On Split-CPs, Uninterpretable Features, and the 'Perfectness' of Language

Frederick J. Newmeyer
University of Washington
fjn@u.washington.edu

Abstract

This paper discusses critically a number of developments at the heart of current syntactic theory. These include the postulation of a rich sequence of projections at the left periphery of the sentence; the idea that movement is tied to the need to eliminate uninterpretable features; and the conception put forward by Chomsky and others that advances in the past decade have made it reasonable to raise the question about whether language might be in some sense 'perfect'. However, I will argue that there is little motivation for a highly-articulated left-periphery, that there is no connection between movement and uninterpretable features, and that there is no support for the idea that language might be perfect.

1   Introduction

The scope of this paper is quite wide-ranging, but hopefully not excessively so. Its goal is to discuss and relate a number of proposals whose relation might not be immediately obvious. These include the postulation of a sequence of projections at the left periphery of the sentence, each encoding some aspect of the sentence's information structure; the linking movement to the elimination of uninterpretable features; and the idea that advances in the past decade have made it reasonable to raise the question about whether language might be in some sense 'perfect'. Here is a typical passage from Chomsky to that effect:

> [Is] it possible that the system of language itself has a kind of optimal design, so, is language perfect? Back in the early 1980s that was the way I started every course — 'Let's ask: could language be perfect?' — and then I went on the rest of the semester trying to address the question, but it never worked, the system always became very complicated. What happened in the early 1990s is that somehow it began to work; enough was understood, something had happened, it was possible to ask the question in the first session of a course: could language be perfect? and then get some results which indicated it doesn't sound as crazy as you might think. (Chomsky 2002: 96-97)

For each trend, I am going to take what one might call a 'cautious', perhaps even a 'conservative', position. I will conclude that there are no strong arguments for a highly-articulated left periphery, that there is no connection

whatsoever between movement and uninterpretable features, and that the idea that language might be perfect is at best incoherent.

The paper is organized as follows. Section 2 outlines the split-CP hypothesis and section 3 shows how the Nested Dependency Constraint is able to account for much of the data that has been appealed to in order to motivate the hypothesis. Section 4 argues against the covert movement of focus and topic and section 5 against attracting focus projections. Section 6, by pointing to the disparity between information structure and structural positions, reinforces the conclusions of the previous two sections. The claim that language might be 'perfect' is put under the microscope in section 7, followed by a brief conclusion in section 8.

2   The Split-CP Hypothesis

By the mid 1980s, a near consensus had arisen on one aspect of clause structure. In the then-dominant view, CP and IP occupied root position and were regarded as full participants in X-bar system, as in (1):

(1) $_{CP}$[SPEC,CP $_{C'}$[C $_{IP}$[SPEC,IP $_{I'}$[I VP …

The dominant trend within Principles-and-Parameters syntax today is to split CP into an indefinite number of separate projections, as in (2):

(2) $_{\alpha P}$[SPEC, αP $_{\alpha'}$[α $_{\beta P}$[SPEC, βP $_{\beta'}$[β $_{\gamma P}$[SPEC, γP $_{\gamma'}$[γ … IP …

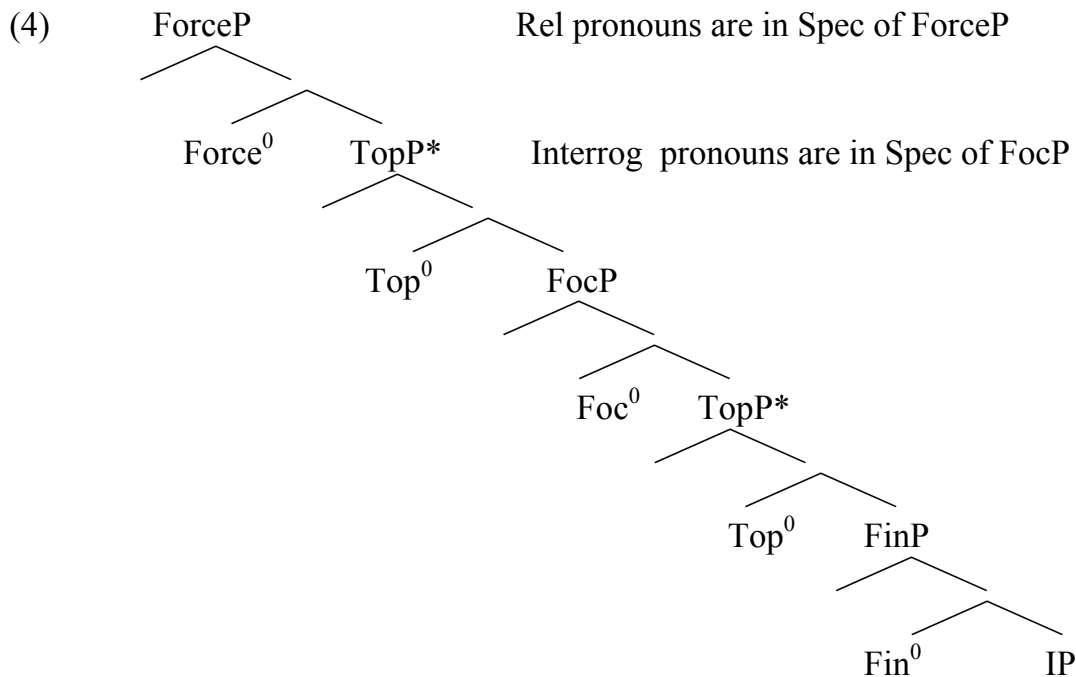Three widely accepted theoretical proposals of the 1990s made the split-CP hypothesis almost inevitable:

(3)   a. There is no optionality in grammar; hence elements move only when they are 'required to' (Chomsky 1995).

   b. Movement must be triggered by a feature on a functional head (Chomsky 2000).

   c. Features of the 'peripheral system' (force, topic, focus, etc.) trigger A' movement (Chomsky 2000).

(3a), in effect, requires that seemingly optional variants have different underlying structures. Since few if any structural variants have the same information structure properties, it seemed reasonable to locate their structural differences in projections representing properties of information structure. Such an idea was reinforced by the adoption of (3b) and (3c).

Luigi Rizzi has been the prime mover in implementing the idea that the elements above IP (the 'left periphery') encode semantic and pragmatic properties of the sentence:

> Syntactic movement … must be triggered by the satisfaction of certain quasi-morphological requirements of heads. … [S]uch features have an interpretive import (Wh, Neg, Top, Foc, …): they determine the interpretation of the category bearing them and of its immediate constituents …, function as scope markers for phrases with the relevant quantificational force in a local configuration, etc.… (Rizzi 1997: 282)

The tree in (4) illustrates his concrete proposal:

(4)

```
        ForceP                        Rel pronouns are in Spec of ForceP
        /    \
   Force⁰    TopP*                Interrog  pronouns are in Spec of FocP
            /    \
        Top⁰     FocP
                /    \
             Foc⁰    TopP*
                    /    \
                 Top⁰    FinP
                        /    \
                     Fin⁰     IP
```

(Rizzi 1997; see also Hatakeyama 1997 for a similar proposal)

In this view, then, CP is split into a number of semantically-relevant projections. The highest projection, Force Phrase, encodes the illocutionary force of the sentence (whether it is declarative, interrogative, exclamatory, and so on). Below are Topic Phrases and a Focus Phrase. The lowest CP-type projection is 'FinP', which specifies whether the IP below it is finite or non-finite. Notice that Topic Phrases can iterate both above and below the Focus Phrase, at least in Italian:

(5)   Credo che a Gianni, QUESTO, domani, glidovremmo dire
        C   Top   Foc   Top   IP
      'I believe that to Gianni, THIS, tomorrow we should say'

In Rizzi's account, information structure is directly encoded in the LF representation, as in (6a, b). What follows the Topic is claimed to be the Comment; what follows the Focus the Presupposition:

(6)   a. Top^0 [Comment]

  b. Foc^0 [Presupposition]

Since 1997, the 'cartography of the C domain' (as it is often referred to) has become a topic of much discussion. It seems fair to say that most work in the general envelope of the MP has accepted the idea of multiple projections on the left-periphery. Indeed, the trend seems to be to propose *more* projections than those posited by Rizzi. The next few sections argue against the idea of splitting CP into a multitude of distinct projections.
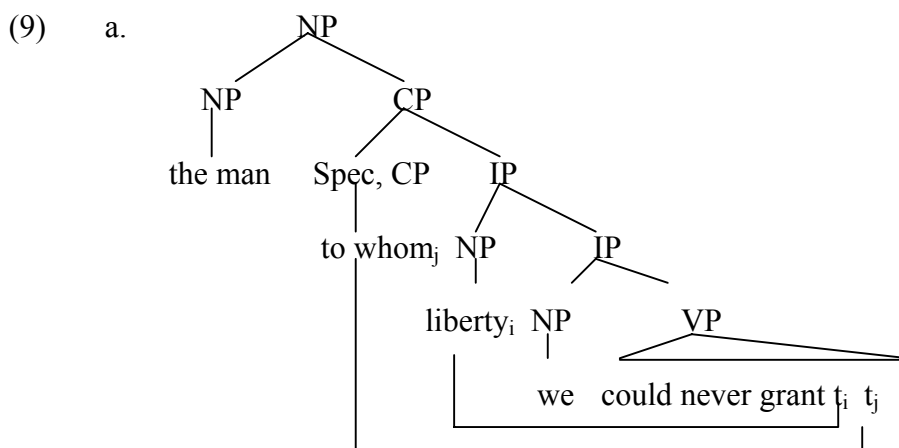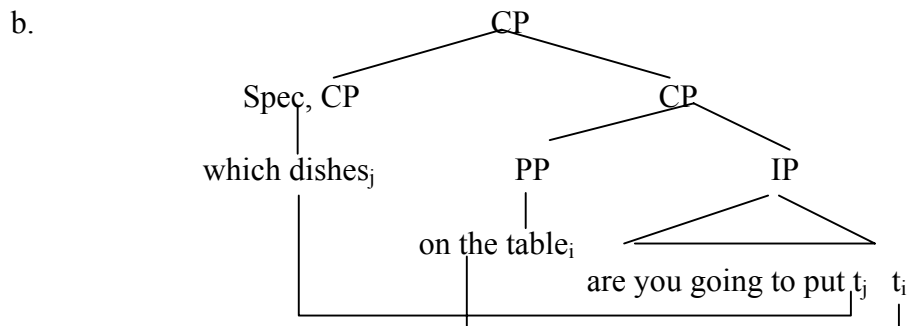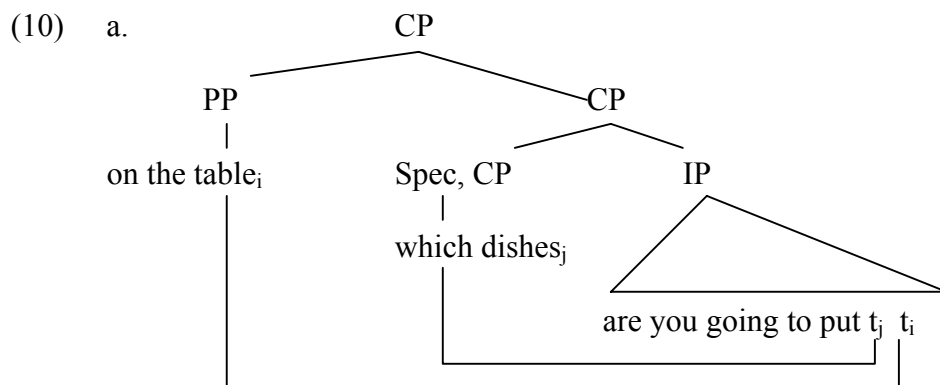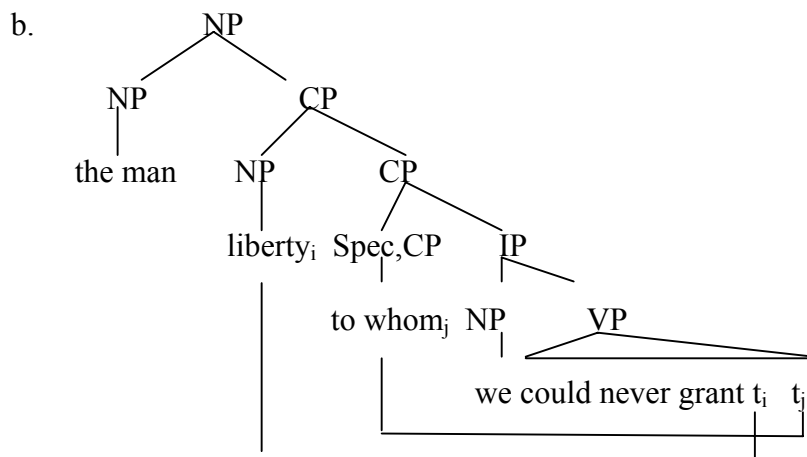
3   Ordering on the Left Periphery and the NDC

The primary support for the articulated structure of (4) comes from what appear to be strict ordering restrictions on the left periphery. For example, Rizzi points out that preposed topics follow *wh*-relative pronouns, but precede *wh*-interrogative pronouns, as in (7) and (8):

(7)   a.   the man to whom liberty, we could never grant (Baltin 1982)

  b. *the man, liberty, to whom we could never grant

(8)   a.   On the table, which dishes are you going to put?

  b. *Which dishes are, on the table, you going to put? (Culicover 1991a)

Such facts would seem to support the idea of a *wh*-relative projection above a topic projection, in turn above a *wh*-interrogative projection.

   These facts, however, follow from an independently-needed principle. Consider the structures of (7a, b) and (8a, b), namely (9a, b) and (10a, b), respectively:

(9)   a.

b.

```
                        NP
                  NP         CP
              the man    NP        CP
                       liberty_i  Spec,CP    IP
                             to whom_j  NP      VP
                                        we could never grant t_i  t_j
```

(10)   a.

```
                        CP
                  PP           CP
              on the table_i  Spec, CP       IP
                            which dishes_j
                                        are you going to put t_j  t_i
```

b.

```
                              CP
                  Spec, CP            CP
              which dishes_j      PP        IP
                               on the table_i
                                        are you going to put t_j  t_i
```

As the trees clearly indicate, the grammatical expressions comply with the Nested Dependency Constraint (11), while the deviant ones violate it.

(11) NESTED DEPENDENCY CONSTRAINT (Fodor 1978, 1984; Pesetsky 1987): Multiple filler-gap dependencies may be disjoint or nested, but not intersecting.

Hence there is no need to posit separate projections for interrogative and relative pronouns. The bulk of the ordering restrictions among overt elements on the left periphery fall out in an analogous way.

I should point out that there is a wide range of acceptability regarding sentences like (8a). (8a) itself I find to be of only of marginal acceptability. Other sentences like it are clearly bad, for example (12a, b) and others much better, for example (12c, d):

(12) a. *The book to whom did you give?   (Baltin 1982)

b. *Robin who will talk to?   (Culicover 1991a)

c. ?On that subject, who should I consult with
                                        (Haegeman & Guéron 1999)

d. During the holidays, which book will you read?
                                        (Haegeman & Guéron 1999)

I will return to these sentences in §6.

4   Against Covert Movement of Focus and Topic

The grammatical element acting as topic and focus in English and other languages is very frequently not on the left-periphery on the surface. Hence, given a Rizzian approach, non-overtly fronted topics and focuses raise in LF to the Spec of their relevant projections, attracted by the features of the heads of these projections. The idea of LF focus movement goes back to work by Chomsky in the 1970s (Chomsky 1976, 1981). There is no difficulty, of course, in positing the LF movement of an element from some clause-internal position X to some left-peripheral position Y. The only question is how motivated the movement is. I will argue now that the motivation for LF movement of focus and topic is very weak.

Rizzi gives a number of arguments for LF focus movement, the most important of which is the fact that focus, unlike topic, sets up an operator-variable relation. For example, focuses, but not topics, give rise to weak cross-over effects (Culicover 1991b; Lasnik & Stowell 1991; Rizzi 1997):

(13) a. $John_i$, his_i mother loves t_i (*John* a topic)

b. *$JOHN_i$ his_i mother loves t_i (not Paul) (*John* a focus)

Even in situ focus can generate weak cross-over effects, as pointed out by Lappin (1982) and Lasnik & Stowell (1991):

(14) ??His_i mother shot $JOHN_i$

At first thought, the quantificational nature of focus would seem ideal for an LF movement analysis. But the problem with a movement approach to focus is

that basically *anything* can be in focus. That means that focus movement obeys or violates island constraints willy-nilly. So as noted by Reinhart (1991), the stressed NPs in (15) are contained in strong (ungoverned) islands:

(15) a. [IP [CP That Linda argued with THE CHAIRMAN] is surprising].

     b. [IP [NP Even the paper that LUCIE submitted to our journal] was weak].

Extraction of the focused elements should therefore be impossible. Horvath (1999) made a similar point with the examples in (16) and (17) and Jackendoff (1997) with the ones in (18) and (19):

(16) Q. Do people wonder where Mary was last night?

     A. No, people wonder where [Mary's BOYFRIEND] was last night.

(17) Q. Have you shown Bill the book that I gave you for your birthday?

     A. No, I have (only) shown him the book that you gave me for CHRISTMAS.

(18) a. Is John certain to WIN the election?, which could not have LF:

     b. *[win$_i$ [is John certain to t$_i$ the election]]

(19) a. Does Bill eat PORK and shellfish?
        which could not have LF:

     b. *[pork$_i$ [does Bill eat t$_i$ and shellfish]]

Notice that pied-piping is not an option for deriving the sentences of (16)-(17), given the reasonable prohibition against LF pied-piping. And even if covert pied-piping were permitted, it should not apply in sentences (16) and (17), since if a *wh*-word is substituted for the focus, pied-piping is impossible:

(20) a. *Whose boyfriend do people wonder where was last night?

     b. *For what (holiday) did you show him the book that you gave me?

But what then about the weak-crossover effects observed with in situ focus? Erteschik-Shir (1997) has explained how the cases we have been looking at can be explained without focus movement. In sentences like (14), coreference licensing is violated since stress on the antecedent blocks its interpretation as a topic.

The idea of LF movement to Spec,FocP would be supported if the configuration resulting from this movement were sufficient to guarantee an operator-variable relation between filler and gap. It would indeed be impressive if the hypothesized LF-movement correlated with facts about quantification. But Rizzi postulates parallel movements where there *is no* operator-variable relation. That is, he has LF-topicalization as well. As Rizzi himself points out and discusses in detail, topicalization is not quantificational

(as the contrast between (13a) and (13b) illustrates). Unfortunately, the non-quantificational nature of topics does not follow from anything in Rizzi's analysis per se. He is forced to stipulate that a focus-moved element binds a variable and that a topic-moved element does not. In other words, these fundamental semantic differences between topics and focuses give no support to the split-CP hypothesis.

It is also worth pointing out with respect to topicalization that topics do not obey Rizzi's own Relativized Minimality principle. That is, they do not interfere with each other. Rizzi's paper is full of examples of multiple topicalization where one topic has clearly moved over another. Such seems to provide another reason to abandon topic projections.

A related problem has been pointed to by (Shaer 2003a,b). Shaer argues that the fronted adverbials in (21) bear no syntactic relation to the sentence they are associated with. In the terminology of Haegeman (1991), they are 'orphans':

(21) a. *With perfect technique*, John executed the triple somersault.

b. *As passionately as he could*, Kim kissed Sandy.

c. *With nothing but a crowbar and a ballpeen hammer*, I very much doubt that Terry will be able to repair the Vax in our office.

To summarize Shaer's arguments, the sentence-initial adverbs in (21) are set off intonationally from the rest of the sentence; do not give rise to Principle C violations when they contain R-expressions coindexed with a subject NP; show no evidence of being subject to movement constraints; never trigger subject-auxiliary inversion; resemble parentheticals in their behavior with respect to VP ellipsis; can be interpreted as applying to both conjuncts when they occur with conjoined sentences; and are unable to license negative polarity items, where the licensor needs to be in a c-command relation to the NPI. The problem for the Split-CP hypothesis is that these initial adverbials have the same sort of information structure properties that other fronted elements have. One would assume that if CP were split, then they would occupy the Spec of TopP or FocP, as the case may be. The conclusion that *some* fronted elements do not participate in an articulated CP cartography suggests that such an approach *in general* is on the wrong track.

Complicating matters still further is the fact that Rizzi implies that not all topics and focuses that are in situ on the surface should be fronted in LF. For example, he makes a statement that I find quite cryptic:

> [I]t is reasonable to assume that the topic-focus system is present in a structure only 'if needed', i.e. when a constituent bears topic or focus features to be specified by a Spec-head criterion'. (1997: 288)

But what determines which such features need to be specified and which do not? Rizzi has given a possible answer in class lectures. He postulates what he calls 'Subj-Pred Articulation' and suggests that in situ subject-topics are interpreted as in (22c), parallel to (22a) and (22b):

(22) a. [XP [Top Comment]]

     b. [XP [Foc Presupposition]]

     c. [XP [Subj Predicate]]

I have no idea what the discourse interpretation of 'Subj Predicate' might be that would distinguish it from Top Comment. In fact, I doubt very strongly that 'subject' and 'predicate' are information-structure concepts at all. Rizzi does suggest a criterion to distinguish elements in a (22a) configuration from those in a (22c) configuration. He suggests that the fronted topics of (22a) need to be D-linked, while the subjects of (22c) need not be. But in fact, not all fronted topics are D-linked. So, I can observe some workers moving heavy furniture into an office and felicitously remark to my friend (23):

(23) A filing cabinet that heavy you would never get me to lift.

A filing cabinet can be a discourse topic without being D–linked. In fact, sentence (23) could be uttered without any discourse antecedence at all.

    Feature-driven topic and focus movement leads to other curious problems. It has been a staple of discourse studies since the early work of the Prague School that non-contrastive focus in English is a function of the information contour and is therefore final in its domain. Indeed, Chomsky endorsed that idea over 30 years ago:

> … an apparent alternative would be to determine focus and presupposition in terms of surface structure: the focus is the phrase containing the intonation center, and the presupposition is determined by the replacement of the focus by a variable … (Chomsky 1971: 200).

So consider Chomsky's sentence (24):

(24) Was he warned to look out for [an ex-convict with a red shirt]?

<div align="center">

SHIRT

RED SHIRT

A RED SHIRT

WITH A RED SHIRT

EX-CONVICT WITH A RED SHIRT

AN EX-CONVICT WITH A RED SHIRT

</div>

All of the capitalized constituents are possible focuses. Nevertheless, they are not all are phrasal constituents. For example, the noun *shirt*, head of the NP *a red shirt*, is not phrasal. Unless one were to take the position that all constituents on right branches are maximal projections, the movement of *shirt* to SPEC, FocP violates constraints on movement going back to Emonds' Structure Preserving Constraint (Emonds 1976).

Focus movement becomes even more problematic, given that Rizzi and others assume that *contrastive* focus fits into the schema. But not only do contrastive focuses not have to be maximal projections, they do not have to be on right branches either. Examples from Ladd 1978 and Lambrecht 1994 illustrate:

(25) Q. Has John read *Slaughterhouse-Five*?

A. No, John doesn't READ books.

(26) Mary is THE boss.

Also, note contrastive focus can be formally discontinuous (Jacobs 1984, 1988,1992; Krifka 1991; Rooth 1985; von Stechow 1989):

(27) Q: Did Mary wash the car?

A: No, TOM washed the WINDOWS.

Chomsky (1971) has even given an example of where the focused element is a prefix:

(28) John is more concerned with AFfirmation than with CONfirmation.

It is not clear how a prefix could occupy [Spec, FocP].

In other words, the idea of moving focus to the left periphery creates more problems than it solves.

5   Against Attracting Focus Projections

Let me turn to another problem with the idea that A′-movement is driven by features housed in functional projections with semantic or pragmatic relevance. For some such movements, there is no plausible semantically appropriate functional projection to act as a landing site. So consider Heavy-NP Shift:

(29) I gave [to Mary] [all my books on the phonetic foundations of conventional implicature].

The classical analysis of this construction involves adjunction to the right. However, there have been several attempts to derive heavy NPs through leftward movement (Kayne 1994; Larson 1988, 1990; Rochemont 1998;

Rochemont & Culicover 2000). For Rochemont, the movement involves attraction to a feature of a head. The question is the nature of the attracting projection. The usual assumption that it is a Focus Phrase and so movement is to Spec of a FocP. The problem is that heavy NPs are not necessarily in focus. As Arnold, et al. (2000) have shown, postposing is a complex function of information structure and processing requirements. These requirements sometimes work together, but not always. So, for example, in (30b), the phrase *to Mary* is the focus, not the postposed heavy NP:

(30) a. To whom did you give all your books on the phonetic foundations of conventional implicature?

b. I gave [TO MARY][all of those unbelievably ridiculous publications].

But if Heavy-NP Shift is movement to the left to check off a feature, and it is not necessarily a focus feature, then what semantic or information-structure motivation could the projection have? And if none, then what is the empirical motivation for the leftward movement in the first place?

Another argument that has been made for the direct syntactic encoding of a focus projection is that in some languages the movement appears to be instantiated overtly to some designated structural focus position. Hungarian is the best-studied instance. So — the reasoning goes — one should find parallel covert movement in other languages. And if such movement exists, then why not regard it too as attraction by a focus feature?

The problem is that in a lot of languages the overt movement of focus would have to be supplemented by a subsequent covert movement to get the scope facts right. Take the Bantu language Aghem, as discussed by Watters (1979) and Horvath (1995):

(31) a. fíl á mɔ̀ zɨ́ GHɛ́ bɛ́-ˈkɔ́
    friends SM P₂ eat WHERE fufu
    'Where did the friends eat fufu?'

b. fíl á mɔ̀ zɨ́ ÁN ˈSÓM bɛ́-ˈkɔ
    friends SM P₂ eat IN FARM fufu
    'The friends ate fufu IN THE FARM.'

Aghem is an SVO language with a syntactic focus position immediately following V — the position in which *wh*-elements and other focused phrases occur. So one would posit a FocP in that position triggering movement into that position. But as Horvath points out, the focused phrase has scope over the entire sentence. So assuming that scope is determined on the basis of c-command relations, one would still need a covert focus movement in Aghem.

In other words, overt scope-position languages provide no support for an analogous covert movement in languages like English. Some overt movement languages would have to have covert movement as well.

The same point can be made for Chinese (Niina Zhang, personal communication). In Chinese, referential DO's can move to a position between the subject and the verb, creating SOV order:

(32) Háizi   bǎ   shū   mǎi   le
     child   OBJ   book   buy   ASP
     'What the child did to the book was buy it.'

But the object in this position is either a focus (if it is stressed) or a topic. So again, covert movement would be needed to get the scope facts right.

Wh-movement manifests the same problem. In Malayalam and a number of other OV languages, question words move immediately to left of V, but have scope over the entire sentence:

(33) a.   niṅ-ṉe   aarə   talli   (Jayaseelan 2002)
       you-ACC   who   beat (PAST)
       'Who beat you?'

     b.   *aarə   niṅ-ṉe   talli
       who   you-ACC   beat(PAST)

So, again, to get the scope facts, one would need to posit a second covert movement. The fact that both overt and covert *Wh*-movement have apply in the same language makes the idea that *Wh*-movement is attraction to [Spec,FocusP] lose a lot of its appeal.

6   The Disparity between Information Structure and Structural Positions

Confounding the split-CP hypothesis still further is that fact that preposed *wh*-elements themselves have information structure properties. For example, in the unmarked case, *wh*-relative pronouns are topics and *wh*-interrogative pronouns are focuses. These facts create still other conflicts. Let us reexamine tree (4). Rizzi posits relative pronouns in Spec, ForceP and interrogative pronouns in Spec of FocP. This hypothesis gets the ordering facts right, as well as the generalization that *wh*-interrogative pronouns are focuses. But it gets other facts wrong. If relative pronouns occupy [Spec, ForceP], then we lose the generalization that they are topics and it mistakenly implies that they type the illocutionary force of the S. Likewise, if interrogative pronouns are in [Spec, FocP], as suggested by Rizzi, then one loses the generalization that it is the nature of ForceP that types a sentence as declarative, interrogative, and so

on. It is the *wh*-phrase in questions that conveys the illocutionary force of questioning. Thus, the interrogative head feature of ForceP would seem to need an interrogative phrase in its Specifier to meet the *Wh*-Criterion, (34):

(34) *WH-CRITERION* (Rizzi 1996: 64, revising May 1985):

    A. A *wh*-operator must be in a Spec-head configuration with X^0 [+wh].

    B. An X^0 [+wh] must be in a Spec-head configuration with a *wh*-operator.

In other words, to get the syntactic facts, the *wh*-operator has to be in a low projection. But to get the illocutionary force facts, the *wh*-operator has to be in a high projection. The obvious way to avoid this contradiction is to *unsplit* CP, that is, to posit one CP with one head, which can be + or - *wh*.

Sobin (2002) has called attention to an analogous problem. Rizzi posits that the complementizer *that* is in Force, because it types a sentence as declarative and occurs before embedded topics:

(35) I think that, tomorrow, John will leave.

But in languages like Middle English, where fronted *wh*-phrases and the *that* complementizer cooccur, the *wh*-phrase *precedes* the complementizer:

(36) What that I mene, O swete herte deere?
    'What do I mean, oh dear sweetheart?'

Other considerations, however, lead Rizzi to put fronted *wh*-phrases in Spec, FocP, a position below (i.e. following) the Force Phrase and therefore after the *that*-complementizer. In other words, given the split-CP, we have to posit two incompatible landing sites for *wh*-movement. Again, the classical analysis, where the *wh*-phrase is in Spec, CP and the *that* complementizer is C, gets the facts right. In any event, the idea that the illocutionary force of a sentence might be representable as something as simple as a feature of a syntactic projection flies in the face of research on the syntax, semantics, and pragmatics of speech acts going back to Searle's work in the early 1970s.

Just to make a historical point, one of the major advances in syntactic theory was the proposal, first articulated in Bresnan 1970, that what we would later call the specifier of CP was the unique landing site for *Wh*-Movement. Bresnan's hypothesis paved the way for the hypothesis of a single movement rule Move-α and the constraints to which it was subject. Before 1970, as in Ross 1967, relative clause movement and *wh*-question movement were regarded as separate rules. The split-CP hypothesis takes us back to the construction-based theory of early transformational grammar and away from the constraint-based theory developed in the 1970s and 1980s.

Cascades of higher functional projections might not seem unreasonable for languages like Italian and Catalan, where each semantically-relevant

projection seems to have its place. But cross-linguistically it is hard to find good support for the kind of architecture represented in trees like (4). So, some languages allow a single position on the left periphery, which can house either a focus or a topic, but not both (Lopez 2002; Vilkuna 1995). In other languages, like Zapotec, displaced *wh*-elements and displaced focus-elements can cooccur, despite that fact that in Rizzi's model they are supposed to target the same projection (Lee 2001). My point is not to call attention to the existence of language variation. Of course, there is going to be variation. The point is that oceans of functional projections on the left-periphery represent a singularly *unminimalist* framework for capturing this variation. Given that few languages manifest the proposed cartography in its fullness — and that languages differ markedly in the relative hierarchy of those projections that do occur overtly — there is no benefit to proposing that UG provides the set of projections and their ordering. Such a proposal forces recourse to both a fixed architecture and language-particular statements to fit the facts into the mold.

Surely there are more minimalist analyses possible than the CP-hypothesis in all its full articulated glory. For example, one might be to posit, along with Breul (2003), a single target for both fronted interrogatives and other fronted XPs, thereby accounting for the marginal acceptability or outright ungrammaticality of many sentences like (8a, b) and (12a-d). But the fact that *some* sentences of this form are possible suggests that the best option might be to posit however many adjunction rules are necessary for the particular language along with a characterization of the information-structure properties of the results of the movement. The desirability of an adjunction analysis over a TopP analysis for topicalization was argued for persuasively by Lasnik & Saito (1992). To the best of my knowledge, their arguments have never been successfully rebutted.

For a time, Chomsky seemed to be leaning back to the idea of topicalization by adjunction. In 'Derivation by Phase' (Chomsky 2001), he suggested that Topicalization might be a free option, conditioned by interpretive principles and he opened the possibility that a wide variety of syntactic operations that were once deemed triggered by features of functional categories might apply in the phonology. We will see how the MP develops in this regard.

Even within the general envelope of the MP, there are approaches where triggered movement is not posited. For example, in Erteschik-Shir 1997, topic and focus constituents are marked at 'F-structure', an annotated S-structure. The rule of predication takes F-structures as input and produces well-formed discourse representations. And in Fox 1995 and Reinhart 1997 (the latter based on Cinque's (1993) theory of phrasal stress), focus is not encoded in the syntactic derivation, but arises exclusively at the PF interface. 'Interface economy' explains why elements not in canonical position can be stressed focuses. (See also a parallel optimality-theoretic account in Szendrői 2001.)

I believe it to be the case that most generative syntacticians who work on information structure have excluded topic and focus-related rules from narrow syntax (see also Zubizarreta 1998). They all both point out that focus cannot be defined until the derivation is a single P-marker. There is a loss of generalization and a violation of the inclusiveness condition if some movement in the narrow syntax is triggered by focus features. Note also that focus assignment depends on information contours having been assigned.

7 On the Perfectness of Language

What does the Split-CP hypothesis have to do with Chomsky's idea that language might be in some sense 'perfect'? As we will see in this section, the idea is directly related to the triggering of and landing sites for movement and, in particular, the kinds of A′-movements that we have been looking at.

When Chomsky talks about perfection and optimality, he means it in a fairly narrowly conceived way. Language might be perfect from the point of view of its internal structure, that is, for interaction with systems that are internal to the mind, not in terms of its use as a vehicle for communication. In Chomsky 2002, he gives an interesting analogy, not for the first time comparing language to the liver. The liver in Chomsky's opinion is not particularly well designed from the point of view of its external function. After all, it does not take too much drinking or exposure to toxic chemicals to wreck it completely. But the liver is, Chomsky speculates, beautifully designed from the point of view of its internal function, that is, for interaction with the circulatory system, the kidneys, and so on.

Given the preceding, there are two ways that Chomsky seems to view language as possibly 'perfect'. In some places he flirts with an extremely strong criterion for perfection, along the lines of (37):

(37) STRONG PERFECTION: Language is perfect/optimal because all that is needed in grammar are the elementary constructs of the Minimalist Program (Select, Merge, Move, etc.).

As Chomsky notes, distinct principles of UG challenge (37). Consider his remarks on two of the most celebrated such principles ever proposed:

> For example, take the LCA [Linear Correspondence Axiom (Kayne 1994)]. If that theory is true, then the phrase structure is just more complicated. … If the ECP [Empty Category Principle (Chomsky 1981)] really works, well, too bad; language is more like the spine [i.e., poorly designed — FJN] than like a snowflake [i.e., optimally designed]. (Chomsky 2002: 136)

In other words, Chomsky regards these restrictive principles as inherently undesirable.

If something like Strong Perfection were correct, then one sign of the optimality of language would be that, with the passage of time and the increase of understanding, fewer and fewer principles like LCA and ECP would be proposed. But that is certainly not the case. As many distinct UG principles are being proposed today as were proposed 20 years ago. I would go so far as to claim that no paper has ever been published within the general rubric of the MP that does not propose some new UG principle or make some new stipulation (however well motivated empirically) about grammatical operations that does not follow from the bare structure of the MP. In fact, just in order to describe the facts in question, virtually every minimalist-oriented paper posits an ancillary set of hypotheses, either directly borrowed from the GB-era or posited ad hoc for the analysis in question. And for those UG principles that are proposed, it is difficult to see how they make language any more perfect or less perfect. So consider the discussion in Belletti & Rizzi 2002. They also take the position that language might be perfect and endorse principles such as movement-as-last-resort, the idea that derivations take place in 'phases' corresponding to VP and CP; the principle of Relativized Minimality; and the copy theory of movement. These principles might be right or they might be wrong, but language would be no more or no less 'perfect' if movement were utterly forbidden, if DP and AP were also phases, if minimality were not relativized, and if when an element moves, it moves, *tout court*. So it seems pretty far-fetched that language might be perfect in this strong sense.

In most of his discussions of perfection, Chomsky gives a somewhat weaker characterization of perfection in language:

(38)   WEAK PERFECTION: 'In a perfectly designed language, each feature would be semantic or phonetic, not merely a device to create a position or to facilitate computation.' (Chomsky 2000: 109)

I am not sure why (38) would be a marker of perfection, if it were right, but it is certainly not right. Chomsky himself gives a number of counterexamples to perfection in the sense of (38). The most serious is the fact that languages contain features that make no contribution to interpretation. Take, for example, features such as structurally-defined Case and agreement. Chomsky does, however, speculate that these uninterpretable features exist for a reason:

[Uninterpretable features such as structurally-defined Case and agreement] are there as perhaps an optimal method of implementing something else that must be there, namely dislocation.' (Chomsky 2002: 113)

So here is where we find the link between dislocation and supposed perfection that unites the two parts of this paper. In fact, I feel that there are serious problems tying movement to the need to eliminate uninterpretable features and will return to these problems at the end. But a logically prior question is why

movement itself should exist. In other words, why don't grammars simply manifest a one-to-one relation between surface position and meaning (however broadly one might want to define meaning), with no movement at all? Chomsky gives a functional explanation for movement, though not a particularly convincing one. He distinguishes between 'deep semantics' (thematic relations among elements) and 'surface semantics' (information structure properties such as 'topic' and 'focus') and writes that:

> … if surface semantics were signaled by inflection [like deep semantics is; FJN], the underlying morphological system would be complicated. For elements with Inherent Case, there would be double inflection if they have distinct surface-semantic properties; for elements lacking Inherent Case, they would have inflection only in this case. In contrast, if surface properties are signaled configurationally, at the edge, the morphological system is uniform throughout: a single Case inflection always (whether manifested phonologically or not). (Chomsky 2002: 122)

That speculation of Chomsky's is called into question in about half a dozen different ways. (I-VI) enumerate a number of difficulties with trying to explain movement in terms of the need to signal surface semantics differently from deep semantics:

(I) Many languages signal surface semantics inflectionally (i.e. without movement), as is evidenced by the widespread use of topic and focus inflections in the languages of the world:

   (39) John-wa   Mary-oaisi-te i-ru (Japanese)
        John-THEME Mary-OBJ lov-ing is
        'John loves Mary.'

(II) Other languages signal surface semantics by means of a (non-inflectional) marker, again without movement:

   (40) Cali baa sheegay in Muuse yimid (Somali)
        Cali FOC said that Muuse came
        'CALI said that Muuse has come.'

(III) Others signal surface semantics by intonational means (without accompanying movement), say, by placing heavy stress on a focused element:

   (41) MARY left (not Sue).

(IV) Still others use canonical (i.e. non-dislocated) positions to indicate topic and focus. So for example, in English subject position is conventionalized as topical.

(V) Not even the focusing implicit in direct questioning requires movement. Only about half of the world's language have *wh*-movement (Dryer 1991):

(42) ni   kanjian-le shei   (Chinese)
you see-ASP   who
'Who did you see?'

(VI) 'Double inflection' could easily be avoided by means of a single inflection that coded, say, both topichood and agency, just as in many Indo-European languages a single inflection codes gender, number, and case.

I therefore fail to understand why the distinction between the two types of semantic properties would explain why movement exists. Why does movement exist then? I am quite sure that there is not one sole answer. Movement certainly can function to highlight information structure differences, as in (43a, b), even if there are other ways of signaling the same thing:

(43) a. Mary contacted John about the party.   (*Mary* is 'thematic')

b. John was contacted by Mary about the party. (*John* is 'thematic')

But movement also serves the needs of language processing, where semantics is not involved at all. Consider the postposing of 'heavy' elements in English:

(44) a. ?I consider [everybody who agrees with me and my disciples about the nature of the cosmos] smart.

b. I consider smart [everybody who agrees with me and my disciples about the nature of the cosmos].

(44a, b) mean the same thing, but (44b) is easier to process. Moving the heavy phrase in brackets allows a shorter recognition time for the constituents of the sentence. It seems pretty clear to me that Chomsky would not want to integrate the idea of movement to serve processing ease into his ideas about the perfection of language, given that he has ruled out the desirability of grammatical operations whose only role is to facilitate computation.

   Let us now look more closely at Chomsky's idea that uninterpretable features like structural case and agreement exist to trigger movement. To a certain extent this idea is a minimalist implementation of a mechanism that was present in GB, namely the Case Filter. Movement in GB was optional, but if it failed to apply to *Mary* and *John* in structures like (45a, b), a violation of the filter resulted:

(45) a. *e* was seen Mary

b. *e* seems John to have left

In the Minimalist Program, movement is obligatory, triggered by the need to eliminate the uninterpretable (structural) nominative Case feature of the subject. Given the VP-internal subject hypothesis, which requires even seemingly in situ subjects to move, the MP derives the same results as GB, as

far as subjects are concerned. For objects, however, things are less clear. In GB, the apparent rarity of movement into direct object position was attributed to the idea that verbs assign thematic roles to the elements that they subcategorize. So movement into object position would violate the Theta Criterion. Given the unclear status of the Theta Criterion in the MP, I am not sure a parallel explanation of the relative rarity of such movement is still available. If not, we would predict (incorrectly) widespread examples of movement into object position, which would take place in order to eliminate the uninterpretable feature of structural object Case. Going along with that, there is no non-stipulative mechanism for deriving non-dislocated objects via movement. That is, there is nothing with parallel motivation to the VP-internal subject hypothesis that would require objects to move.

Consider now grammatical gender. This feature is as 'uninterpretable' as structural Case and agreement. One would assume that Chomsky sees no semantic basis for distinctions such as that between *la lune* and **der Mond**. Since eliminating the uninterpretable gender feature cannot be tied to movement, we are left with one such feature that requires some other mechanism for its 'elimination'.

Problems continue when we look at A'-movements such as *wh*-movement and topic and focus movements. As far as the former is concerned, surely the *wh*-feature is *interpretable*, being the locus for whatever semantic features distinguish *wh*-expressions from their 'un-*wh*'ed' counterparts. Rizzi seems to be of the same opinion, as is evidenced by the quotation in §2 of the present paper. So in order to attribute *wh*-movement to the need to eliminate uninterpretable features, one would need to posit additional, uninterpretable and for the most part unnecessary, *wh*-features. In my view, so doing reduces the hypothesized linkage between uninterpretable features and dislocation to vacuity.

What is totally frustrating for a dyed-in-the-wool generative grammarian like myself is that Chomsky's appeal to the seeming 'simplicity' of a theory as grounds for adopting it was something that he always used to oppose. Let's go back in time 30 years back to the heyday of generative semantics. Generative semanticists also believed that their theory provided the optimal (maximally elegant) mapping between sound and meaning and for that reason should be accepted as the null hypothesis regarding grammatical architecture. Recall that in its early days the generative semantic model contained nothing but a semantic level and surface level and a transformational mapping between them. As Paul Postal remarked:

> What I wish to suggest briefly is that because of its a priori logical and conceptual properties, [Generative Semantics] is the basic one which generative linguists should operate from as an investigatory framework, and that it should be

abandoned, if at all, only under the strongest pressures of empirical disconfirmation. (Postal 1972: 135)

Chomsky at the time dismissed claims such as Postal's as misguided. What he stressed was not a conceptually simple organization for a theory, but rather a rich internal structure, capable of narrowing down the class of possible grammars:

> Notice that it is often a step forward, then, when linguistic theory becomes complex, more articulated and refined — a point that has been noted repeatedly. … Thus it is misleading to say that a better theory is one with a more limited conceptual structure, and that we prefer the minimal conceptual elaboration, the least theoretical apparatus. Insofar as this notion is comprehensible, it is not in general correct. (Chomsky 1972: 67-68)

Ironically, the Chomsky of the 21st century reads like a reincarnation of the Postal of the 20th. There is next to nothing on the question of restrictiveness in Chomsky's recent writings, but page after page extolling the 'a priori logical and conceptual properties' of the MP.

8   Conclusion

For many decades, displaced elements have been the testing ground for linguistic theories. In the 1950s and 1960s, their distinctive properties provided the strongest arguments for discarding earlier structuralist accounts and adopting generative grammar. The battles between mainstream transformational approaches and lexicalist approaches again have focused on displaced elements as providing the best evidence for one or the other. And I hope to have shown that their importance continues today. In particular, when one looks at displaced elements closely, one sees that they provide no evidence for the more elaborate views of phrase structure and licensing that have become part and parcel of the minimalist program, and no evidence for the exaggerated claims about the perfection of language that have tended to accompany these views.

Acknowledgement

References

Arnold, Jennifer E., Thomas Wasow, Anthony Losongco and Ryan Ginstrom. 2000. Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering. *Language* 76: 28-55.

Baltin, Mark. 1982. A Landing Site Theory of Movement Rules. *Linguistic Inquiry* 13: 1-38.

Belletti, Adriana and Luigi Rizzi. 2002. Editors' Introduction. In *Noam Chomsky: On Nature and Language.* Cambridge University Press, Cambridge, pp. 1-47.

Bresnan, Joan W. 1970. On Complementizers: Toward a Syntactic Theory of Complement Types. *Foundations of Language* 6: 297-321.

Breul, Carston, 2003. *Focus Structure in Generative Grammar: An Integrated Syntactic, Semantic, and Intonational Approach to the Categorial/thetic/identificational Distinction*. Benjamins, Amsterdam.

Chomsky, Noam. 1971. Deep Structure, Surface Structure, and Semantic Interpretation. In D. Steinberg and L. Jakobovits, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. Cambridge University Press, Cambridge, pp. 183-216.

Chomsky, Noam. 1972. Some Empirical Issues in the Theory of Transformational Grammar. In P. S. Peters, ed., *Goals of Linguistic Theory*. Prentice-Hall, Englewood-Cliffs, pp. 63-130.

Chomsky, Noam. 1976. Conditions on Rules of Grammar. *Linguistic Analysis* 2: 303-351. Reprinted in Noam Chomsky, 1977, *Essays on Form and Interpretation*. North-Holland, New York, pp. 163-210.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Studies in Generative Grammar, Vol. 9. Foris, Dordrecht.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik,* MIT Press, Cambridge, MA, pp. 89-155.

Chomsky, Noam. 2001. Derivation by Phrase. In M. Kenstowicz, ed., *Ken Hale: A Life in Language.* MIT Press, Cambridge, MA, pp. 1-52.

Chomsky, Noam. 2002. *On Nature and Language.* Cambridge University Press, Cambridge, MA.

Cinque, Guglielmo. 1993. A Null Theory of Phrase and Compound Stress. *Linguistic Inquiry* 24: 239-297.

Culicover, Peter W. 1991a. Polarity, Inversion, and Focus in English. *Proceedings of the Eighth Eastern States Conference on Linguistics*: 46-68.

Culicover, Peter W. 1991b. Topicalization, Inversion, and Complementizers in English. In E. A. D. Delfitto, ed., *OTS Working Papers: Going Romance and Beyond*. University of Utrecht, Utrecht.

Dryer, Matthew. 1991. SVO Languages and the OV:VO Typology. *Journal of Linguistics* 27: 443-482.

Emonds, Joseph E. 1976. *A Transformational Approach to English Syntax*. Academic Press, New York.

Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge University Press, New York.

Fodor, Janet D. 1978. Parsing Strategies and Constraints on Transformations. *Linguistic Inquiry* 9: 427-473.

Fodor, Janet D. 1984. Constraints on Gaps: Is the Parser a Significant Influence? In B. Butterworth, B. Comrie and Ö. Dahl, eds., *Explanations for Language Universals*. Mouton, Berlin, pp. 9-34.

Fox, Danny. 1995. Economy and Scope. *Natural Language Semantics* 3: 283-341.

Haegeman, Liliane. 1991. Parenthetical Adverbs: The Radical Orphan Approach. In S. Chiba, et al., eds., *Aspects of Modern English Linguistics: Papers Presented to Masatomo Ukaji* on his Sixtieth Birthday. Tokyo, Kaitakusha, pp. 232-254.

Haegeman, Liliane and Jacqueline Guéron.1999. *English Grammar: A Generative Perspective*. Blackwell, Oxford.

Hatakeyama, Yuji. 1997. An Analysis of Inverse Copula Sentences and the Theoretical Consequences for Clause Structure: A Feature Compositional Approach to the Split-CP Hypothesis. *Linguistic Analysis* 27: 26-65.

Horvath, Julia. 1995. Structural Focus, Structural Case, and the Notion of Feature-Assignment. In K. É. Kiss, ed., Discourse Configurational Languages. Oxford University Press, Oxford, pp. 28-64.

Horvath, Julia. 1999. Interfaces vs. the Computational System in the Syntax of Focus. In H. Bennis and M. Everaert, eds., *Interface Strategies*. Royal Academy of the Netherlands, Amsterdam, pp. 183-205.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. MIT Press, Cambridge, MA.

Jacobs, Joachim. 1984. Funktionale Satzperspektive und Illokutionssemantik. *Linguistische Berichte* 91: 25-28.

Jacobs, Joachim. 1988. Fokus-Hintergrund-Gliederung und Grammatik. In H. Altmann, ed., *Intonationsforschungen*. Niemeier, Tübingen, pp. 89-134.

Jacobs, Joachim. 1992. Focus Ambiguities. In H. Jakob, ed., *Proceedings of the Conference on Focus and Intonation*. Fraunhofer-Institut, Stuttgart.

Jayaseelan, K. A. 2002. Question Movement and how Some SOV Languages Handle It. Paper presented at GLOW in Asia Colloquium.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, MA.

Krifka, Manfred. 1991. A Compositional Semantics for Multiple Focus Constructions. In S. Moore and A. Z. Wyner, eds., *Proceedings of the First Semantics and Linguistic Theory Conference (= Cornell University Working Papers in Linguistics*, Number 10). Cornell University, Ithaca.

Ladd, D. Robert. 1978. *The Structure of Intonational Meaning: Evidence from English*. Indiana University Press, Bloomington.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge University Press, Cambridge.

Lappin, Shalom. 1982. Quantified Noun Phrases and Pronouns in Logical Form. *Linguistic Analysis* 10:131-159.

Larson, Richard K. 1988. On the Double Object Construction. *Linguistic Inquiry* 19: 335-392.

Larson, Richard K. 1990. Double Objects Revisited: Reply to Jackendoff. *Linguistic Inquiry* 21: 589-632.

Lasnik, Howard and Mamoru Saito. 1992. *Move α: Conditions on Its Application and Output*. Current Studies in Linguistics, Vol. 22, MIT Press, Cambridge, MA.

Lasnik, Howard and Timothy A. Stowell. 1991. Weakest Crossover. *Linguistic Inquiry* 22: 687-720.

Lee, Felicia 2001. WH- and Focus Are Not the Same Projection. *West Coast Conference on Formal Linguistics* 20: 346-357.

Lopez, Luis. 2002. Toward a Grammar without TopP or FocP. Paper presented at Workshop on the Syntax-Semantics Interface in the CP-Domain.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* MIT Press, Cambridge, MA.

Newmeyer, Frederick J. 2003a. Review of *On Nature and Language* by Noam Chomsky. *Language* 79: 583-599.

Newmeyer, Frederick J. 2003b. Against the Split-CP Hypothesis. *Generative Linguistics in Poland* 5.

Pesetsky, David M. 1987. Wh-in-Situ: Movement and Unselective Binding. In E. J. Reuland and A. G. B. ter Meulen, eds., *The Representation of (in)Definiteness*. MIT Press, Cambridge, MA, pp. 98-129.

Postal, Paul M. 1972. The Best Theory. In P. Stanley Peters ed., *Goals of Linguistic Theory*. Prentice-Hall, Englewood Cliffs, pp. 131-170.

Reinhart, Tanya. 1991. Elliptic Conjunctions — Non-Quantificational LF. In A. Kasher, ed., *The Chomskyan Turn: Generative, Linguistics, Philosophy, Mathematics, and Psychology*. Blackwell, Oxford, pp. 360-384.

Reinhart, Tanya. 1997. Interface Economy: Focus and Markedness. In C. Wilder, H.-M. Gärtner and M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 146-169.

Rizzi, Luigi. 1996. Residual Verb Second and the *Wh*-Criterion. In A. Belletti and L. Rizzi, eds., *Parameters and Functional Heads: Essays in Comparative Syntax.* Oxford University Press, Oxford, pp. 63-90.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In L. Haegeman, ed., *Elements of Grammar: Handbook of Generative Syntax*. Kluwer, Dordrecht, pp. 281-337.

Rochemont, Michael S. 1998. Phonological Focus and Structural Focus. In P. Culicover and L. McNally, eds., *Syntax and Semantics, Volume 29: The Limits of Syntax*. Academic Press, San Diego, pp. 337-364.

Rochemont, Michael S. and Peter W. Culicover. 2000. Deriving Dependent Right Adjuncts in English. In D. Beermann, D. Leblanc and H. v. Riemsdijk, eds., *Rightward Movement*. John Benjamins, Amsterdam, pp. 234-257.

Rooth, Mats. 1985. *Association with Focus*. Ph. D. diss., University of Massachusetts.

Ross, John R. 1967. *Constraints on Variables in Syntax*. 1967 Ph.D. diss., MIT, Cambridge, MA. Published in 1985 as *Infinite Syntax!* Ablex, Norwood, NJ.

Shaer, Benjamin. 2003a. On The Syntactic Status of Certain Fronted Adverbials in English. *Berkeley Linguistics Society* 29.

Shaer, Benjamin. 2003b. An 'Orphan' Analysis of Long and Short Adjunct Movement in English. *West Coast Conference on Formal Linguistics* 22**:** 450-463.

Sobin, Nicholas. 2002. A Note on Echo Questions and the Split CP Hypothesis. Unpublished ms, University of Wales.

Stechow, Arnim von. 1989. *Focusing and Backgrounding Operators*, Universität Konstanz, Fachgruppe Sprachwissenschaft, Arbeitspapier nr. 6, Konstanz.

Szendröi, Kriszta. 2001. *Focus and the Syntax-Phonology Interface.* Ph.D. diss., University College London.

Vilkuna, Maria. 1995. Discourse Configurationality in Finnish. In K. É. Kiss, ed., *Discourse Configurational Languages*. Oxford University Press, Oxford, pp. 244-268.

Watters, J. 1979. Focus in Aghem. In L. M. Hyman, ed., *Aghem Grammatical Structure*: *Southern California Occasional Papers in Linguistics* 7, Los Angeles.

Zubizarreta, Maria Luisa. 1998. *Prosody, Focus, and Word Order*. Linguistic Inquiry Monograph, Vol. 33, MIT Press, Cambridge, MA.