *The elements of grammar: Handbook in Generative syntax*. Dordrecht: Kluwer Academic Publishers. pp. 281-337.
- Roberts, I. 1997. *Comparative syntax*. New York: Edward Arnold.
- Uriagereka, J. 1999. Multiple spell-out. In S. Epstein and N. Hornstein (eds.) *Working Minimalism*. Cambridge, MA: MIT Press. pp. 251-282.
- Van Craenenbroeck, J. and M. Van Koppen. 2002. *Merge versus Long-Distance Agree: The case of complementizer agreement*. Talk presented at GLOW, 9–11 April 2002, University of Amsterdam and University of Utrecht.
- Van Koppen, M., N. Corver and H. Kranendonk. 2009. *Quantitative er: A new micro-comparative view on an old puzzle*. Presentation at the 24<sup>th</sup> Comparative Germanic Syntax Workshop 24 at the Center for Research in Syntax, Semantics and Phonology, 28–29 May 2009, Hogeschool Universiteit Brussel.
- Zwart, J.-W. 1994. Dutch is head-initial. *The Linguistic Review* 11(3): 377-406.
- Zwart, J.-W. 2011. Structure and order: Asymmetric merge. In C. Boeckx (ed.) *The Oxford handbook of linguistic minimalism*. Oxford: Oxford University Press. pp. 96-118.