

Projection, Heads, and Optimality*

Jane Grimshaw

Department of Linguistics and Center for Cognitive Science

Rutgers University

May 1995

Department of Linguistics
18 Seminary Place
New Brunswick NJ 08903

grimshaw@rucss.rutgers.edu

* This research was generously supported by Rutgers University. The analysis has benefited immensely from discussions with, and (often extensive) comments from, the following colleagues: Peter Ackema, David Adger, Maria Bittner, Eric Bakovic, Claudia Borgonovo, Christine Brisson, George Broadwell, Daniel Büring, Harald Clahsen, Veneeta Dayal, Viviane Déprez, Liliane Haegeman, Hubert Haider, Katharina Hartmann, Marco Haverkort, Teun Hoekstra, Norbert Hornstein, Riny Huybregts, Sabine Iatridou, Ingeborg Lasser, Geraldine Legendre, Jim McCloskey, Armin Mester, Gereon Müller, Ad Neeleman, Jaye Padgett, Susan Pintzuk, Andrew Radford, Vieri Samek-Lodovici, Yael Sharvit, Nomi Shir, Sten Vikner, Fred Weerman, Ken Wexler, Edwin Williams, Ellen Woolford, Raffaella Zanuttini. Many suggestions came from audiences at Brandeis University, CUNY, University of Delaware, Girona International Summer School in Linguistics, Lund University, University of Massachusetts at Amherst, MIT, the University of Pennsylvania, Rutgers University, Universität Stuttgart, and Wuppertaler Linguistisches Kolloquium. I owe a particular debt to Alan Prince and Paul Smolensky, for their extremely generous contributions to this work, and to Vieri Samek-Lodovici for his invaluable scrutiny of the manuscript. Two reviewers provided very helpful commentaries.

1. Introduction

The goal of this paper is to show that the distribution of heads in English clauses, including inversion of the subject and auxiliary verb, and the appearance of the auxiliary verb *do* and the complementizer *that*, can be explained by the interaction of universal constraints, under Optimality Theory as proposed in Prince and Smolensky 1993. The core of Optimality Theory lies in these ideas:

Constraints are universal.

Constraints can be violated.

Grammars are rankings of constraints.

The optimal form is grammatical, all non-optimal candidates are ungrammatical.

An optimal output form for a given input is selected from among the class of competitors in the following way: A form which, for every pairwise competition involving it, best satisfies the highest-ranking constraint on which the competitors conflict, is *optimal*.

The constraints which play a central role in the proposal are these:¹

Constraints related to Specifiers:

OPERATOR IN SPECIFIER (OP-SPEC) position	-- syntactic operators must be in specifier position
CASE-MARKING (CASE)	-- DPs must be case marked
SUBJECT (SUBJ)	-- clauses have subjects

Constraints related to heads:

OBLIGATORY HEADS (OB-HD)	-- a projection has a head
HEAD LEFT (HD-LFT)	-- the head is leftmost in its projections
HEAD RIGHT (HD-RT)	-- the head is rightmost in its projections
NO MOVEMENT OF A LEXICAL HEAD (NO-LEX-MVT)	-- a lexical head cannot move

Government Constraints:

TRACE GOVERNMENT (T-GOV)	-- a trace is governed
TRACE IS LEXICALLY GOVERNED (T-LEX-GOV)	-- a trace is lexically governed

Others:

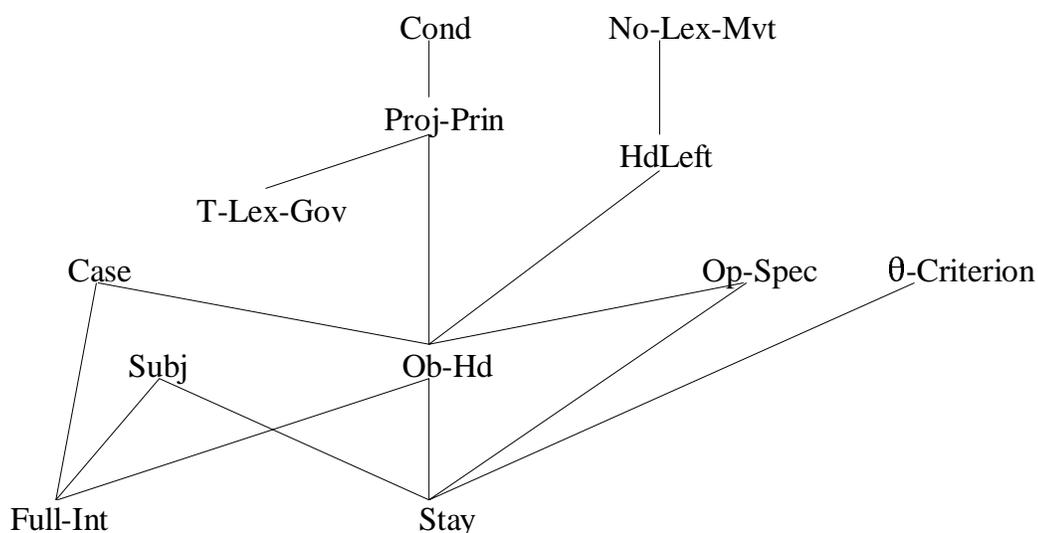
PROJECTION PRINCIPLE (PROJ-PRIN) movement into the head of a subordinate clause	-- no adjunction to subordinate clauses; and no movement into the head of a subordinate clause
ECONOMY OF MOVEMENT (STAY)	-- trace is not allowed
FULL INTERPRETATION (FULL-INT)	-- lexical conceptual structure is parsed
CONDITIONAL (COND)	-- a dependent head c-commands the extended projection containing it

¹ Grimshaw 1993 and 1994 assumed an additional constraint, "Minimal Projection", which required that a functional projection make a contribution to the functional representation of the extended projection that it is part of, thus ruling out entirely empty projections. In the current system, there is no need to stipulate such a constraint: the effects follow from OB-HD and STAY, as will be seen in Section 2.

I will discuss the formulation of the constraints and their relationship to ideas in the literature at the appropriate points in the discussion to come. The constraints are assessed at s-structure, and the account is, as far as I can tell, entirely consistent with the view that s-structure is freely constructed from the primitives of the theory, including traces, and then evaluated by well-formedness constraints, among them the ones in this paper.

Because OT constraints are violable, they are not in general “surface-true”. Under standard assumptions, positing a constraint which is violated requires corrective work. The constraint may be modified to a less general form so that no violation occurs, or taken to be satisfied by an invisible element or piece of structure. Or it may be relegated to some level of representation where it can be held to be unviolated, LF having proved most useful for this purpose. Under OT, violability is the norm, and it is this which makes it possible to have general constraints freely interacting. The generality of the constraints leads to systematic conflicts. For example, OB-HD may mandate movement to provide a head for a projection, yet STAY is always violated by movement. Similarly, PROJ-PRIN will prohibit head movement if the projection is a subordinate clause. Such conflicts, and their resolutions, prove essential to explaining the empirical generalizations at issue.

When constraints conflict, it is their relative ranking that determines which will be satisfied and which violated. The following rankings will be crucial in the proposal to come: Both OB-HD and OP-SPEC dominate STAY, OP-SPEC dominates OB-HD. FULL-INT is dominated by OB-HD, NO-LEX-MVT, SUBJ and CASE. PROJ-PRIN dominates OB-HD and T-LEX-GOV. HD-LFT dominates OB-HD. An ordering which is consistent with all of these dominance relations is: COND, PROJ-PRIN, OP-SPEC, HD-LFT, OB-HD, CASE, NO-LEX-MVT, FULL-INT, STAY, T-GOV, T-LEX-GOV. The constraints are cited in this order throughout.



The competitors, Prince and Smolensky's “candidate set”, are alternative realizations of an input.

The INPUT for a verbal extended projection is a lexical head plus its argument structure and an assignment of lexical heads to its arguments, plus a specification of the associated tense and semantically meaningful auxiliaries.

The input is passed to GEN (cf. Prince and Smolensky 1993), which generates all extended projections which conform to X' theory, i.e. in which all projections are of the right basic structure. Thus GEN can introduce such functional heads as do not appear in the input, due to their lack of semantic content, *that* being an example discussed in this paper. I assume a minimal X' theory, which simply says that each node must be a good projection of a lower one, if a lower one is present. The presence of a head is mandated by OB-HD, which is violable. (For a very different position on the head position see Bakovic (1995), where properties of the Spanish inversion system are argued to follow if OB-HD is considered to be part of GEN, hence inviolable.)

An extended projection (Grimshaw 1991; see also Haider 1989, van Riemsdijk 1990 for related proposals) is a unit consisting of a lexical head and its projection plus all the functional projections erected over the lexical projection. The smallest verbal projection is VP, but IP and CP are both extended projections of V. Other projections such as Agr-s, Agr-o can be incorporated into the overall program, but are not discussed here.

Suppose, for example, that the lexical head is *see*, and *John* is assigned to its external argument and *who* to its internal argument: $see(x,y)$, $x=John$, $y=who$. GEN will construct all extended projections of this lexical specification. Among the candidates will be some with *who* in object position, some with *who* in Spec of a functional projection, some with *that* included, some with *do* and so forth.² Competing candidates are evaluated as analyses of the same lexical material because of the way the input is defined. Thus, say, *John saw who* as an analysis of the input above is compared to other analyses of the same input, and not to, say, analyses of an input with a different lexical head, or with different arguments assigned to the predicate. I will assume that competing candidates have non-distinct logical forms, in a sense which must be made precise by further research, but which certainly must entail that they are truth functionally equivalent. Reinhart (1993) and Iatridou and Embick (1994) discuss this issue with respect to minimalism and economy (Chomsky 1991, 1992). It may turn out that the input should include a specification of LF-related properties, such as scope, as suggested in Smolensky et al. (1995).

The entire set of candidates is compared with respect to conformity to the set of violable constraints provided by UG and ranked by the grammar of the language, and the optimal one(s) survive as grammatical.

² Among other candidates are those which do not analyze elements of the input: those with violations of the Prince and Smolensky (1993) PARSE constraints. FULL-INT is a constraint of this type -- see 3.3.

2. The Basics of Inversion

Patterns of inversion follow from the interactions among universal constraints affecting heads in the verbal extended projection. The key constraints for this initial exploration are STAY, OP-SPEC, OB-HD.

First, consider an English matrix declarative sentence, where inversion is neither required nor allowed.

- (1) a. They will read some books
b. *Will they read some books

What well-formedness requirements are relevant for a matrix declarative? In an IP analysis of a matrix declarative with an auxiliary verb, such as the one in (1a), both projections have heads: a lexical V heads the VP and the auxiliary *will* heads IP. Thus there is no head missing from the structure, and here lies the crucial contrast between a declarative and an interrogative. *Wh* questions are operator variable constructions and the *wh* phrase is subject to OP-SPEC. OP-SPEC is based on the insight of Rizzi (1991) and Haegeman (1992), that there is a special relationship between the Specifier position and a syntactic operator, a scope-bearing expression which takes its scope by virtue of its syntactic position. All phrases which are marked as *wh* by virtue of percolation (through an extended projection) of a *wh* feature from a *wh* head or specifier (as in Grimshaw (1991)) count as *wh* operators. OP-SPEC requires that such expressions be Specifiers, motivating one more projection than occurs in the corresponding declarative. Specifier of VP is filled by the underlying subject, specifier of IP by the surface subject, hence the operator must be in specifier of an additional projection. This extra projection has no head and thus violates OB-HD, which requires a projection to have a head (either lexically realized or occupied by a trace) much as in Haider (1989). Head movement provides a head for the CP. *Wh* movement thus violates STAY in order to satisfy OP-SPEC. Similarly, inversion violates STAY, but it results in satisfaction of OB-HD.

- (2) a. Which books will they read ?
b. *Which books they will read ?

In questions introduced by *how come* there is no empty C because *how come* itself is a C (Collins 1991) and OB-HD is satisfied without inversion. To complete the explanation for (2), it is necessary to show that there is no C to head the CP in (2b), an obvious candidate being the complementizer *that*. Section 6 addresses this issue.

The basic idea, then, is that subject-auxiliary inversion in interrogatives is movement to C, following den Besten (1983), due to the OB-HD constraint, which induces movement to provide a head for a projection which would otherwise lack one. The projection is absent altogether in declaratives, hence no movement is necessary for the satisfaction of OB-HD.³

³ Verb second languages show a much-studied paradigm in which inversion is found in matrix declaratives. While there are a number of different accounts of V2, many view the typological

The tableau in (T1) shows the key competitors for the interrogative. If the interrogative is just an IP it will inevitably violate OP-SPEC (except when the *wh* phrase is the subject, see 3.4). The CP structures with no *wh* movement also violate OP-SPEC, and without inversion they violate OB-HD too. The CP structure with *wh* movement but no inversion violates OB-HD and STAY. The CP with *wh* movement and inversion violates STAY twice. It is optimal.

(T1) Matrix interrogatives

Candidates	OP-SPEC	OB-HD	STAY
[_{IP} DP will [_{VP} read what]]	*!		
[_{CP} e [_{IP} DP will [_{VP} read what]]]	*!	*	
[_{CP} what e [_{IP} DP will [_{VP} read t]]]		*!	*
 [_{CP} what will _i [_{IP} DP e _i [_{VP} read t]]]			**
[_{CP} will _i [_{IP} DP e _i [_{VP} read what]]]	*!		*

This illustrates some of the fundamentals of Optimality Theory. Every candidate, including the grammatical one, violates some constraint. The optimal one, however, marked with the pointing finger, violates only STAY, while the competitors all violate at least one constraint that is higher-ranked than STAY. (Because of the way constraint violations are reckoned, no number of violations of STAY would evict the winning candidate from its position as optimal, see Prince and Smolensky (1993).) Since OP-SPEC is the highest ranked of these constraints, any candidate which violates it will fail if there are candidates which do not violate it, as there are in this case.

Several rankings are crucial here. OP-SPEC must out-rank, or “dominate” STAY, otherwise the candidates with no *wh* movement would be preferred over candidates which satisfy OP-SPEC. OB-HD must out-rank STAY, for similar reasons: this is why head movement is possible if it leads to satisfaction of OB-HD. Furthermore, we might wonder why the *wh* phrase cannot be base generated in Specifier position, avoiding violation of STAY and satisfying OP-SPEC. This possibility is

property that distinguishes the V2 languages from the others as a property of the C position (see for example, the work reviewed in Weerman (1989) and Vikner (in press)), taking the filling of Spec of CP as a secondary property. The analysis of English given here invites a different perspective, in which what is special about these languages is that they require the presence of a Spec position of a particular kind, hence an extra projection can appear and OB-HD motivates V2 to fill the empty head position. (The specifier-related constraint in question must outrank STAY in V2 languages, and not in English -- see 3.3 and 7 for discussion of effects of re-ranking.) The correct formulation of the constraint requiring the Spec position is a complex issue, however, and I will not pursue it further.

eliminated, for arguments at least, if the Theta Criterion is invoked: either as a violable constraint dominating STAY or as an inviolable constraint (Vieri Samke-Lodovici p.c.). The result is that the system prefers to generate the *wh* phrase inside the VP to satisfy the Theta Criterion, at the cost of violating STAY in satisfying OP-SPEC.

In the declarative, OP-SPEC is always satisfied, vacuously, since there is no operator present. Hence the extra structure is not needed:

(T2) Matrix declaratives

Candidates	OP-SPEC	OB-HD	STAY
$\text{I} \rightarrow$ [_{IP} DP will [_{VP} read books]]			
[_{CP} e [_{IP} DP will [_{VP} read books]]]		*!	
[_{CP} will _i [_{IP} DP _{e_i} [_{VP} read books]]]			*!

When the matrix is an IP, every constraint considered so far is respected, and this is the optimal candidate. OB-HD is satisfied by *will* in I and by a lexical head in V. OP-SPEC is vacuously satisfied. When the matrix is a CP there are two alternative patterns. If inversion does not occur then OB-HD is violated, while if inversion does occur, STAY is violated. (This is why there is no need to appeal to the constraint “Minimal Projection” of Grimshaw (1993, 1994) to eliminate empty projections. Joan Bresnan points out, however, (p.c.) that empty adjunction structures are not excluded by the system.) Thus of the three competing possibilities discussed, the IP is the best. In effect, there is no point to inversion, since it leads to violation of STAY, with no compensating improvement on a higher-ranked constraint. In interrogatives, in contrast, the STAY violation has the benefit of allowing OB-HD to be satisfied.

Under the VP-internal subject hypothesis (Zagona 1982, Koopman and Sportiche 1991, Kitagawa 1986, McNally 1992, Burton and Grimshaw 1992), the subject in both the declarative and the interrogative has raised from Spec of VP to Spec of IP, satisfying CASE and SUBJECT but in violation of STAY. I will factor this out for now, and not represent it in the analysis, until it becomes strictly relevant, in 3.5.

Returning to the interrogatives, OP-SPEC merely requires that the *wh* phrase appear in a Specifier position — it says nothing about the position of this specifier relative to the remainder of the clause. Why then does the extra projection go on top of IP, instead of, for example, between IP and VP? This is due to the scopal properties of *wh* phrases: they take scope by virtue of their syntactic position, and must have scope over the entire propositional structure (roughly speaking IP) in order to perform their semantic function. There are several different ways to incorporate this into the present solution. The assumption in Grimshaw (1993) was that structures with the *wh* phrase in the

wrong Specifier would be uninterpretable and hence would never be grammatical. Here instead, I will take the tack that OP-SPEC requires an operator to be in the Specifier from which it takes its scope. Thus OP-SPEC for a *wh* operator will be satisfied only by movement to a position from which the operator c-commands the entire extended projection.

This solution predicts that when more than one *wh* phrase is present in a clause, only one moves, so that (3b), for example, is ungrammatical.

- (3) a. What will they put where?
b. *What will where they put?

This is because the movement of *where* to an internal specifier violates STAY, but fails to satisfy OP-SPEC since it is not moving the *wh* phrase into a scope position. Thus both candidates have one OP-SPEC violation, and the decision falls to STAY, which selects the candidate with fewest movements. The (simplified) tableau in (T3) illustrates this.

(T3) Multiple interrogatives

Candidates	OP-SPEC	OB-HD	STAY
☞ [CP what will [IP they [VP put t where]]]	*		*
[CP what will [where [IP they [VP put t t]]]]	*		***!

Movement to Specifier induces inversion, as we have seen. Adjunction, however, does not. Adjunction leaves the number of projections present in the representation unchanged:

- (4) a.
$$\begin{array}{c} \text{XP} \\ / \quad \backslash \\ \text{Spec} \quad \text{X}' \\ \quad / \quad \backslash \\ \quad \text{X} \quad \text{YP} \end{array}$$
- b.
$$\begin{array}{c} \text{XP} \\ / \quad \backslash \\ \text{ZP} \quad \text{XP} \\ \quad / \quad \backslash \\ \quad \text{Spec} \quad \text{X}' \\ \quad \quad / \quad \backslash \\ \quad \quad \text{X} \quad \text{YP} \end{array}$$

There are no more heads to be filled in an adjoined structure than in a structure with no adjunction and OB-HD will be satisfied in exactly the same way in both. Inversion will therefore never be necessary and therefore never be possible, since inversion always violates STAY. Along with Baltin (1982) and Lasnik and Saito (1992), I take topicalization (perhaps also the *wh* movement in exclamatives) to be a case of adjunction. See Cinque (1990), Lasnik and Stowell (1991), Postal

(1992), Müller and Sternefeld (1993) for recent discussion of differences between topicalization and true operator movements. Possibly English exclamatives, which have no inversion, are also adjunction structures.

In English matrix yes-no questions, inversion is again required.

- (5) a. Will they read some books?
b. *They will read some books?

The most straightforward way to subsume them under the same system is to posit a null operator in Spec of CP.⁴ Then matrix yes-no questions are CPs, the C position will be empty without inversion, and OB-HD will induce inversion. The assumption that yes-no questions involve an operator is crucial for other aspects of the syntax of interrogatives, such licensing of polarity items, and relativized minimality effects for interrogative complements (Rizzi 1990).

Empty layers of structure in the system presented here always violate OB-HD without movement and STAY with movement. The effects of a constraint against empty structure follow without stipulation.⁵ It is a consequence of competition among competing candidates that a clause is only as big as it needs to be. It is an IP unless it has to be a CP. A clause always has the minimal structure consistent with maximal satisfaction of constraints, under the OT theory of ranking. There is no fixed structure for clauses; no unique answer to the question of whether they are IPs or CPs. (or indeed VPs as we will see below). Sometimes they are one, sometimes the other, depending on what well-formedness conditions are relevant.

It should be noted that is crucial that all extended projections which can be formed above the lexical projection be in the candidate set, otherwise both the optimal CP candidates and the optimal IP candidates would be grammatical, giving inversion in declaratives alongside a no-inversion structure, and admitting an interrogative in which neither *wh* movement of the operator nor inversion occurred.

3. *Do Support*

The generalization governing *do* in English is simple to state but has proved challenging to formalize: that *do* is possible only when it is necessary (Chomsky 1957, 1991). The proposal developed so far in this paper allows us to make precise this conceptualization of *do*, given two

⁴ George Broadwell points out (p.c.) that cross-linguistically, overt yes-no question markers seem to be X-zeros rather than XPs, casting some doubt on analyses which treat yes-no questions and *wh* questions in parallel. This suggests exploring an alternative which would more closely resemble the proposal made in Section 6 for conditionals, where the operator is taken to be a head rather than a phrase.

⁵ In Grimshaw (1993,1994), however, “Minimal Projection” was necessary, because STAY was not among the constraints.

assumptions. First, *do* is a semantically and functionally empty verbal head: this seems to be the minimal specification we can give to *do*. As a result every occurrence of *do* violates FULL-INT. Second, OB-HD outranks FULL-INT, so *do* occurs when its presence results in satisfaction of OB-HD. The distribution of *do* then follows. The consequences are that *do* is impossible in (positive) matrix declaratives and subordinate interrogatives, required in matrix interrogatives with no auxiliary, never co-occurs with other auxiliaries and never co-occurs with itself.

3.1 Background

Emonds (1978) and Pollock (1989) argue that in French, the main verb raises to I, while in English the main verb appears to stay in its base position (see these works on English and French, Vikner (in press) on Scandinavian). If we assume, more or less as they do, that I is the locus of verbal inflection, then the French system is immediately comprehensible, since V raises to I in order to combine with its inflection. On the other hand, the English system becomes quite mysterious, since it is self-evident that the verb and its inflection do in fact combine, yet there is no obvious way for this result to be achieved. There have been a variety of responses to this problem. Under “checking theory” English main verbs raise at Logical Form (Chomsky 1993). In “Distributed Morphology” tense and the verb are merged in Morphological Structure (Halle and Marantz 1993). Williams (1994) argues that inflectional features are not syntactic nodes, but part of X-bar projections: under this view there is no V-to-I movement even in French type languages. Here I will take a position which maintains features of both the Emonds/Pollock proposal and the alternatives. Suppose that the difference between the French system and the English system is that in the English system inflection is morphologically associated with a V, i.e. it is lexically attached to a V head, while in French it is syntactically projected as head of a projection. (Whether inflection is syntactically projected as an I, or morphologically attached to V, the result is the same in one respect, namely the entire verbal extended projection has an inflectional specification, regardless of which (extended) head of the projection is the source of it). French will then have V to I raising, while English will not. If, as has been suggested by Pollock (1989), Platzack and Holmberg (1989), there is an important relationship between the existence of V-to-I movement and richness or “strength” of inflection, this should now be understood as reflecting a relationship between properties of inflection and existence as an independently projected head.

For concreteness I will adopt the position that the first auxiliary verb, like a finite main verb, is morphologically associated with phi features. It is generated in I. I assume a constraint, not further discussed here, which requires that the finite verb must c-command all other proposition-related heads in the extended projection. (This automatically penalizes extended projections containing more than one finite verb, and will in practice rule them out.) CASE and SUBJ constrain the position of the subject relative to functional heads in the extended projection. Their role will be examined in the discussion of negation, but in the meantime I consider only candidates which satisfy them both. A consequence of this analysis is that a tensed clause which contains a main verb and no auxiliaries, must be a VP in English, since inclusion of an IP in its representation will violate OB-HD with no

compensating effects.⁶ Thus main clauses can be VPs, IPs, or CPs, depending on their contents.

3.2 The Distribution of *do*

I will first show how the analysis works out for the contrast between matrix declaratives and matrix interrogatives, starting with declaratives. (I will omit CP representations for declaratives, always non-optimal for reasons given above.) The first issue, then, is why *do* does not occur in declaratives, even when there is no other auxiliary present.

- (6) a. She said that
b. *She did say that

(T4) Matrix declaratives with and without *do*

Candidates	OP-SPEC	OB-HD	FULL-INT	STAY
 [VP DP V that]				
[IP DP <i>do</i> [VP V that]			* !	
[IP e [VP DP V that]		*!		

(T4) shows how this result is obtained. Note that FULL-INT, which was satisfied in all previous cases, and was omitted from the tableaux for reasons of presentation, is now crucial. The VP form of the clause, with no auxiliary verb projection, is the optimal one, since it involves no violations, while the alternative with *do* violates FULL-INT, and the final alternative which has an extra projection but no *do*, violates OB-HD. Thus auxiliary *do* cannot occur in a declarative. The crucial difference between *do* and other auxiliaries is that they differ from *do* in having semantic content. They are part of the input.

In interrogatives, *do* must appear in the absence of another auxiliary verb.

- (7) a. What did she say?
b. *What she said?
c. *What she did say?

⁶ Adjunction to X-bar will have to be admitted, to accommodate the position of the adverb in *John usually likes tomatoes*, as pointed out to me by Sten Vikner and Marco Haverkort. If adjunction is regulated only by scope considerations and by PROJ-PRIN, this is not a problematic conclusion. Haverkort suggests that adjunction to X-bar is needed anyway for examples like the following: *Peter said that Carl, had he been on time, would have caught the train.*

I simplify by considering only those forms which do not violate OP-SPEC, i.e. where *wh* movement has occurred. Further, I will temporarily ignore candidates in which the main verb moves to supply the otherwise missing head, returning to the analysis of these cases in 3.3.

(T5) Matrix Interrogatives with and without *do*

Candidates	OP-SPEC	OB-HD	FULL-INT	STAY
 $[_{CP} \mathbf{wh} \mathbf{do}_i [_{IP} \mathbf{DP} \mathbf{e}_i [_{VP} \mathbf{V} \mathbf{t}]]]$			*	**
$[_{CP} \mathbf{wh} \mathbf{e} [_{IP} \mathbf{DP} \mathbf{e} [_{VP} \mathbf{V} \mathbf{t}]]]$		* ! *		*
$[_{CP} \mathbf{wh} \mathbf{e} [_{VP} \mathbf{DP} \mathbf{V} \mathbf{t}]]$		*!		*
$[_{CP} \mathbf{wh} \mathbf{e} [_{IP} \mathbf{DP} \mathbf{do} [_{VP} \mathbf{V} \mathbf{t}]]]$		*!	*	*

The optimal form is the one in which *do* occurs and inverts to C: this one violates only FULL-INT and STAY. Inversion of *do* satisfies OB-HD, which is violated in all other candidates. The ranking of OB-HD over FULL-INT is crucial here: the opposite ranking would select as grammatical a form in which the C is empty and *do* does not appear.

Thus we obtain the desired result: *do* is possible only where it is necessary, and it is necessary when its presence makes a clause obey a constraint that has a higher ranking than FULL-INT, in this case OB-HD.⁷ We will see in Section 4 that in subordinate interrogatives, OB-HD is violated anyway, regardless of whether an auxiliary verb is present or not. As a result there is no possible advantage to the inclusion of *do*, so it never occurs in embedded questions.

This analysis explains why *do* never co-occurs with another auxiliary verb, or with another token of *do*. Consider, for example, a matrix interrogative clause in which *do* and *will* co-occur. Movement of *will* to C satisfied OB-HD. There is no advantage to including *do* and its projection, which will always add a FULL-INT violation, hence only the *will* version is possible.

- (8) a. What will she say?
 b. *What will she do say?
 c. *What does she will say?

⁷ Although Roberts (1992: 293-294) notes that *do* insertion appears to have been freely available in 16th century English.

(T6) Co-occurrence of *do* and another auxiliary in a matrix interrogative

Candidates	OP-SPEC	OB-HD	FULL-INT	STAY
 $[_{CP} \mathbf{wh} \text{ will}_i [_{IP} \text{DP } \mathbf{e}_i [_{VP} \mathbf{V} \mathbf{t}]]]$				**
$[_{CP} \mathbf{wh} \text{ will}_i [_{IP} \text{DP } \mathbf{e}_i [_{XP} \mathbf{do} [_{VP} \mathbf{V} \mathbf{t}]]]]$			* !	**
$[_{CP} \mathbf{wh} \mathbf{do}_i [_{IP} \text{DP } \mathbf{e}_i [_{XP} \text{will} [_{VP} \mathbf{V} \mathbf{t}]]]]$			* !	**

Main *be* and *do* will not occur together. Since *be* has the capacity to move to C, to satisfy OB-HD, the form with *do* will have no competitive advantage over the one without, and the *do*-less form will be optimal.

Obviously, auxiliary *do* will also not repeat. Clauses containing *do* will be in competition with all otherwise equivalent clauses containing more and fewer occurrences of *do*. Each occurrence of *do* yields a violation of FULL-INT and no occurrence of *do* after the first one can ever improve the success of the structure with respect to the constraints. Hence a clause with more than one instance of *do* can never be optimal. This is illustrated with respect to a matrix question with no semantic auxiliary, where one *do* is possible:

- (9) a. What did she say?
 b. *What did she do say?
 c. *What did she do do say?

(T7) Multiple occurrences of *do* (illustrated for matrix interrogative)

Candidates	OP-SPEC	OB-HD	FULL-INT	STAY
 $[_{CP} \mathbf{wh} \mathbf{do}_i [_{IP} \text{DP } \mathbf{e}_i [_{VP} \mathbf{V} \mathbf{t}]]]$			*	**
$[_{CP} \mathbf{wh} \mathbf{do}_i [_{IP} \text{DP } \mathbf{e}_i [_{XP} \mathbf{do} [_{VP} \mathbf{V} \mathbf{t}]]]]$			** !	**
$[_{CP} \mathbf{wh} \mathbf{do}_i [_{IP} \text{DP } \mathbf{e}_i [_{XP} \mathbf{do} [_{XP} \mathbf{do} [_{VP} \mathbf{V} \mathbf{t}]]]]]]$			** ! *	**

The fundamental point of analysis for *do*, then, is that it violates FULL-INT, hence it is present only when its presence leads to improvement on a higher ranked constraint. Since OB-HD >> FULL-INT, *do* will appear in order to satisfy OB-HD.

3.3 Constraint Ranking and the Lexicon: Main Verbs and the Availability of *do*

In 3.2, I set aside a central property of the English inversion system, namely that main verbs do not invert, *do* occurring instead.

- (10) a. *What said she?
b. What did she say?

This motivates a constraint which penalizes movement of a lexical head, a the constraint named “NO-LEX-MVT”. This constraint is ranked above FULL-INT in English, therefore English will violate FULL-INT (introducing *do*) in order to avoid violating NO-LEX-MVT. Thus the full constraint analysis is the one given in (T8), which compares the winning candidate from (T5) with the alternative involving movement of a main verb.

(T8) Inversion of a Main Verb versus presence of *do*

Candidates	OP-SPEC	NO-LEX-MVT	OB-HD	FULL-INT	STAY
$[_{CP}wh V_i [_{VP} DP e_i t]]$		*!			**
 $[_{CP}wh do_i [_{IP}DP e_i [_{VP}V t]]]$				*	**

In English, then, we know that OB-HD >> FULL-INT, and that NO-LEX-MVT >> FULL-INT. There is no evidence concerning the ranking of OB-HD with NO-LEX-MVT, since the winning candidate violates neither. Evidence for ranking NO-LEX-MVT over OB-HD is sketched in footnote 15.

What of a language that allows main verbs to move? Such a language must have both OB-HD and FULL-INT dominating NO-LEX-MVT. This ranking yields a system which selects filling a head position with a main verb over leaving the head position unfilled or filling it with a meaningless item. This is schematized in (T9):

(T9) Effect of re-ranking on Inversion of a Main Verb versus presence of *do*

Candidates	OP-SPEC	OB-HD	FULL-INT	NO-LEX-MVT	STAY
 $[_{CP}wh V_i [_{VP} DP e_i t]]$				*	**
$[_{CP}wh do_i [_{IP}DP e_i [_{VP}V t]]]$			*!		**
$[_{CP}wh e [_{VP}V t]]$		*!			*

As in the English case, there is no crucial ranking here of OB-HD and FULL-INT — either ranking will yield the desired result.

The difference between the English ranking and the alternative is farther-reaching than might appear at first glance: it does not just concern the ability of a main verb to move. A grammar with NO-LEX-MVT dominated by both FULL-INT and OB-HD will be inconsistent with the use of a semantically empty verb like English *do* in inversion. In fact it will be inconsistent with the *existence* of such a verb, which can never appear. Thus we can derive a gap in the behavior of verbs from the ranking of constraints, following the line of reasoning developed by Prince and Smolensky (1993) for segmental inventories.

Pursuing this point further, we can also ask whether it is an accident that English has a semantically empty verb, and whether it is an accident that the morpheme *do* is the one which has this use. Can we characterize the presence of *do* in English as a necessary consequence of the constraint rankings, just as its absence in the grammar of (T9) is a necessity? The starting point for the answer is that there is only one *do* in English. When its lexical conceptual structure (*lcs*) is parsed, *do* is a theta-marking and argument-taking predicate. However, its *lcs* can also be unparsed, and this is the source of "auxiliary" *do*.

(11)

- a. parsed *lcs*: *do* ---- (ACT x)
- b. unparsed *lcs*: *do* --/-- (ACT x)

When its *lcs* is unparsed, *do* is a "light verb" in the sense of Grimshaw and Mester (1988), using a term from Jespersen (1954). Light *do* will always vacuously satisfy NO-LEX-MVT. Only an element with an *lcs* is "lexical", hence only such an element is subject to the constraint. This is why *do* shows the characteristic auxiliary-verb property of appearing outside the lexical projection: it patterns with the semantic auxiliaries and presumably *be* in having no *lcs*. Light *do* will, however, violate FULL-INT, since it will have no *lcs* interpretation associated with it. With parsed *lcs*, FULL-INT will be satisfied, so "main" or "heavy" *do* does not violate FULL-INT. Thus, light *do* will not be affected by NO-LEX-MVT, but it will violate FULL-INT, while heavy *do* will be regulated by NO-LEX-MVT but will satisfy FULL-INT.

We can now address the question of why *do* is the morpheme that appears when an auxiliary is required by OB-HD. This follows if, as seems reasonable enough, *do* is the lexically simplest verb of the language, though perhaps it is necessary to appeal to further details of its analysis in order to explain why *do* violates FULL-INT minimally, rather than *be*. At any rate, the reason *do* appears and not *shout* or *obfuscate* is that the *lcs* of *do* is simpler, hence there is less in the *lcs* of *do* to be unparsed. Failure to parse the *lcs* of *do* results in a less severe violation of FULL-INT than failure to parse the *lcs* of any other verb in the language. This proposal treats FULL-INT as a "gradient constraint" (see Prince and Smolensky (1993) for theoretical background.)

We can also see that it is inevitable that light *do* exists in the language, given the constraint rankings. A grammar with the rankings of English will necessarily select the least offensive FULL-INT violation in order to satisfy OB-HD, and since *lcs*-unparsing is freely available there will always be (at least) one candidate which minimally violates FULL-INT and which is therefore optimal.

Finally, one might ask why "auxiliary" *do* must occur even when the main verb is itself *do*, if they are really both the same verb. The answer has already been implicitly given. When *do* has its *lcs* parsed it can theta-mark but cannot occur outside VP. When its *lcs* is unparsed it can occur outside VP but cannot theta-mark. Thus (12a) violates NO-LEX-MVT when *do* has a parsed *lcs*, but when it lacks one, the sentence violates the theta criterion, since *do* cannot theta-mark when its *lcs* is unanalyzed. Assuming that the theta criterion dominates FULL-INT (perhaps because the theta criterion is never dominated), the optimal form will contain both occurrences of *do*.

- (12) a. *What did he e t?
b. What did he do t?

(T10) shows the interaction of the Theta Criterion and the other constraints in the sentences of (12), for candidates which satisfy OB-HD and OP-SPEC.

(T10) Co-occurrence of heavy and light *do*

Candidates	THETA-CRIT	OP-SPEC	NO-LEX-MVT	OB-HD	FULL-INT	STAY
$[_{CP}wh \mathbf{do}_i [_{VP} he e_i t]]$ + <i>lcs</i>			*!			**
$[_{CP}wh \mathbf{do}_i [_{VP} he e_i t]]$ - <i>lcs</i>	*!				*	**
 $[_{CP}wh \mathbf{do}_i [_{IP}DP e_i [_{VP}\mathbf{do} t]]]$ - <i>lcs</i> + <i>lcs</i>					*	**
$[_{CP}wh \mathbf{do}_i [_{IP}DP e_i [_{VP}\mathbf{do} t]]]$ + <i>lcs</i> - <i>lcs</i>	*!					**
$[_{CP}wh \mathbf{do}_i [_{IP}DP e_i [_{VP}\mathbf{do} t]]]$ + <i>lcs</i> + <i>lcs</i>			*!			**
$[_{CP}wh \mathbf{do}_i [_{IP}DP e_i [_{VP}\mathbf{do} t]]]$ - <i>lcs</i> - <i>lcs</i>	*!				**	**

An advantage of this solution is that it eliminates the positing of ambiguity for *do*. That is, we

no longer have to say that English has a main verb *do* and an auxiliary verb which is accidentally identical to a main verb which itself is semantically extremely vague. There is only one G43 verb, recruited to occur in non-lexical positions because it is the least costly choice. Note that Japanese *suru* is similarly “ambiguous” between and light and heavy use (Grimshaw and Mester (1988)), inviting the same analysis. See Ritter and Rosen (1993) for related ideas about *have*.

In this view, there can not be a language which lexically lacks a semantically empty verb, any more than there can be a language which lexically lacks an epenthetic vowel. The occurrence of such items is not regulated by lexical stipulation. Rather, the semantically simplest verb will appear if the constraint rankings for the language induce its presence, and not otherwise, just as an epenthetic vowel appears when it is induced by constraint interaction. Thus we can predict both that a language with the constraint rankings in (T9) will not have a light verb like *do* and that a language with the constraint rankings of English will inevitably have such a verb. In this way, we avoid lexical stipulations concerning the inventory of the language and instead explain the visible lexical items as the result of constraint interaction. This general point is taken up again in Section 7.

3.4 *Do* and *Wh* Subjects

The one matrix interrogative configuration in which *do* does not appear is one with a *wh* subject:

- (13) a. Who saw it?
b. *Who did see it?

The key here is the formulation of OP-SPEC, which does not require that a *wh* operator occur in Specifier of CP, just that it occur in a Specifier position from which it c-commands the verbal extended projection. An interrogative clause is therefore not required to be a CP, but just a verbal projection with a *wh* specifier. When a *wh* phrase is a non-subject it will always have to move to a Spec position. The candidate positions are basically VP, IP, and CP. Both of the first two are filled, so non-subject *wh* phrases move to Spec of a higher projection, which we label CP. Specifier of VP is not a possible position for a *wh* operator, both because it is (usually) filled and because it does not give the *wh* operator the right scope.

The one case where the requirements of an interrogative are met without any movement is when a *wh* phrase is a subject. It is already in Specifier position, moreover it is in Specifier of the highest phrase in the verbal extended projection. Thus an IP (when the clause contains an auxiliary verb) or a VP (when there is no auxiliary present) will be a perfectly good interrogative, provided it has a *wh* phrase as subject.

As a result, a subject *wh* phrase can never occur in Spec of CP, because the STAY violation incurred by the movement will not be offset by satisfaction of OP-SPEC: OP-SPEC is already satisfied. By the same reasoning, a subject *wh* phrase cannot occur in Spec of IP when there is no auxiliary present: the *wh* phrase in this case would be moving from Spec of VP to Spec of IP, violating STAY, again with no improvement on OP-SPEC. This is why *do* never occurs in subject *wh* clauses.

(T11) The position of subject *wh* phrases

Candidates	OP-SPEC	NO-LEX-MVT	OB-HD	FULL-INT	STAY
 [VP <i>wh</i> V ...]					
[IP <i>wh</i> e [VP <i>t</i> V ...]]			* !		*
[IP <i>wh</i> do [VP <i>t</i> V ...]]				* !	*

Thus we conclude that subject *wh* phrases do not move, but remain in Specifier of the verbal head with phi features, i.e. IP or VP. This Spec position must then count simultaneously as an A-position (by virtue of its relationship to phi features) and an A-bar position (by virtue of being a scope position.) The idea that subject *wh* phrases do not move has been frequently entertained (see Gazdar (1981), Chung and McCloskey (1983), Haider (1989) (Rizzi 1991) for example, also Travis (1991) and Vikner and Schwartz (1991) Cardinaletti (1992) for discussion of similar issues in V2 systems).

It is important to counter an apparently overwhelming argument against the idea that *wh* questions with subject *wh* phrases are just VPs (when they contain no auxiliary verb) or IPs (when an auxiliary is present): this is based on selection. The very same predicates select questions with *wh* subjects as those with other *wh* phrases, despite the VP versus IP versus CP distinctions. How is it possible to explain why they are selected by the same predicates? This problem does not arise under the “Type-Category” theory of selection (Grimshaw 1991), in which any verbal projection with a [+wh] Spec is equivalent to any other verbal projection with the same property, as far as selection is concerned, hence it necessarily follows that a verb which takes an interrogative will combine with all of these sub-cases. See Section 8 for further discussion.

Interrogatives with *wh* subjects show that OP-SPEC does not stipulate which specifier must house the *wh* phrase. This suggests the desirability of a theory in which reference cannot be made to such information, in which notions like “CP”, “IP” really have no status. I return to this issue in Section 9 below..

3.5 Negative clauses with *do*

The appearance of negation, like inversion, induces the presence of *do*. The core idea is that when *not* is present, so is a higher projection which is absent from clause structure when *not* is absent. The head of this projection is an auxiliary verb, or *do* if no other auxiliary verb is present in the clause. In this way, the analysis of *do* with negation resembles the analysis for interrogatives, although, as

we will see, the situations are not identical.⁸

What requires the presence of the extra projection? There are a number of possibilities which could be explored. One would be to appeal to the idea that there is a requirement that T must c-command negation (Laka 1990). This will force the generation of a tensed verb above *not*, and when the only contentful verb is a main verb, it will force the generation of tensed *do* as a head c-commanding negation. A second line might try to use the idea that *not* is a clitic and a verb must appear to its left to support it. This is essentially the proposal of Roberts (1992), and as he argues, it explains why in English *do* support came in at the same time as *n't*. However, while a clitic analysis may be correct for *n't* (though see Williams 1994 for arguments that *n't* is lexically attached) it is less clearly so for *not*, which occurs unsupported elsewhere, e.g. in subjunctives, as we will see below.

The proposal I will make here attempts to use independently necessary constraints on subjects and verbs to achieve the desired result. The first constraint, SUBJECT, corresponds essentially to the “Extended Projection Principle” (Chomsky, 1981), and says that a clause must have a subject. There are two alternative formulations, either of which will be satisfactory here. The constraint may require that the highest A Specifier in a clause must be filled. Alternatively it could require that the Specifier of the highest I-related head be filled, where “I-related” includes V, T, Agr, Neg etc., i.e. the cluster of heads that are neither lexical nor type-affecting. Which of these formulations is to be preferred depends on the theoretical development of the notions “A-position” and “I-related”, and I will not address these questions here. The second constraint, CASE requires that the head of a DP A-chain be in a case position.

CASE must dominate STAY, because the Spec of VP raises to Spec of IP to get case, when I contains the case assigner.

In the positive declaratives in (14), both constraints are satisfied: by a DP in Spec of VP in (14a) and by a DP in Specifier of IP in (14b), where movement has occurred from Specifier of VP to the higher Specifier position, assuming the VP-internal subject hypothesis (see references in Section 2). The reason is that the DPs in (14a) and (14b) are both in case positions, and the clause has a subject. Inclusion of a *do* could not improve the fate of the examples on these or any other constraints, hence the FULL-INT violation that accompanies the presence of *do* is fatal.

- (14) a. John left
b. John will t leave

When a negative is present, however, the situation is disrupted. If *not* is a head, as in Roberts (1992) Williams (1994), it is not one that has case-assigning ability, so if the subject is in its Spec, CASE is violated, (15a). On the other hand, if the DP is in Spec of VP CASE is obeyed but SUBJ is

⁸ I will not discuss emphatic *do* here. It may well be that the general line of explanation will support the proposal of Laka (1990) that emphasis induces a projection, although this projection, like all the others discussed, will not be present in all clauses.

violated (assuming that the Specifier position is an A position, see Haegeman 1992 for an alternative perspective): (15b).

- (15) a. *John not t left
b. *Not John left

Even if *not* is a specifier as in Zanuttini (1991), the conflict still arises, since the subject in (15a) must then be in a higher specifier, and again CASE is violated. (A more precise formulation of SUBJ must be given to eliminate the possibility that *not*, if it is a specifier, might satisfy the constraint. Perhaps SUBJ should require that an argument fill the position.) In contrast, the presence of an intervening adjoined element such as *rarely*, in *John rarely left on time*, does not affect satisfaction of the two constraints. Thus they conflict when negation is present but not otherwise.

The candidate in (16) does succeed in satisfying both SUBJ and CASE but at the cost of a NO-LEX-MVT violation, since the main verb is heading a non-lexical projection. The ranking of NO-LEX-MVT higher than FULL-INT, established above for inversion in interrogatives, thus properly eliminates this candidate.

- (16) *John left not

The conflict caused by *not* is resolved by the presence of *do*, which is both the feature-carrying head and the highest I-related head. in (17), just as *will* is in (14b). Provided that FULL-INT be dominated by both SUBJ and CASE, this candidate will be the optimal one.

- (17) John did not leave

Tableau (T12) displays the analysis for a clause with no meaningful auxiliary. (I have omitted OP-SPEC from the table since it is vacuously satisfied by all candidates.) The tableau assumes that *not* is a head, although as previously mentioned, this is not crucial. It also assumes that subjects are generated VP-internally, and shows the resulting traces, factored out elsewhere in the paper.

(T12) Presence of *do* with negation

Candidates	HD-LFT	NO-LEX-MVT	OB-HD	SUBJ	CASE	FULL-INT	STAY	HD-RT
[John not [_{VP} t left]]					*!			*
[not [_{VP} John left]]				*!				
 [John did [t not [_{VP} t leave]]]						*		**
[John left[t not [_{VP} t t]]]		*!						***

One instance of *do* makes it possible to satisfy both subject-related constraints: additional instances would constitute fatal violations of FULL-INT, although I do not show these cases. Note the crucial ranking of FULL-INT with both SUBJ and CASE. If FULL-INT dominated both constraints, then (17) could not be the optimal candidate: instead either (15a) or (15b) would be, depending on the ranking between SUBJ and CASE. If FULL-INT dominated just SUBJ, and was dominated by CASE then (15b) would be optimal, since it satisfies CASE and FULL-INT, violating only SUBJ. If FULL-INT dominated just CASE and was dominated by SUBJ, then (15a) would be optimal, since it satisfies SUBJ and not CASE. The ranking between NO-LEX-MVT and FULL-INT is crucial also, as pointed out already.

In the account given here, (15a) is ungrammatical because there is an alternative which better satisfies the constraints, given the constraint rankings for English. This solution contrasts with recent proposals by Pollock (1989) and Chomsky (1991), discussed in Baker (1991), which make appeal to hidden elements or movements. Pollock's account posits a null counterpart of *do*, which must move to T, the movement being blocked when *not* is present. In a similar vein, Chomsky (1991) uses obligatory LF V-to-I raising, triggered by an invisible affixal Q morpheme, but blocked by *not*.

The full system of constraints predicts the distribution of *do* in a negative question: here *do* will appear even when the *wh* phrase is in subject position, for the same reason it appears in negative declaratives, and only one *do* is permitted.

- (18) a. Who did not leave?
b. *Who not left?

Construction of the tableaux for these examples is left for the reader.

Where a clause does not contain a head bearing phi-features, what happens? One answer is that CASE is simply violated in this configuration. This is the case in infinitives, and in subjunctives (thanks to Hubert Haider for drawing the importance of subjunctives to my attention).⁹ The conflict which gives rise to *do* support in (17) is not in effect, with the result illustrated for subjunctives in (19).

- (19) a. I insist that John not leave
b. *I insist that not John left
c. *I insist that John do not leave
d. *I insist that John leave not

The extended projection in (19a), with the subject in specifier of *not*, satisfies SUBJ and violates CASE. Since none of the other candidates can satisfy CASE either, and they all do worse on one of the other relevant constraints, (19a) is optimal, and the variant which includes *do* is impossible.

⁹ Imperatives, such as *Do not collect \$200*, do not pattern with subjunctives and infinitives, although the reason is unclear, having to do perhaps with their morphology (note that they take finite tag questions) or the missing subject.

(Calculations of STAY violations are omitted, as they involve many irrelevant commitments and have no effect on the outcome here).

(T13) Negation in Subjunctives

Candidates	NO-LEX-MVT	OB-HD	SUBJ	CASE	FULL-INT	STAY
☞ [IP <i>John</i> [not [VP t leave]]]				*		?
[not [VP <i>John</i> leave]]			*!	*		?
[IP <i>John do</i> [t not [VP t leave]]]				*	*!	?
[IP <i>John</i> leave [t not [VP t t]]]	*!			*		?

Thus the different behavior of *not* in tensed and non-tensed clauses follows from the effect of CASE. Nothing need be said about *not* itself. This contrasts with the analysis given in Williams (1994), where *not* has an arbitrary [-tense]specification, and hence does not take a tensed complement.

Under this proposal, the nature of the problem solved by “*do*-support” with negation is a little different from the nature of the problem solved by “*do*-support” in interrogatives. In the interrogative case the role of *do* is to fill an otherwise empty head. In the negation case the role of *do* is to provide a head that is structurally higher than *not* and that agrees with the subject. More precisely, the occurrence of *do* in inversion depends on the ranking of OB-HD with FULL-INT. The occurrence of *do* with negation depends on the ranking of CASE and SUBJECT with FULL-INT. But both depend on the ranking of NO-LEX-MVT and FULL-INT. This partial separation receives indirect support from the historical development of English: Roberts (1992) notes that the use of *do* in interrogatives pre-dates its occurrence in negatives. This suggests that the two instances of *do* support represent the same solution to a different problem, rather than the same solution to the same problem. In any case, it is no accident that the same verb appears in both circumstances: the verb which minimally violates FULL-INT.

In this analysis, then, the finite auxiliary is freely generated in any head of the verbal extended projection, its distribution being reined in by a set of constraints. One consequence of this is that there is no movement involved even in cases where the auxiliary precedes *not*. This will later turn out to be important: we will see that the effects of raising an auxiliary verb to C are detectable from their interaction with the Projection Principle, while there is no such effect of the presence of negation.

4. Adjunction, Heads and the Projection Principle

As is well known, inversion in interrogatives is limited to main clauses in most varieties of English (see McCloskey (1992) for analysis of inversion in Hibernian English).

- (20) a. They found out when they should take the train.
 b. *They found out when should they take the train.

This simple fact is initially quite puzzling. OP-SPEC will be violated unless the *wh* phrase is in a Specifier in subordinate interrogatives, for just the same reason as in matrix interrogatives. Yet if this is correct, OB-HD is regularly violated in examples like (20a). I will develop a solution here which attributes the difference between matrix and subordinate interrogatives to a constraint which conflicts with OB-HD, namely the Projection Principle. The constraints conflict because filling the head position by head movement would violate PROJ-PRIN. Since PROJ-PRIN is the dominant constraint, and since it is not possible to satisfy both PROJ-PRIN and OB-HD, the structure with no head is well-formed. It is the optimal structure, hence it is the only one possible.

4.1 The Projection Principle

PROJ-PRIN is loosely related to the principle proposed in Chomsky (1981), and it prohibits movement into the head of, and adjunction to, a subordinate clause. This is a development of two proposals in the literature. The first (Rizzi and Roberts (1989)), is that the root nature of certain head movements (see Emonds 1975) follows from the PROJ-PRIN. In particular, they propose that head movement which is direct substitution is disallowed in selected contexts. The second is the argument in McCloskey (1992) that configurations in which inversion is ruled out seem to be systematically related to configurations in which adjunction is disallowed, the correspondence being particularly clear in the case of arguments. The constraint proposed by McCloskey, based on that in Chomsky (1986), states: "Adjunction to a phrase which is s-selected by a lexical head is ungrammatical". The contrast in (21) from McCloskey (1992), illustrates the motivation for the constraint.

- (21) a. It's probable that in general/most of the time he understands what is going on.
 b. *It's probable in general/most of the time that he understands what is going on.

(Note also the observation made by Rochemont (1989) that topicalization adjoins to either CP or IP in a matrix, but only to IP in a subordinate clause).

Both of these proposed constraints make specific reference to selection. That is, they take the primary cut to be between selected arguments on the one hand, and matrix and adjunct clauses on the other (see also Kayne (1982, 1983) and den Besten (1983)). This split was followed in Grimshaw (1993): there is, however, a problem with such a formulation. McCloskey shows that adjunction to relative clauses and adjuncts is not possible, despite the fact that they are not selected.

- (22) a. *The people [when you get home [who want to talk to you right away]]...
 b. *I graduated [while at college [without having really learned anything]]]

This suggests that all non-root clauses, including adjuncts and relatives, are subject to the constraint. Moreover, as Rizzi and Roberts (1989) note, inversion is impossible in relative clauses, again suggesting that all subordinate clauses behave in the same way. This is supported by patterns such as the one displayed by temporal adjuncts and relative clauses introduced by *when*, illustrated in (23). They behave exactly like the corresponding indirect questions in (23c), and unlike the corresponding matrix questions in (23d).

- (23) a. I left when **he had** / ***had he** arrived.
 b. The day when **he had** / ***had he** arrived ...
 c. I found out when **he had** / ***had he** arrived
 d. When ***he had** / **had he** arrived?

If PROJ-PRIN governs adjuncts, then both patterns have the same explanation: in each case inversion violates PROJ-PRIN and since there is no movement-inducing constraint dominating PROJ-PRIN the optimal forms violate OB-HD.

Examples with *whether*, raised by Armin Mester and Andrew Radford (p.c.), make the same point: there is no inversion in adjuncts introduced by *whether*, just like complements.

- (24) a. Whether we can agree or not, we have to make a decision
 b. *Whether can we agree or not, we have to make a decision

The most attractive position is obviously that the absence of inversion in all of these cases reflects a general property of adjuncts and complements, i.e. of subordinate clauses. Hence the Projection Principle must ban adjunction and head movement for all subordinate clauses:

PROJ-PRIN: No adjunction to subordinate clauses and no movement into the head of a subordinate clause.

I will leave a number of questions open here, including whether head movement is substitution (as Rizzi and Roberts argue for I to C movement) or adjunction. Also, it is possible that PROJ-PRIN is properly understood as two constraints, one on adjunction and one on head movement, which both regulate subordinate clauses. In this case, it will be possible to rank them separately. For English I see no reason to separate them so I will continue to treat PROJ-PRIN as a single constraint.

It is important that the definition of “clause” and “head of a clause” be made clear, as it will be crucial to understanding the effect of PROJ-PRIN. A clause is the highest extended projection of V: i.e. if a CP is the top projection then it is the clause, and other extended projections of V, such as XP, IP, etc. are not clauses, merely parts thereof. The head of a clause is the head of the highest extended projection of V. i.e. C is the head if CP is highest, I is the head if there is no CP etc. PROJ-PRIN is

therefore violated by head movement or adjunction to complements to lexical heads, relative clauses and adjuncts, which are all subordinate clauses. It is not violated by head movement in matrix clauses, or in complements to functional heads, since the former are not subordinate and the latter are not clauses.

This leaves us with one problem: if adjunction and head movement are both ruled out in adjuncts, including relative clauses, why is head movement admitted in one kind of adjunct clause, namely conditionals, even though adjunction is disallowed?

- (25) a. Had I learned anything at college, I would be better off now
 b. *While at college had I learned anything, I would be better off now.

This is a dilemma which can be dissolved under OT, if PROJ-PRIN is violated in (25a) due to the effects of a higher ranked constraint. I will suggest in Section 6 that this is in fact the case here.

4.2 The Projection Principle and Subordinate Interrogatives

PROJ-PRIN is irrelevant for matrix clauses, clearly, in fact it is vacuously satisfied in all of the cases we have looked at so far, and has been omitted from tableaux. Subordinate interrogatives are subordinate clauses in the sense relevant for PROJ-PRIN however. Let us for now consider just the CP version of the complement, with *wh*-fronting. These candidates satisfy OP-SPEC at the cost of a STAY violation. With respect to the other constraints there are two possibilities: if inversion occurs OB-HD will be satisfied (at the cost of a STAY violation), but PROJ-PRIN will be violated, since the CP is a subordinate clause. If inversion does not occur, OB-HD will be violated but PROJ-PRIN will be satisfied (and STAY will be violated only once, although this is not relevant to determining the optimal form). Since the clause with no inversion is the grammatical one, we conclude that PROJ-PRIN >> OB-HD, hence the optimal form is one which satisfies PROJ-PRIN and OP-SPEC but violates OB-HD. Ranking PROJ-PRIN above OB-HD thus derives the absence of inversion in subordinate interrogatives.

(T14) Subordinate Interrogatives — Candidates which Satisfy OP-SPEC

Candidates	PROJ-PRIN	OP-SPEC	CASE	OB-HD	SUBJ	FULL-INT	STAY
$[_{CP} \textit{wh will}_i [_{IP} \textit{DP e}_i [_{VP} \textit{V t}]]]$	*!						**
$[_{CP} \textit{wh e} [_{IP} \textit{DP will} [_{VP} \textit{t V t}]]]$				*			*
$[_{IP} \textit{wh will} [_{VP} \textit{DP V t}]]]$			*!				

There is another candidate to be ruled out here, included as the final candidate in (T14). Here the auxiliary has been generated in C, rather than moving there. Since this strategy evades a violation

of PROJ-PRIN, the candidate threatens the success of the actual winner. However, CASE is violated here, since the DP is never in a Specifier-head relationship with the case-marking head. Thus the final candidate is eliminated. (Having made this point, I will set such candidates aside, and omit SUBJ and CASE from tableaux in this section, along with NO-LEX-MVT.) If the picture is as in (T14), then CASE >> OB-HD. If SUBJ is also violated by the final candidate, which depends on the exact formulation of the constraint, we can conclude only that at least one of CASE and SUBJ dominates OB-HD.

Some alternative candidates which do not satisfy OP-SPEC, are illustrated in (T15). A CP with no *wh* movement but with inversion violates PROJ-PRIN, which guarantees that it will lose to one of the other candidates. A CP with no *wh* movement and no inversion, the last candidate in (T15), satisfies PROJ-PRIN, but at the cost of a violation of OP-SPEC and OB-HD, so it too compares unfavorably with the optimal candidate in the first row. An informative comparison is between the optimal candidate and the last candidate: an IP with no *wh* movement. This extended projection necessarily violates OP-SPEC, but it satisfies all the other constraints. Comparison of this alternative with the grammatical one reveals that there is a crucial ranking between OP-SPEC and OB-HD: OP-SPEC >> OB-HD. Otherwise the IP structure would be optimal, since it violates OP-SPEC but not OB-HD.

(T15) Candidates which violate OP-SPEC, compared to the optimal candidate

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
$\text{[}_{\text{CP}} \text{wh e [}_{\text{IP}} \text{DP will [}_{\text{VP}} \text{V t]]]}$			*		*
$\text{[}_{\text{CP}} \text{will}_i \text{ [}_{\text{IP}} \text{DP e}_i \text{ [}_{\text{VP}} \text{V wh]]}$	*!	*			*
$\text{[}_{\text{IP}} \text{DP will [}_{\text{VP}} \text{V wh]}$		*!			
$\text{[}_{\text{CP}} \text{e [}_{\text{IP}} \text{DP will [}_{\text{VP}} \text{V wh]]}$		*!	*		

We can now see why *do* never occurs in subordinate interrogatives, given the analysis of *do* from Section 3.

- (26) a. *I don't know what did she say
 b. *I don't know what she did say
 c. I don't know what she said

When no auxiliary verb is present and there is no inversion, a subordinate interrogative violates OB-HD because the head of the projection housing the *wh* operator is empty. Including a *do* adds a FULL-INT violation, and if it inverts, a PROJ-PRIN violation. Since OB-HD is violated in subordinate interrogatives in all candidates which respect PROJ-PRIN, and since the only virtue of *do* is that it can satisfy OB-HD, the presence of a *do* can only add violations in this situation, it can never reduce them. Hence there can be no advantage to the presence of *do*, hence it is impossible.

(T16) Subordinate Interrogatives with *do*

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
$[_{CP} \textit{wh} \textit{do}_i [_{IP} \textit{DP} \textit{e}_i [_{VP} \textit{V} \textit{t}]]]$	* !			*	**
 $[_{CP} \textit{wh} \textit{e} [_{VP} \textit{DP} \textit{V} \textit{t}]]$			*		*
$[_{CP} \textit{wh} \textit{e} [_{IP} \textit{DP} \textit{do} [_{VP} \textit{V} \textit{t}]]]$			*	* !	*

If *if* and *whether* are heads, then both inversion and the appearance of *do* will be ruled out in subordinate yes-no questions because the C position will be filled by a complementizer. As a result, the *if* and *whether* forms satisfy OB-HD, without inversion or the appearance of *do*. (T17) shows the optimal candidate, for an input which includes an auxiliary, under this analysis. Since inversion violates STAY, and *do* violates FULL-INT, it is clear that the inverted and *do* forms will always lose to the optimal ones.

(T17) Subordinate Interrogatives with *if* and *whether* as heads

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
 $[_{CP} \textit{if/whether} [_{IP} \textit{DP} \textit{will} [_{VP} \textit{V}]]]$					

- (27)
- a. They asked if/whether he will leave
 - b. *They asked if/whether will he leave
 - c. They asked if/whether he left
 - d. *They asked if/whether he did leave

If, on the other hand, *whether* is a Spec of CP, as in Kayne (1991), then the analysis for clauses with *whether* is essentially the same as for other subordinate interrogatives, except for the fact that no *wh* movement is involved, hence there is no STAY violation. Comparison of the optimal *if* clause in (T17) and the optimal *whether* clause in the Specifier analysis, shows that under Kayne's proposal, the two clause types must not be competitors. If they were then the *if* variant would be optimal because it does not violate OB-HD. Perhaps the structural difference between them is sufficient to guarantee that their LFs are distinct. However, another reason why they might not be in the same

candidate set is the analysis suggested by Donca Steriade (p.c.) in which the *if* clause is really a kind of adjunct. This is supported by Steriade's observation that *if* clauses occur naturally only with verbs that allow Null Complement Anaphora in the sense of Grimshaw (1979).

So far, then, three questions concerning inversion patterns have been addressed: why there is no inversion in matrix declaratives, why there is inversion in matrix interrogatives and why matrix and subordinate interrogative clauses should show different inversion patterns. OB-HD makes heads obligatory except where PROJ-PRIN makes them impossible. The fact that an empty head is possible is a side effect of the fact that movement is not. Since PROJ-PRIN does not affect matrix clauses, a head can be empty only in a complement clause or adjunct.

4.3 Some Alternatives

The essential property of the solution for inversion is that each component principle is fully general: none of the principles is specific to interrogatives or to inversion, for example. In fact, there is no theory of inversion; it is just the result of OB-HD, whose effects are seen whenever the effects of PROJ-PRIN do not obscure them. Conflict between general constraints, and the resolution of the conflict, lies behind the observed patterns.

Consider an alternative to the constraint conflict proposal for the absence of preposing in subordinate interrogatives, namely that there is a null C, or a C filled by +wh, in subordinate interrogatives. Now it is necessary to distinguish in a principled way between the empty head in this case, which by hypothesis would be filled by a null element, and other empty heads, such as the one in a matrix interrogative (and others to be analyzed in Section 8), which cannot be filled by a null complementizer. One might take the position that only selected heads can be null. This would run into serious empirical problems with adjuncts (see 4.1), and *that* complements (see 8). But it is worth dwelling on, because it has a revealing property — it builds into the principle governing empty heads the effects of PROJ-PRIN. The real situation is that heads can be empty in exactly the case where PROJ-PRIN will not allow them to be filled. This is a direct consequence of constraint conflict but not of an empty heads solution, where elaboration of the constraint, reflecting the effects of conflict must be stipulated. A useful example is the following:

Specifier Licensing Condition (Plunkett 1990, 128)

If a maximal projection is in a non-subcategorized position, its specifier may not be filled at s-structure unless its head position has also been filled by that time.

This is an accurate description of the empirical situation, setting aside adjuncts; examining each part of the condition reveals that it states the effects of the interaction of the relevant constraints. The highly specific character of this kind of solution entails that it cannot extend over the range of cases which follow from the optimality theoretic proposal: this will be particularly clear in the case of obligatory complementizers (Section 8). It is, moreover, inevitably language particular (see Sect. 7.)

Similar points hold with respect to the recent proposal for inversion developed by Rizzi (1991) Haegeman (1992). It uses the idea that a head with a certain feature has to raise in cases of inversion

in order to participate in a Specifier-head relationship with a Specifier with the same feature, to meet a well-formedness principle, called the “*Wh* Criterion” for the interrogative cases. An important insight in this work, that the relationship of Specifiers and heads lies behind inversion, is incorporated into the present proposal.

A large part of the work done by the *Wh* Criterion and related principles results from an auxiliary hypothesis concerning the initial distribution of the feature: it is on I at d-structure in matrix clauses, hence it must raise to C, and it is on C already in subordinate clauses, hence inversion is not required to put it in the right relationship to Specifier of CP. What is the result of constraint interaction in the present paper is instead the result of feature distribution in the Rizzi-Haegeman approach. From the present perspective, the *Wh* Criterion, like the Specifier Licensing Condition, builds in the effects of the independently existing PROJ-PRIN. As expected, this is accompanied by significant loss of generality. For example, there is no relationship between the explanation for inversion and the explanation for the obligatoriness of *that* discussed below: inversion is necessary to get features in the right place, *that* is necessary for reasons of selection, as touched on in Section 8.

The constraint conflict proposal is more general in another way — it predicts that any constraint that forces an element to occur in Spec will have the effect of inducing inversion. English offers a number of other instances of inversion, some of which do not seem to be insightfully subsumed under a feature-based account. These inversions have the same distribution as inversion with negative preposing, analyzed in Section 5.)

- (28) a. So wealthy will he become that
 b. I will be rich and so will you
 c. I don't like coffee and neither/nor does Bill.
 d. Only under these circumstances will you be able to win

Although of course possible, it seems unlikely that there is a criterion governing Specifier-head relationships for all of these, with a feature generated on I in matrix clauses and so forth. Rather it seems that the patterns in (28) reflect the existence of a set of expressions which must occur in specifier. Inversion is simply a structural consequence.

Maximally general principles will inevitably conflict. The alternative is to formulate more specific principles which are designed never to conflict, and one price is generality. Only by allowing constraints to conflict can we avoid building the effects of every principle into all of the others that it potentially conflicts with. There is another price — universality. These points will be developed further in Section 7.

5. Inversion inside a subordinate clause

We know from the previous discussion that OB-HD will induce inversion wherever PROJ-PRIN does not prevent it. PROJ-PRIN is violated by inversion into the highest head of a subordinate clause, i.e. into the highest head of the extended projection of V. It is not violated, however, by inversion

into other heads of the extended projection, since the projections they head are not subordinate clauses, merely parts of clauses. So inversion should be possible when the relevant head is not the highest head of a subordinate clause, but is contained within the extended projection. This prediction is verified by the pattern of inversion which accompanies preposing of a negative.

If a negative operator is preposed, inversion is required, as illustrated in (29) (Klima 1964, Liberman 1974).

- (29) a. Never/under no circumstances will she work this hard again
 b. *Never/under no circumstances she will work this hard again

In the absence of preposing, inversion is not allowed (see (30)).

- (30) a. She will never work this hard
 b. *Will she never work this hard

This paradigm follows the same pattern as interrogatives: negative operators occur in specifier position, so a projection is present when preposing occurs, which is otherwise absent. The head of this projection is empty. Hence head movement must occur to fill the head, and inversion follows.

What is the projection that is present when negative preposing occurs? It cannot be CP since the negative element follows C in subordinate clauses, where the entire paradigm can be replicated.

- (31) a. She said that never/under no circumstances would she work this hard again
 b. *She said that never/under no circumstances she would work this hard again
 c. She said that she would never work this hard again
 d. *She said that would she never work this hard again

Nor can the projection be IP since the Specifier of IP is already filled by the subject. Thus the projection must be a further member of the verbal extended projection, which intervenes between IP and CP, and which I label "XP". The relative position of the *wh* phrase and preposed negative follows from the nature of the operator. *Wh* movement is type changing, so the *wh* operator must be outside everything pertaining to the propositional structure. Negative preposing on the other hand, is a variety of sentential negation, hence the operator c-commands IP but is c-commanded by C. There is no need, then, to stipulate which specifier position each operator appears in.¹⁰

¹⁰ Teun Hoekstra (p.c.) raises the question of whether *wh* movement and negative preposing can co-occur, in examples like *Which book will never in her life Mary read?* Such an example presumably violates Relativized Minimality (Rizzi 1990), and certainly has a highly marginal status. It seems clear, however, that this version comes closest to well-formedness -- positioning the auxiliary after the preposed negative seems worse, which is what the proposal made here would predict. Exchanging the position of the *wh* phrase and negative gives clear ungrammaticality.

The constraints discussed so far dictate that the structure of the matrix clause in (29a) is (32a), with no CP present just as for matrix declaratives, while the structure of the complement in (31a) is (32b).

- (32) a. $[_{XP} \text{ never/under no circumstances... } X [_{IP} DP I [_{VP} V \dots t]]]$
 b. $[_{CP} C [_{XP} \text{ never/under no circumstances... } X [_{IP} DP I [_{VP} V \dots t]]]]$

The discussion here is simplified by proceeding as if all negative phrase operators are subject to OP-SPEC, and therefore must move. This is approximately true for some, presumably those for which sentential scope is the only possibility:

- (33) a. *She would do this under no circumstances
 b. Under no circumstances would she do this

It is not true for, e.g. *never*. I assume that this reflects the fact that such negatives can take scope from more than one position, unlike *wh* phrases.

(T18) shows how inversion is induced by preposing. If the matrix is an IP then there will be no possibility for preposing in the first place, so OP-SPEC will be violated. If the matrix is an XP there are four possibilities: no inversion and no preposing will violate OB-HD and OP-SPEC, preposing without inversion will violate OB-HD, inversion without preposing will violate OP-SPEC but preposing with inversion will violate nothing other than STAY twice. Since STAY is ranked below both OP-SPEC and OB-HD, as we already know, the form which violates only STAY is optimal, and grammatical. In this way the constraints and rankings predict that preposing “induces” inversion, and inversion is impossible without preposing.

(T18) Negative-Induced Inversion in Matrix Clauses: “*u. no circs.*” stands for “*under no circumstances*”.

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
$[_{IP} DP \text{ will } [_{VP} V \text{ u. no circs.}]]$		*!			
$[_{XP} e [_{IP} DP \text{ will } [_{VP} V \text{ u. no circs.}]]]$		*!	*		
$[_{XP} \text{ u. no circs. } e [_{IP} DP \text{ will } [_{VP} V t]]]$			*!		*
$\text{☞} [_{XP} \text{ u. no circs. will}_i [_{IP} DP e_i [_{VP} V t]]]$					**
$[_{XP} \text{ will}_i [_{IP} DP e_i [_{VP} V \text{ u. no circs.}]]]$		*!			*

(31) shows that inversion with negative preposing does not show a matrix-subordinate contrast, unlike inversion in interrogatives. The reason is that PROJ-PRIN is vacuously satisfied by inversion into X — the XP is not a subordinate clause, hence inversion into it does not violate PROJ-PRIN. The subordinate clause in (31a) is the CP headed by *that*, and not the XP it contains, since CP is the highest extended projection of V. (We will consider what happens when CP is absent in Section 8). PROJ-PRIN is violated by head movement into the complement of a lexical head, but not by movement into the complement of a functional head; this follows from the fact that complements to functional heads are not clauses. XP embedded within a CP is like a matrix clause in the critical respect, and both therefore require inversion.

(T19) Negative-Induced Inversion in Subordinate Clauses

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
$[_{CP} \text{ that } [_{IP} \text{ DP will } [_{VP} \text{ V } u. \text{ no } \text{ circs.}]]]$		*!			
$[_{CP} \text{ that } [_{XP} \text{ e } [_{IP} \text{ DP will } [_{VP} \text{ V } u. \text{ no } \text{ circs.}]]]]]$		*!	*		
$[_{CP} \text{ that } [_{XP} \text{ } u. \text{ no } \text{ circs. } \text{ e } [_{IP} \text{ DP will } [_{VP} \text{ V } t]]]]$			*!		*
$[_{CP} \text{ that } [_{XP} \text{ } u. \text{ no } \text{ circs. } \text{ will } [_{IP} \text{ DP } t [_{VP} \text{ V } t]]]]$					**
$[_{CP} \text{ that } [_{XP} \text{ will}_i [_{IP} \text{ DP } e_i [_{VP} \text{ V } u. \text{ no } \text{ circs}]]]]]$		*!			*

The pattern of constraint violation and satisfaction is exactly the same for negative preposing in a subordinate clause and in a root clause, as can be seen from a comparison of (T19) with (T18). By the same token, (vacuous) satisfaction of PROJ-PRIN admits adjunction to VP (when dominated by IP) and to IP (when dominated by CP), as can be seen in (34), based on McCloskey (1992), where this observation is made.

(34) They announced that at Christmas time $[_{IP}$ the president has generally $[_{VP}$ gone to visit his mother.

The structure assigned to a clause is entirely determined by the constraints. Hence it is not possible for the analyst to simply declare that some clause has a certain structure, without providing the system of constraints which will guarantee that desired result. In the present case, the constraints require that the XP intermediate verbal projection be omitted except when preposing to specifier occurs. An XP which is empty will always violate OB-HD, and moving an auxiliary to fill the head X position will always violate STAY. Thus the optimal representation of a clause will not include XP unless some constraint higher ranked than STAY can be satisfied by virtue of its presence. As a consequence the structure of a subordinate clause with a preposed negative is different from that of one with no preposing. One has the extra XP and the other does not. This is crucial to inversion of course: when the XP is present inversion must occur, when the XP is absent inversion is not possible.

6. Inversion in Adjuncts

PROJ-PRIN is violated by inversion into the head of any subordinate clause, including an adjunct. Hence adjuncts systematically resist XP adjunction and inversion. However, as mentioned above, inversion is possible in conditional adjuncts, which show the pattern illustrated in (35)-(36), recently discussed in Rizzi and Roberts (1989), Pesetsky (1989) Iatridou and Embick (1994).

- (35) a. Had I been on time I would have caught the train
 b. Were he to be asked, he would probably say no.
 c. Should it ever happen, you will be sorry.
- (36) a. *I had been on time I would have caught the train
 b. *He were to be asked, he would probably say no.
 c. *It should ever happen, you will be sorry.

These examples show that a higher-ranked constraint can force violation of PROJ-PRIN. In support of this, note that I-to-C movement is possible in complements sometimes also, just not in English. In the Aux-to-Comp process studied in Rizzi (1983, Ch3) and Raposo (1987), an Aux raises to C to assign case to an otherwise caseless subject. Here too PROJ-PRIN is overridden by another constraint, in a complement this time.

There are two properties of conditionals which might form the basis of the relevant constraint: the morphology of the auxiliary, and its semantics. Unlike conditionals introduced by *if*, which allow any auxiliary to occur, inverted conditionals admit only certain auxiliaries, those in (35), which never occur in this form in declaratives (apart from *should* in its root meaning).¹¹ These auxiliaries, then, are not independent elements, but require a connection to the consequent. COND requires a dependent head to c-command the extended projection containing it. The idea is that the inverting auxiliaries are semantic and/or morphological dependents, and thus must be locally accessible to the supporting main clause, and hence at the top of the extended projection that houses them. COND dominates PROJ-PRIN, hence the inverted form is grammatical despite the violation of PROJ-PRIN.

¹¹ The auxiliary *do* will never invert, because it has neither the semantics nor the morphology to be subject to COND. Pesetsky (1989) attributes absence of inversion with *do* to its language particular status, which makes it inert. In the present system, *do* is most certainly not inert, since OB-HD induces movement of *do* just like any other auxiliary. Indeed, *do* is not so very language particular either, given the proposal in 3.3.

(T20) Inversion in Conditional Adjuncts

Candidates	COND	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
$[_{IP} DP \text{ had } [_{VP} V \dots]]$	*!					
 $[_{XP} \text{ hadi } [_{IP} DP e_i \text{ } [_{VP} V \dots]]]$		*				*

In a clause introduced by the conditional element *if*, there is no dependency of relevance to COND, and the constraint is satisfied without inversion. In this circumstance, inversion will violate STAY and will not improve the status of the extended projection on any other constraint, hence the ungrammaticality of examples like (37b).

- (37) a. If he were to be asked, he would probably say no
 b. *If were he to be asked, he would probably say no.

It is important that the *if* and inversion conditionals are not in the same candidate set, given that inversion conditionals always violate PROJ-PRIN and STAY where *if* conditionals respect both. The inversion conditionals would always be ungrammatical if the two were in competition. The issue arises in any theory with an economy constraints; see Iatridou and Embeck (1994) for discussion of this point, and an argument that the two kinds of conditional have different logical forms.

In sum, inversion in conditionals is not motivated by OB-HD, which is dominated by PROJ-PRIN in English, but by a constraint which dominates PROJ-PRIN. Hence conditionals provide the lone example of inversion in a subordinate clause in the language.

7. Typological Consequences of Constraint Ranking

As Prince and Smolensky (1993) show, positing universal constraints subject to ranking by individual grammars offers a theory of language typology, with often striking predictions. (See also Legendre et al (1993) for an optimality theoretic typology of case systems.) It is not strictly-speaking possible to determine the effects of alternative rankings of constraints without knowing what all the constraints of UG are. Nonetheless it is revealing to examine re-rankings, and in particular to compare re-ranking with parametric accounts of variation.

As an illustration of the systems generated with alternative rankings, consider the interaction of STAY with OB-HD and OP-SPEC, crucially ranked in English with STAY dominated by both OB-HD and OP-SPEC. The six possible rankings of these three constraints are illustrated in (38).

(38)

a.	OP-SPEC	OB-HD	STAY
b.	OB-HD	OP-SPEC	STAY
c.	STAY	OP-SPEC	OB-HD
d.	STAY	OB-HD	OP-SPEC
e.	OB-HD	STAY	OP-SPEC
f.	OP-SPEC	STAY	OB-HD

(38a) shows the English option, giving *wh* movement and inversion. Both (38c) and (38d) yield systems with neither *wh* movement nor inversion, since STAY suppresses the effects of both OB-HD and OP-SPEC. (38e) in fact yields the same result: when OB-HD dominates STAY and STAY dominates OP-SPEC, the effect is the same as when STAY dominates both. This is because the ranking of OP-SPEC relative to STAY will prohibit *wh* movement, hence there will be no CP present and no empty C to fill. Thus rankings (38c,d,e) all give a system with no movement.¹² (38f) corresponds to a system which has *wh* movement but no inversion. These all seem to be natural possibilities.

One a priori possible language type is characterized as impossible, namely one in which there is no general inversion process, yet there is inversion, but not *wh* movement, in interrogatives. This is ruled out by the reasoning governing (38e): without XP movement to Spec there is no head to be filled.

Finally, the ranking in (38b) gives a system in which both *wh* movement and inversion occur in a matrix clause, and also in a subordinate clause if OB-HD dominates PROJ-PRIN. If, however, PROJ-PRIN outranks OB-HD, in subordinate clauses there will be no *wh* movement (Compare the first and third candidates in (T15)). Whether this corresponds to a possible language, or whether some other constraint interferes here, perhaps concerning selection, is a question that I leave open.

Re-ranking constraints gives typological effects of the kinds treated as parametric in other perspectives. Within OT, the constraints are universal. Whether the effects of a constraint are visible in a given language depends on the constraint rankings. In contrast, if only inviolable constraints are admitted into the theory, then some notion like a parameter is essential, since there is no set of non-parameterized constraints such that every language satisfies them. Consider for example, the descriptive generalization cited in 4.3 concerning the distribution of inversion in English, which, as noted in 4.3, is a non-violable counterpart, in some sense, of the proposed OT constraint system.

¹² If Japanese and Korean “*wh*-phrases” are, as first argued in Kim (1989) are simply QPs with no *wh* properties then the absence of movement might be properly attributed to the irrelevance of OP-SPEC.

It is accurate for English, but it is not true of all languages, and it cannot be a universal constraint, but must be parametric. Hence the need for parameters — they accommodate language variation in a system of inviolable constraints. In contrast, when such generalizations are understood, not as part of the grammar of any language but rather as descriptions of the state of affairs that results from the constraints as ranked in a particular grammar, the constraints are seen to be universal, and not parameterized.

There is a systematic relationship between the constraint rankings in an optimality theoretic grammar and the formulation of the consequences of these rankings as an unviolable constraint, varying parametrically. If we reformulate a violable constraint C_v , as an inviolable constraint C_i , what happens is this. If C_v is undominated, no change is required. However, if C_v is dominated, the reformulation will have to build into C_i the effects of its interaction with every dominating constraint. The more numerous the constraints that crucially dominate C_v , the more complicated the formulation of C_i will have to be. Needless to say, the form that C_i eventually takes will vary cross-linguistically, since C_v stands in different dominance relations in different grammars.

Variation determined by constraint interaction entails the existence of an entire set of grammars: there is no way to simply eliminate a system by stipulation. Given the constraints, each grammar is inevitable, as illustrated by (38). Of course the constraints could be wrong or other constraints could be at play, but any change in the posited constraints, or the addition of new ones, will itself make further predictions about possible re-rankings. For this reason, there is no way to surgically remove exactly those systems which we wish to dispose of. This is not true of a parametric system. For the sake of illustration, suppose we posit a parametric system using inviolable constraints to replace the one posited here. We could formulate it like (39).

- (39) Matrix/subordinate interrogatives have *wh* movement/no movement
 Matrix/subordinate interrogatives have inversion/no inversion

This system allows the problematic system yielded by ranking (38b) above. It also allows the grammar which re-ranking of the universal constraints makes impossible — inversion without *wh* movement. But the point is that it can easily be reformulated, in any way we choose. For instance, we could write in a dependency between the specification concerning *wh* movement and the specification concerning inversion. Any other arbitrary dependency could equally well be established. The reformulation may involve complication, but it is always possible. In the limit we can just list the alternative systems, call the list a "parameter", and its members "values" of the parameter. It is arbitrary what appears on the list and what does not, and mere descriptive convenience is the driving force. In this sense, parametric values are isolated from each other, interacting constraints are not. Re-ranking has inevitable consequences for the entire grammatical system in a way that re-writing parameters does not. Re-ranking of universal constraints clearly promises a more illuminating theory of typological variation.

8. Head Position and Subordinate Interrogatives

The interaction of OB-HD and PROJ-PRIN explains why inversion does not occur in subordinate interrogatives. It does not, however, explain why the optimal candidates have no complementizer, rather than some element such as *that*. Such a representation would satisfy both OB-HD and PROJ-PRIN. In Grimshaw (1993) it was assumed that the answer lies in the lexicon of English: *that* is a *-wh* complementizer, hence cannot occur in a Specifier-head relationship with a *+wh* expression, and English offers no alternatives which are compatible both semantically and syntactically with an interrogative structure. However, it seems that despite the apparently parochial character of this puzzle, the answer nonetheless resides in UG.

In order to develop this more interesting alternative, we will need to build a picture of how X'-theory can be treated under OT assumptions, in particular a theory of how heads are positioned. Travis (1989) discusses two cases where a head V is not uniformly initial or final. In Chinese, according to her analysis, the V precedes arguments but follows adjuncts. This follows if there is default head-final, overridden by directionality of theta-marking, which requires theta-marking to the right. In Kpelle, the default is head-initial, and it is overridden by directionality of case-marking, which is leftward. Hence NP complements precede V while PPs follow. (Note that this solution is not coherent under standard assumptions, since it crucially relies on notions like "default" and "override" which have no place in a system of inviolable constraints.) In OT terms, in Chinese Theta-Right >> Head-Right (and both these constraints dominate all others concerning theta-marking, case-marking and head position).¹³ In Kpelle, Case-Left >> Head-Left (and these constraints dominate all relevant others, as for Chinese.) Thus in Chinese and Kpelle head-position constraints, HD-LFT and HD-RT, conflict with other constraints affecting position, and when the other constraints dominate, the pattern of head position is perturbed as a result. These constraints, which require that the head of every projection be leftmost/rightmost, are alignment constraints (Prince and Smolensky 1993, McCarthy and Prince 1993).

English shows a different kind of mixture: it is head final at the XP level and head initial at the X' level. This too is the result of the head position constraints in interaction with others. Suppose that in English HD-LFT >> HD-RT. (If all structure is binary then HD-RT is satisfied whenever HD-LFT is violated). HD-LFT is violated by subjects (Spec of VP when there is no auxiliary, Spec of IP when there is). This is the effect of a dominant constraint on Specifier positions: SPEC-LFT. In this analysis, then, English is a left-headed language because of HD-LFT, except where other constraints demanding other head configurations intervene. More generally, heads are uniformly left or right within a

¹³ The head-finality of nominals in Chinese (Huang 1982) could be due to a constraint treating N as final, N-RT, outranking THETA-RT. More interestingly, it could be attributed to N not theta-marking, as Travis suggests, in which case THETA-RT would not be violated in a head final NP. The issue is subtle, however, in view of the arguments (Grimshaw 1990) that nouns take obligatory arguments just like verbs when they have an event structure.

language, to the extent that they are able to be so. Apparent mixed systems always arise from constraint conflict.

This sketch of the theory of head position in UG makes possible a simple solution to the residual problem of heads in subordinate interrogatives. The candidates of interest are one with no C (the successful one for English), and the alternative in which C is filled with *that*, or with any other morpheme. The crucial comparison is between the pair in (40).

- (40) a. I wonder when I will see such a sight again.
 b. *I wonder when that I will see such a sight again

Both these sentences, in the analyses in (T21), violate STAY twice, because of *wh* movement and movement of the subject DP. Both also violate HD-LFT twice, because V and I are not leftmost in their XP projections, due to the presence of Specifiers. The critical difference is that the candidate with no *that* violates OB-HD, while the candidate with a *that* satisfies OB-HD but at the cost of an additional HD-LFT violation, in CP. We conclude that HD-LFT dominates OB-HD, hence English chooses to have no C rather than a C in the wrong position. (The opposite ranking of OB-HD and HD-LFT would give a language in which a complementizer occurs obligatorily with interrogatives, as well as with adverbial clauses. Perhaps middle English is close to such a system.)

(T21) omits PROJ-PRIN, OP-SPEC and FULL-INT, all satisfied in the two candidates at issue. There is no crucial ranking of SPEC-RT AND HD-RT with respect to each other or with respect to OB-HD and STAY.

(T21) Subordinate Interrogatives, presence versus absence of *that*

Candidates	SPEC-LFT	HD-LFT	SPEC-RT	HD-RT ¹⁴	OB-HD	STAY
[_{CP} <i>wh that</i> [_{IP} DP will [_{VP} t V t]]]		***!		**** **		**
 [_{CP} <i>wh e</i> [_{IP} DP will [_{VP} t V t]]]		**		****	*	**

Why doesn't the null complementizer in the optimal candidate violate HD-LFT also? If it does, then the two candidates will tie on HD-LFT, and the decision will be made by OB-HD, which will choose the wrong candidate. The answer must be that there is no "null complementizer" present. This conclusion allows us to choose between two alternative conceptions of OB-HD. We might have

¹⁴ Violations of HD-RT (strictly irrelevant here) are calculated on the assumption that a complement to the right of X induces two violations, since X is final neither in X' nor in XP. Alternative formulations of the constraint would give different outcomes, but no evidence here bears on the issue.

taken the position that GEN always includes a head X for a projection of X. Then both the candidates in (40) would have a head, and the OB-HD violation would be due to the emptiness of the head. But this interpretation is inconsistent with the explanation just given here for the absence of a complementizer in interrogative CPs. Thus we conclude that GEN includes a head X for a projection of X only when there is an element filling X. Only the first candidate in (41) has a head C present: the other has no C position at all, just a C' dominating IP. OB-HD, then, is violated by this configuration, i.e. it regulates the presence of X^0 in a projection, and not the filling of X^0 with linguistic material. With this interpretation, the problem of why the null complementizer in the optimal candidate does not violate HD-LFT is trivial: there is no null complementizer, so HD-LFT is (vacuously) satisfied.¹⁵

When an auxiliary verb raises to a higher position, as in matrix questions and with negative preposing, OB-HD is satisfied, but what of HD-LFT? The raised head has a specifier on its left in both cases, so it is apparently not in the optimal head position. Since we know that HD-LFT dominates OB-HD, we are in danger of making the incorrect prediction that raising of an auxiliary verb should never be the best response to an empty head, which would undermine the entire analysis of inversion. This suggests that HD-LFT (and presumably HD-RT) holds only of perfect heads in the sense of Grimshaw (1991), i.e. heads which match the projection in all respects. The relationship of I to IP is that of perfect head, but that of any raised head to the projection it raises into, is not. Hence HD-LFT is not violated by moved heads. This will have no effect on X-bar structure in general, since simple X-bar projections are always perfect projections, hence HD-LFT is always effective here.

We know, then, that unprojected heads and moved heads do not figure in calculating violations of HD-LFT (and HD-RT): the same proves to be true for traces. (T5') is a revision of (T5), which now includes HD-LFT and HD-RT, and also shows the trace left by movement of DP from within VP. How many violations of HD-LFT are there? If traces and unprojected heads left by head movement are exempted from the constraint, the violations are those indicated. All candidates except the last incur one violation, hence the decision passes to OB-HD, which prefers the first candidate, with inversion. Note that if unprojected heads alone were exempted from the constraint, we would get the wrong result. The trace of head movement in the first candidate would violate HD-LEFT, thus incurring two violations of the constraint, which would eliminate the actual optimal candidate from the competition. We conclude that the constraint holds of overt perfect heads.¹⁶

¹⁵ Thanks to Alan Prince for aid in developing this proposal.

¹⁶ The other cases of *do* support -- those involving negation provide us with evidence for a ranking. In (T12) the final candidate (with no violation of HD-LFT) will win unless NO-LEX-MVT dominates HD-LFT, a ranking which is entirely consistent with the system, but which is not further explored here.

(T5') Matrix Interrogatives with and without *do* (revised)

Candidates	OP-SPEC	*HD-LFT	OB-HD	FULL-INT	STAY	HD-RT
$\text{[}_{\text{CP}} \mathbf{wh} \mathbf{do}_i \text{[}_{\text{IP}} \text{DP } \mathbf{e}_i \text{[}_{\text{VP}} \mathbf{t} \mathbf{V} \mathbf{t} \text{]}]]]$		*		*	***	*
$\text{[}_{\text{CP}} \mathbf{wh} \mathbf{e} \text{[}_{\text{IP}} \text{DP } \mathbf{e} \text{[}_{\text{VP}} \mathbf{t} \mathbf{V} \mathbf{t} \text{]}]]]$		*	* ! *		**	*
$\text{[}_{\text{CP}} \mathbf{wh} \mathbf{e} \text{[}_{\text{VP}} \text{DP } \mathbf{V} \mathbf{t} \text{]}]]]$		*	*!		*	*
$\text{[}_{\text{CP}} \mathbf{wh} \mathbf{e} \text{[}_{\text{IP}} \text{DP } \mathbf{do} \text{[}_{\text{VP}} \mathbf{t} \mathbf{V} \mathbf{t} \text{]}]]]$		**	*!	*	**	**

Thus UG provides constraints on head position, in the form of HD-LFT and HD-RT. The ranking of HD-LFT relative to OB-HD in English ultimately explains the absence of complementizers in subordinate interrogatives, and a fact which appears to be highly language specific is derived from the interaction of universal constraints.

Note too, that this is another example where a lexical gap is explained as a function of constraint rankings. English has no word to appear in the head position in subordinate interrogatives. The constraint ranking of English makes it impossible for such a word to be used, thus in effect it makes it impossible for it to exist in the lexicon. The same point has been made previously: no language in which NO-LEX-FUNCT is dominated by both FULL-INT and OB-HD can have a semantically empty use for a verb as in English auxiliary *do* (3.3.) Gaps in the lexicon can be epiphenomena of constraint rankings, and the same is true of the existence of lexical items in particular analyses, if the analysis of *do* in 3.3 is correct. These results are quite surprising, in view of the more standard view (Chomsky 1992) that language variation is due to differences in the lexicon, rather than, as here, that lexical variation may be due to differences in the grammar.

This raises the question of whether *all* principled lexical variation, in particular all cross-linguistic variation in functional categories, might be derived from constraint ranking. Typological differences attributed to “strong” versus “weak” features in work in the minimalist program (Chomsky 1992) for example, can be understood as resulting from ranking of checking constraints and STAY. When STAY outranks a checking constraint there will be no movement, when it is outranked by a checking constraint, movement will occur. Minimally, it seems that re-ranking of violable constraints offers an extremely interesting window on the relationship between lexical and syntactic properties.

9. The Obligatoriness of *that*

9.1 Complements with and without *that*

This paper does not aim to characterize inversion, but to derive the distribution of inversion from the properties of heads in general. I will show here that the obligatory appearance of *that* in

subordinate clauses with topicalization or operator movement to Specifier follows from the principles already laid out.

The constraints allow sentential complements to be VPs, IPs or CPs, or in principle XPs, although as we will see below, this possibility is in fact excluded by the constraints. When a complement is a CP, it must have its head filled, because of OB-HD. Hence a clause which is not introduced by a complementizer cannot be a CP. (Doherty (1993) provides extensive argumentation in favour of analyzing clauses without *that* as IPs.) Similarly, a clause with no auxiliary cannot be an IP. Thus verbs like *think* allow three complement structures, as in (41):

- (41) a. I think [_{CP} that it will rain]
 b. I think [_{IP} it will rain]
 c. I think [_{VP} it rained]

An apparent objection to this analysis is that we have to stipulate that *think* and every verb like it selects VP, IP and CP as a complement. This would not only complicate the lexical representation for all of these verbs, it makes it impossible to explain why there isn't a verb of approximately the same semantics as *think*, which takes just one of the three, since such a verb would be exploiting the simplest possible selection option available.

The same issue arose in connection with the claim that interrogatives with *wh* subjects are VPs or IPs, in Section 6, and the same answer holds here, exploiting Type-Category selection (Grimshaw 1991). Recall that in the theory of extended projection, functional heads do not select at all. Lexical heads do select: they c-select the syntactic category and s-select the semantic type of their complements. All members of the verbal extended projection (C, I, V and whatever other heads participate) are of the same syntactic category (verbal). They differ in their functional analysis, not in their syntactic category. It follows that they cannot be distinguished by c-selection. What about s-selection? Suppose finite VP, IP and CP are good realizations of the same semantic type: let us call it "proposition". Then it will follow that all verbs which take propositional arguments take all three realizations (VP, IP and CP) of their arguments. (Those verbs which appear to take just CP, such as factives and manner-of-speaking verbs, have different selectional specifications.)

Given the constraints and rankings developed so far, when the input includes a semantic auxiliary, there are two optimal candidates: a CP with *that* as its head, and an IP headed by the auxiliary. When there is no semantic auxiliary in the input, the two optimal candidates are a CP with *that* and a bare VP. A CP with no head is non-optimal in both cases, since it violates OB-HD. Note that the complementizer *that* does not violate FULL-INT, since it does not have an unparsed *lcs*, unlike *do* as analyzed in 3.3.

(T22) VP and CP propositional complements

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
☞ V [_{VP} DP V ...]					
☞ V [_{CP} that [_{VP} DP V ...]]					

(T23) IP and CP propositional complements

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
☞ V [_{IP} DP will [_{VP} t V ...]]					*
☞ V [_{CP} that [_{IP} DP will [_{VP} t V ...]]]					*

The structure of Optimality Theory makes the survival of more than one candidate a difficult result to achieve. In fact some adjustment of the present proposal will be necessary if the head position constraints include HD-RT, as suggested in the previous section. Otherwise the CP complement will always be eliminated in favor of the VP or IP complement, since the CP complement contains an extra left-headed projection (C'), and hence violates HD-RT one more time than the VP or IP complement. It is also necessary to assume that *that*, and presumably all other functional heads, can be freely included in an extended projection or not, without incurring a violation of constraints of either the FILL or PARSE type. Both of these issues merit further exploration.

In general, then, VP, IP and CP alternate in complement position. Nonetheless, there are certain circumstances in which the complementizer occurs obligatorily. When topicalization or any other adjunction occurs in a subordinate clause, the *that* complementizer is obligatory. We can see this below: (42), with no adjunction, is equally good with and without the complementizer, while (43), with *most of the time* adjoined to IP and construed with the subordinate clause, becomes very seriously degraded if the complementizer is omitted.¹⁷

¹⁷ The complementizer is obligatory also in all positions except when the clause is a complement to a verb (Stowell (1981) Kayne (1981)). This generalization, illustrated for subject clauses and complement clauses in (i), is more robust than the obligatory *that* effect analysed here. There are

- (42) a. She swore/insisted/thought that they accepted this solution most of the time.
 b. She swore/insisted/thought they accepted this solution most of the time.
- (43) a. *She swore/insisted/thought(,) most of the time(,) they accepted this solution.
 b. She swore/insisted/thought that(,) most of the time(,) they accepted this solution.

The solution builds on a suggestion by Eric Hoekstra (p.c.): when the complement is a CP then adjunction to the IP is possible, whereas when the complement is an IP then adjunction to IP will be ruled out. Hence, only when there is a CP projection over the IP projection will the IP projection be a possible adjunction site.

Let us consider first of all the situation for adjunction. If the complement is an IP, PROJ-PRIN will be violated, because the IP is a subordinate clause. If the complement is an XP (not shown in the tableau), or a CP with no C, OB-HD will be violated, and if inversion takes place PROJ-PRIN will be violated. But if the complement is a CP headed by *that*, PROJ-PRIN and OB-HD will both be respected. Hence the optimal configuration is a CP with a filled head, hence this is the only grammatical configuration.

(T24) Adjunction to IP and CP complements

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
V [_{IP} adjunct [_{IP} DP will [_{VP} t V ...]]]	*!				*
V [_{CP} will _i [_{IP} adjunct [_{IP} DP e _i [_{VP} t V ...]]]]]	*!				**
V [_{CP} e [_{IP} adjunct [_{IP} DP will [_{VP} t V ...]]]]]			*!		*
 V [_{CP} that [_{IP} adjunct [_{IP} DP will [_{VP} t V ...]]]]]					*

Obligatoriness of *that* is predicted both with adjunction and with negative preposing. The reasoning for negative preposing parallels that for adjunction. If the complement is an XP, either

speakers who systematically accept adjunction and negative-related inversion even in the absence of the complementizer, (Andrew Radford p.c.), but such speakers still reject examples like (ib).

- (i) a. That he left so early shows (that) he was tired.
 b. *He left so early shows (that) he was tired.

The explanation for this effect seems to concern the root character of complements to some verbs, but the issue will not be developed further here.

PROJ-PRIN or OB-HD must be violated: PROJ-PRIN if inversion to X occurs, and OB-HD if inversion to X does not occur. When the complement is a CP headed by *that*, OB-HD is satisfied in CP, and inversion to X is not prohibited by PROJ-PRIN, and satisfies OB-HD for X. This is the optimal structure, and it includes *that*. As (44) shows, *that* is indeed required with negative preposing.

- (44) a. She swore/insisted/thought that never in her life would she accept this solution
 b. *She swore/insisted/thought never in her life would she accept this solution

(T25) Negative Preposing in XP and CP complements

Candidates	PROJ-PRIN	OP-SPEC	OB-HD	FULL-INT	STAY
V [_{XP} never e [_{IP} DP will [_{VP} V ...]]]			*!		*
V [_{XP} never will [_{IP} DP e [_{VP} V ...]]]	*!				**
V [_{CP} that [_{XP} never e [_{IP} DP will [_{VP} V ...]]]]			*!		*
 V [_{CP} that [_{XP} never will _i [_{IP} DP e _i [_{VP} V ...]]]]					**
V [_{CP} e [_{XP} never will _i [_{IP} DP e _i [_{VP} V ...]]]]			*!		**

The CP in all of these cases is present to protect the projection below it from the effects of PROJ-PRIN. When the IP or XP is a subordinate clause it cannot be adjoined to, and its head cannot be filled by movement. It is no longer necessary to appeal to the idea that selectional properties of *that* (its ability to select a CP complement for "CP recursion") are responsible for its appearance here, (Rizzi and Roberts (1989), Cardinaletti and Roberts (1991), McCloskey (1992), Vikner in press). Such a solution rests on a lexical stipulation, while the solution based on constraint conflict rests on general properties of the grammatical system of English. (It seems very likely that the same basic analysis explains the obligatoriness of a complementizer in some embedded V2 systems, see Vikner (in press)).

Comparison of the situation with negative preposing and *wh* movement in subordinate clauses shows an interesting contrast. In subordinate interrogatives, inversion simply fails, since PROJ-PRIN and OB-HD conflict and PROJ-PRIN outranks OB-HD. Why then does inversion not fail also when negative preposing occurs in the XP complement to V? If we get a grammatical sentence by failing to invert in a subordinate interrogative, why don't we get a grammatical sentence by failing to invert in an XP complement to a V? The two sentences have the same constraint profile, and certainly look identical:

- (45) a. He wondered when she would arrive
 b. *He said never he had arrived so late

The solution rests on the comparative nature of optimality theory. In the XP case there is another candidate available, namely the one that includes the CP, which is more successful. In contrast, for an interrogative complement, there is no way to enlarge the extended projection to protect the CP, making inversion possible. The reason is that any additional projection will add an OB-HD violation, given that *that* does not occur outside a interrogative. So the uninverted form of the interrogative is optimal, but the inverted form of the negative is optimal. The very same pattern of constraint satisfaction and violation can yield a grammatical sentence or an ungrammatical sentence, depending on the competition.

Why does *that* not occur outside the *wh* phrase? More generally, why does English not have a way to fill a higher functional head above the projection used for *wh* fronting, as does Spanish (Suñer 1991). Sten Vikner suggests (p.c.) that this follows from OP-SPEC. In Section 2 it was suggested that OP-SPEC is satisfied for a *wh* operator only when the operator is in a position from which it c-commands the entire extended projection. When a higher projection is present, then, OP-SPEC will be violated for *wh* operators. The successful candidate will be one with the *wh* phrase in Specifier of the highest projection in the extended projection. Hence there is no way to enlarge the extended projection to allow inversion with the interrogative.

In sum, the question of why the complementizer is obligatory with topicalization or preposing to Specifier is answered in terms of the principles laid out here. The question of why *that* is obligatory reduces to the question of why a CP projection is obligatory, and this follows from the Projection Principle. OB-HD, responsible for the inversion patterns analyzed in earlier sections, forces the presence of *that* in the optimal CP. A possible extension of these results (suggested by Sten Vikner and Viviane Déprez p.c.) would take the general obligatoriness of the complementizer in Romance and Germanic languages to follow from their having V-to-I movement in finite clauses. This will cause them to violate PROJ-PRIN unless the complementizer is present. In contrast, since there is no head movement involved in the English auxiliary system, *that* is not obligatory in subordinate clauses in general.

9.2 A Note on *that*-trace Configurations

Extraction of an adjunct or a complement is unaffected by the presence of *that*, while the extraction of a subject is possible only if *that* is omitted, except in relative clauses where *that* must be present when it is the highest subject that is extracted. This effect has been widely attributed to the ECP (e.g. Kayne (1981), Lasnik and Saito (1984), Rizzi (1990), and the proposal here is based on such solutions. It builds on the insight of Déprez (1991, 1993) that English *that*-trace configurations are ungrammatical because English offers an alternative, in the form of a *that*-less clause. The proposal to be given here depends crucially on the results presented above concerning clause structure, and on the assumption that (all) heads govern both their complements and the specifiers of their complements. We then posit these government constraints:

T-GOV Trace is governed
 T-LEX-GOV Trace is lexically governed

T-GOV is violated when a trace is not governed by any head, T-LEX-GOV when a trace is not governed by a lexical head.

Let us begin by examining the extraction of complements and adjuncts, both of which are unaffected by the presence or absence of *that*. The reason is that in each case the candidates with and without *that* are equally successful at satisfying the constraints.

- (46) a. Who do you think (that) they will see t?
 b. When do you think (that) they will see them t?

When the object is extracted as in (46a), both constraints are satisfied, so both candidates are optimal, as can be seen in the tableau (T26), which shows only the government constraints (the bold and unindexed trace is the relevant one).

(T26) Extraction of an object

Candidates	T-GOV	T-LEX-GOV
\mathbb{R} V [_{IP} DP _i I [_{VP} t _i V t]]		
\mathbb{R} V [_{CP} that [_{IP} DP _i I [_{VP} t _i V t]]]		

When the adjunct is extracted, as in (46b), neither constraint is satisfied, since an adjunct is not governed at all. Again, then, the candidates with and without *that* are equally successful and both are grammatical.

(T27) Extraction of an adjunct

Candidates	T-GOV	T-LEX-GOV
\mathbb{R} V [_{IP} DP _i I [_{VP} t _i V] t]	*	*
\mathbb{R} [_{CP} that [_{IP} DP _i I [_{VP} t _i V] t]]	*	*

It is precisely in the case of extraction of a subject that the presence of *that* makes a difference.

- (47) a. Who do you think will see them?
 b. *Who do you think that t will see them?

(T28) Extraction of a subject

Candidates	T-GOV	T-LEX-GOV
V [_{IP} t_i I [_{VP} t_i V ...]]		
V [_{CP} that [_{IP} t_i I [_{VP} t_i V ...]]]		* !

When *that* is present, it governs the trace, so the trace is governed, but it is not lexically governed, since C is not lexical. Hence the second candidate violates T-LEX-GOV. When the complement is just an IP, however, as in the first candidate, the complement-taking verb governs the trace, which is therefore lexically governed and satisfies both constraints. Hence the *that*-trace configuration is ungrammatical because the V-trace configuration is optimal.

Relative clauses, being adjuncts, are not governed. It follows that the specifier of a relative clause is not governed by any element from outside the clause. This is the key to the paradigm in (48).

- (48) a. *The people t will see them ...
 b. The people *that* t will see them ...

The trace in Spec of IP is governed by *that* in (48b), hence it is governed, although not lexically governed. The trace in Spec of IP is not governed by any head in (48a), hence this candidate violates both government constraints.¹⁸

¹⁸ I have represented the relative clause with no *that* as just an IP in (T29). A more traditional analysis, based on Chomsky (1977), would posit a CP here with a null operator in its Spec.

- (i) The people [Op e [t will ...]]
 (ii) The people [Op *that* [t will...]]

Assuming that HD-LFT is violated in the second structure, if HD-LFT is outranked by T-GOV, we will get the right result under this representation. There is a problem, however, since if HD-LFT is violated because of the empty operator to the left of C, the prediction is that the candidate with no *that* will be optimal whenever anything other than the highest subject has been relativized, contrary to fact. If, on the other hand HD-LFT is not violated because the OP is empty, then (ii) will still be the optimal version, but we encounter a different problem: OB-HD will require *that* to introduce all relative clauses. This is the reason for the choice of the IP analysis.

(T29) Extraction of a subject in a relative clause

Candidates	T-GOV	T-LEX-GOV
N] [IP <i>t</i> I [VP V ...]]	* !	*
 N] [CP that [IP <i>t</i> I [VP V ...]]]		*

The essential point of this solution is that the *that*-trace configuration is not always ruled out: under certain circumstances it may be the optimal configuration, and be grammatical. Whether it is good or bad for a trace to be governed by C depends on what the alternatives are: it is better to be governed by C than not to be governed at all.

This point holds cross-linguistically as well. The prediction is that languages in which there is no better candidate will simply not show *that*-trace effects, so that extraction corresponding to (47b) will be grammatical. This is exactly the situation in Dutch, as illustrated in this example from Weerman (1989):

- (49) Wie denk je dat *t* ons gezien heeft?
Who think you that *t* us seen has

The observation that *that*-trace configurations are not universally ruled out has been challenging to understand (see e.g. Maling and Zaenen 1978, Sobin 1987, Bayer 1984, Bennis and Haegeman 1984). But in the optimality-theoretic constraint satisfaction perspective, it is just what is expected. In a language which does not allow complementizer-less subordinate clauses, the route to optimality followed by English is not available, so the language must settle for only non-lexical government in this configuration.

The violability of the constraint system governing *that*-*t* effects is visible within English, and lies behind the final puzzle to be analyzed here. (Culicover 1993) shows that the presence of an expression adjoined to IP can make a *that*-trace configuration legitimate:

- (50) a. *Who did she swear(,) most of the time(,) *t* accepted this solution?
b. Who did she swear that(,) most of the time(,) *t* accepted this solution.

Culicover concludes that ECP-based accounts of the effect must be incorrect. However note that in exactly these configurations, there is a conflict between PROJ-PRIN and the lexical government constraint. Tableau (T30) shows the effects.

(T30) PROJ-PRIN conflicts with T-LEX-GOV:

Candidates	PROJ-PRIN	OP-SPEC	O B-HD	FULL-INT	STAY	T-GOV	T-LEX-GOV
☞ V [_{CP} that [_{VP} <i>adjunct</i> [_{VP} <i>t</i> V]]]					*		*
V [_{VP} <i>adjunct</i> [_{VP} <i>t</i> V]]	* !				*		

PROJ-PRIN requires *that* here because of the adjunction. But if *that* is present, then T-LEX-GOV is violated, since the trace is not lexically governed. The correct ranking of these two constraints will therefore automatically give the right result: PROJ-PRIN dominates T-LEX-GOVERNMENT.

10. General Discussion

Every projection is optional, and only present if it is needed. The size of an extended projection is variable, and depends on the effects of grammatical constraints (see Haider (1989), Ackema et al. (1992), Heycock and Kroch (1993) for related proposals). This is inconsistent with the hypothesis that what constructs phrase markers is either phrase structure rules or selectional statements. Optional projections are difficult to make sense of in such systems, because they always involve complicating the PS rules or selection statements. Instead, it seems that what is involved is free combination of heads and their projections, regulated by constraints of the kind discussed here, among others. At this point, however, we can take a more radical step: there is no reason to label the projections which make up extended projections. Every projection is just that, a projection, which has a grammatical category, may have a lexically realized head, may have acquired a head by movement, or may lack one completely. The properties of a projection are just a function of what happens to head it. If it has no head it has no properties other than those imposed by its role in the extended projection. E.g. “XP” is just a verbal projection. Strongly consistent with this is the fact that selectional statements make no reference to distinctions among projections, at least those that are of the same category, such as VP, IP and CP. Moreover, none of the constraints refer to projection labels: a *wh* phrase need not be in Specifier of CP for example, any specifier which c-commands the extended projection will suffice. Adjunction occurs freely, its effects being reined in by PROJ-PRIN. From this perspective, there is, of course, no such thing as “CP recursion” (Rizzi and Roberts (1989) Vikner and Schwartz (1991) McCloskey (1993)), even for cases like Dutch with its three “C” positions (Hoekstra 1993). The “extra” projections that have sometimes been given this analysis are really no different from all other projections: just part of the indefinitely expandible verbal extended projection, which is regulated by the constraints examined in this paper, among others.

Since there is no fixed limit on the number of projections that can be included in an extended projection, the number of competitors also has no fixed limit. However, there is a way to eliminate all but a few candidates; adding projections eventually and reliably leads to worsening status on the constraints. If the addition of, say, two projections leads to a less satisfactory candidate, the addition

of yet another projection will never yield improvement. Thus candidates which are yet larger need not be considered.

There is a clear affinity between the optimality theoretic model presented here and research using notions like “economy”. As I pointed out earlier, STAY is an economy principle, which chooses representations containing the fewest traces. Optimality theory provides a way to understand such notions as “economy”, as just subcases of the total universal set of violable constraints. Moreover, under OT there is an explicit way to determine how constraints will interact with each other. This is of course essential to making sense of any theory in which there is constraint interaction. Consider, for example, the idea that short derivations are less costly than longer ones, and that universal devices are less costly than language particular ones, as in Chomsky (1991). Without a means of computing the comparative cost of the two expensive items, there is no way to calculate the results of interaction between them. What happens when we have to choose between a derivation with fewer steps but more language particular devices and one with more steps and fewer language particular devices? OT provides a theory of constraint interaction which makes such questions answerable: the choice will depend on the ranking of the constraints in the grammar.

References

- Ackema, Peter, Ad Neeleman and Fred Weerman. 1992. Deriving functional projections. *Proceedings of NELS*.
- Baker, L. 1991. The Syntax of English *Not*: The Limits of Core Grammar. *Linguistic Inquiry* 22:387-429.
- Bakovic, Eric. 1995. A Markedness Subhierarchy in Syntax. ms. Rutgers University.
- Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13: 1-38.
- Bayer, Joseph 1984. Towards an explanation for certain that-t phenomena: The COMP-node in Bavarian. In *Sentential Complementation*, ed. W. de Geest and Y. Putseys, 23-32. Dordrecht: Foris.
- Besten, Hans den. 1983. On the interaction of root transformations and lexical deletive rules. In *On the formal syntax of the Westgermania*, ed. Werner Abraham, 47-131. Amsterdam: Benjamins.
- Brisson, Christine. 1994. Noun Phrases, *of*, and Optimality. Ms. Rutgers University, New Brunswick, NJ.
- Burton, Strang and Jane Grimshaw. 1992. Coordination and VP-internal subjects. *Linguistic Inquiry* 23: 305-313.
- Cardinaletti, Anna. 1992. Spec CP in Verb-Second Languages. in M. Starke ed. *Geneva Generative Papers*, Volume 0 Number 0, 1-9. Departement de Linguistic Generale, Geneva.
- Cardinaletti, Anna and Ian Roberts. 1991. Clause-structure and X-Second. Ms. Università di Venezia and Université de Genève, Venice and Geneva, Switzerland and Italy.
- Cheng, Lisa L. S. 1991. On the typology of Wh-questions. Doctoral dissertation. MIT, Cambridge, MA.
- Chomsky, Noam. 1957. *Syntactic Structures*. den Haag: Mouton.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417-453. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1992. A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics*: Number 1. Cambridge, MA: MIT Press.

- Chung, Sandra and James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14: 704-13.
- Cinque, Guglielmo. 1990. *Types of A-Bar dependencies*. Cambridge, MA: MIT Press.
- Collins, Chris. 1991. Why and How Come. *MIT Working Papers in Linguistics* 15.
- Déprez, Viviane. 1991. Economy and the that-t effect. Proceedings of ESCOL.
- Déprez, Viviane. 1991. Wh-movement: adjunction and substitution. *Proceedings of WCCFL X*.
- Déprez, Viviane. 1993. A minimal account of the that-t effect. To appear in *Festschrift for Richard Kayne*.
- Doherty, Cathal. 1993. Clauses without *that*: the case for bare sentential complementation in English. Doctoral dissertation, University of California, Santa Cruz.
- Dwivedi, Veena. 1993. Topicalization in Hindi and small pro. Presented at the Workshop on Ergativity, Rutgers University.
- Emonds, Joseph. 1975. A transformational approach to English syntax. Academic Press: New York, NY.
- Emonds, Joseph. 1978. The verbal complex of V'-V in French. *Linguistic Inquiry* 9:151-175.
- Engdahl, Elisabet. 1980. Wh-constructions in Swedish and the relevance of Subjacency. Proceedings of NELS 10.
- Erteschik-Shir, Nomi. 1992. Focus Structure and Predication, the case of negative wh-questions. *Belgian Journal of Linguistics* 7: 35-51.
- Erteschik-Shir, N. In preparation. *Focus*.
- Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12: 155-184.
- Geest, W. de and Y. Putseys, eds. 1984. *Sentential Complementation*. Dordrecht: Foris.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10: 279-326.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Grimshaw, Jane. 1991. Extended Projection. Ms, Brandeis University, Waltham, MA.

Grimshaw, Jane 1994. Minimal Projection and Clause Structure. In B. Lust, J. Whitman and J. Kornfilt eds. *Syntactic Theory and First Language Acquisition: Cross Linguistic Perspectives - Volume I: Heads, Projections, and Learnability*, Erlbaum. 75-83.

Grimshaw, Jane and R. Armin Mester. 1988. Light verbs and Theta-Marking. *Linguistic Inquiry* 19: 205-232.

Grimshaw, Jane and Vieri Samek-Lodovici. 1994. Optimal Subjects. Ms, Rutgers University, New Brunswick NJ.

Haegeman, Liliane. 1992. Sentential negation in Italian and the Neg Criterion. In M. Starke ed. *Geneva Generative Papers*, Volume 0 Number 0, 10-26. Departement de Linguistic Generale, Geneva.

Haegeman, Liliane and Hans Bennis. 1984. On the status of agreement and relative clauses in West-Flemish. In *Sentential Complementation*, ed. W. de Geest and Y. Putseys, 33-53. Dordrecht: Foris.

Haider, Hubert. 1989. Matching projections. In *Constituent Structure: Papers from the 1987 GLOW Conference*, ed. Anna Cardinaletti, Guglielmo Cinque and G. Giusti, 101-122. Dordrecht: Foris.

Halle, Morris, and A. Marantz. 1993. Distributed Morphology. In Kenneth Hale and Samuel Jay Keyser ed. *The View from Building 20*. MIT Press, Cambridge, MA. 111-176.

Heycock, Caroline and Anthony Kroch. 1993. Verb movement and coordination in the Germanic languages: evidence for a relational perspective on licensing. Ms. University of Pennsylvania, Philadelphia, PA.

Hoekstra, Eric. 1993. On the parametrisation of functional projections in CP. In *Proceedings of the North East Linguistic Society* 23, volume 1, 191-204.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation. MIT, Cambridge, MA.

Iatridou, Sabine and David Embick. 1994. Conditional inversion. In *Proceedings of NELS* 24.

Iatridou, Sabine and Anthony Kroch. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Syntax* 50: 1-24.

Jespersen, O. 1954. *A Modern English Grammar on Historical Principles*, George Allen & Unwin, London, and Ejnar Munksgaard, Copenhagen.

- Kayne, Richard. 1981. ECP extensions. *Linguistic Inquiry* 12: 93-133.
- Kayne, Richard. 1982. Predicates and arguments, verbs and nouns. Abstract of paper presented to the 1982 GLOW conference. *GLOW Newsletter* 8: 24.
- Kayne, Richard. 1983. Chains, categories external to S, and French complex inversion. *Natural Language & Linguistic Theory* 1: 109-137.
- Kayne, Richard. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22: 647-686.
- Kim, Soo Won. 1990. *Chain scope and quantificational structure*. Doctoral dissertation. Brandeis University, Waltham, MA.
- Kitagawa, Yoshihisa. 1986. *Subjects in Japanese and English*. Doctoral dissertation. University of Massachusetts, Amherst, MA.
- Klima, Edward. 1964. Negation in English. In *The Structure of Language*, eds. Jerry Fodor and Jerrold Katz. New York: Prentice-Hall.
- Koopman, Hilda and Dominique Sportiche. 1991. The position of subjects. In *The syntax of verb-initial languages*, ed. James McCloskey. Elsevier. [Published as a special issue of *Lingua*].
- Laka Mugarza, M. Itziar. 1990. *Negation in syntax: On the nature of functional categories and projections*. Doctoral dissertation. MIT, Cambridge, MA.
- Lasnik, Howard and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235-289.
- Lasnik, Howard and Mamoru Saito. 1992. *Move α* . Cambridge, MA: MIT Press.
- Lasnik, Howard and Tim Stowell. 1991. Weakest crossover. *Linguistic Inquiry* 22: 687-720.
- Legendre, Géraldine, William Raymond and Paul Smolensky. 1993. An Optimality-Theoretic typology of case and grammatical voice systems. *Proceedings of the Berkeley Linguistics Society* 19.
- Liberman, Mark. 1974. On conditioning the rule of Subject-Auxiliary Inversion. *Proceedings of NELS V*, 77-91.
- McCarthy, John and Alan Prince. 1993. Generalized Alignment. In *Yearbook of Morphology 1993*, ed. Geert Booij and Jaap van Marle, 79-153. Dordrecht: Kluwer.
- McCloskey, James. 1992. Adjunction, selection and embedded verb second. *Linguistics Research Report LRC-92-07*. University of California at Santa Cruz, Santa Cruz, CA.

- McNally, Louise. 1992. VP-coordination and the VP-internal subject hypothesis. *Linguistic Inquiry* 23: 336-341.
- Müller, Gereon and Wolfgan Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461-507.
- Pesetsky, David. 1989. Language particular processes and the Earliness Principle. Unpublished ms. Cambridge, MA: MIT.
- Platzack, Christer and Anders Holmberg. 1989. The rôle of AGR and finiteness. *Working Papers in Scandinavian Syntax* 31: 9-14.
- Plunkett, Bernadette. 1990. Inversion and early questions. In *Papers in the acquisition of Wh*, ed. T. L. Maxfield and B. Plunkett. *University of Massachusetts Occasional Papers*, special edition. University of Massachusetts, Amherst, MA.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424
- Postal, Paul. 1992. Overlooked extraction distinctions. Ms. T. J. Watson Research Center, IBM, Yorktown Heights, NY.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. RuCCS Technical Report #2. Rutgers University Center for Cognitive Science, Piscataway, NJ.
- Raposo, Eduardo. 1987. Case Theory and Infl-to-Comp: the inflected infinitive in European Portuguese. *Linguistic Inquiry* 18: 85-109.
- Reinhart, Tanya. 1993. Wh-in-situ in the framework of the Minimalist Program. Ms. Utrecht and Tel-Aviv.
- Riemsdijk, Henk van. 1990. Functional prepositions. In *Unity in Diversity*, ed. H. Pinkster and I. Genee. Dordrecht: Foris.
- Ritter, Elizabeth and Sara Thomas Rosen. 1993. Deriving causation. *Natural Language & Linguistic Theory* 11: 519-555.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1991. Residual Verb Second and the Wh Criterion, *Technical Reports in Formal and*

Computational Linguistics 2. Faculté des Lettres, Université de Genève, Geneva.

Rizzi, Luigi and Ian Roberts. 1989. Complex inversion in French. *Probus* 1.1.

Roberts, Ian. 1993. *Verbs and diachronic syntax*. Dordrecht: Foris.

Rochemont, Michael. 1989. Topic islands and the Subjacency parameter. *Canadian Journal of Linguistics* 43: 145-170.

Sadock, Jerrold. 1984. The pragmatics of subordination. In *Sentential Complementation*, ed. de Geest and Putseys, 205-213. Dordrecht: Foris.

Safir, Kenneth. 1993. Perception, selection, and structural economy. Ms. Rutgers University, New Brunswick, NJ.

Smolensky, Paul, Colin Wilson, Géraldine Legendre, Kristin Homer and William Raymond. 1994. Scope and wh-movement in Optimality Theory. Presented at the Society for Philosophy and Psychology, Memphis, TN.

Sobin, Nicholas. 1987. The variable status of COMP-trace phenomena. *Natural Language & Linguistic Theory* 5: 33-60.

Stowell, Tim. 1981. *Origins of phrase structure*. Doctoral dissertation. MIT, Cambridge, MA.

Suñer, Margarita. 1991. Indirect questions and the structure of CP. In *Current issues in Spanish linguistics*, ed. H. Campos and F. Martínez-Gil. Washington, DC: University of Georgetown Press.

Travis, Lisa. 1989. Parameters of phrase structure. In *Alternative Conceptions of Phrase Structure*, ed. Mark Baltin and Anthony Kroch. Chicago, IL: University of Chicago Press.

Travis, Lisa. 1991. Parameters of phrase structure and V2 phenomena. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin. Cambridge, MA: MIT Press.

Vikner, Sten. In press. *Verb movement and expletive subjects in the Germanic languages*. Oxford University Press.

Vikner, Sten and Bonnie Schwartz. 1992. The verb always leaves IP in V2 clauses. To appear in *Parameters and functional heads*, ed. Adriana Belletti and Luigi Rizzi. Oxford: Oxford University Press.

Weerman, Fred. 1989. *The V2 conspiracy*. Dordrecht: Foris.

Williams, Edwin. 1994. *Thematic structure in syntax*. Cambridge, MA: MIT Press.

Zagona, Karen. 1982. *Government and Proper Government of verbal projections*. Doctoral dissertation. University of Washington, Seattle, WA.

Zanuttini, Raffaella. (1991) *Syntactic properties of sentential negation: a comparative study of Romance languages*. Doctoral dissertation. University of Pennsylvania, Philadelphia, PA.