

V°-TO-I° MOVEMENT AND DO-INSERTION IN OPTIMALITY THEORY

Sten Vikner

Institut für Linguistik/Germanistik, Universität Stuttgart,
Postfach 10 60 37, D-70049 Stuttgart, Germany

vikner@rus.uni-stuttgart.de

<http://pinguin.philosophie.uni-stuttgart.de/ifl/Vikner.html>

To appear in Géraldine Legendre, Jane Grimshaw, & Sten Vikner (eds.):

Optimality-theoretic Syntax, MIT Press, Cambridge MA.

Final version, October 1999.

I would like to thank Eric Baković, Fabian Heck, Gunnar Hrafn Hrafnbjargarson, Jóhannes Gísli Jónsson, Takeo Kurafuji, Edward Keer, Gereon Müller, Nicole Nelson, Alan Prince, Luigi Rizzi, Ian Roberts, Ramona Römisch-Vikner, Vieri Samek-Lodovici, Tanja Schmid, Halldór Ármann Sigurðsson, Paul Smolensky, Melissa Trachtenberg, Carl Vikner, Ralf Vogel, Colin Wilson, and Heike Zinsmeister, and in particular Jane Grimshaw and Géraldine Legendre. Thanks also to participants in classes at the Universities of Stuttgart, Geneva, Iceland, Lund, and Rutgers University, and to audiences at talks given at the Hopkins Optimality Theory Workshop/University of Maryland Mayfest (Baltimore), at the Minimal Link Condition Workshop (Potsdam), and at the Universities of Amsterdam, Cologne, Groningen, Leiden, Tilburg, Tromsø, Tübingen, Utrecht, Venice, Johns Hopkins University, the Netherlands Institute for Advanced Study (Wassenaar), and the Arbeitsgruppe Strukturelle Grammatik (Max-Planck-Gesellschaft, Berlin). The research reported here was supported by the Deutsche Forschungsgemeinschaft as part of the Project *The Optimality-theoretic Syntax of German from a Comparative Germanic Perspective* (Grant MU 1444/2-1, Principal Investigators: G. Müller, S. Vikner).

1. Introduction

The empirical phenomena to be discussed in this paper are the position of the finite verb in embedded clauses (preceding or following a sentential adverbial, i.e. in I° or in V°) and the possibility or impossibility of *do*-insertion. For reasons of space, it will only be sketched at the very end how the analysis may be extended to cover those cases (mostly main clauses) where the verb may or may not move to a position to the left of the subject (i.e. to C°).

To a large extent, I will base my analysis on Grimshaw's (1997) analysis of English, and attempt to modify and extend it to also cover the variation found between English, Danish, French and Icelandic. One of my goals is to illustrate how Optimality Theory (OT) can account for language variation in terms of differences in constraint rankings between languages.

A central feature of OT (see the introductory chapter of the present volume and references there) is that constraints are taken to be relative ("soft") rather than absolute ("hard"):

(1) a. **ABSOLUTE:**

"If a sentence violates constraint C, it is ungrammatical"

b. **RELATIVE:**

"That a sentence violates constraint C may have a certain cost, but this cost is smaller than the cost of violating constraint B, which again is less costly than a violation of constraint A"

In other words: Although there is a price to be paid every time a constraint is violated, the price is not always the grammaticality of the sentence in question.

As will become clear, my particular version of OT owes a heavy debt to the Principles and Parameters framework (see Chomsky 1986, Rizzi 1990, and e.g. Vikner 1995:3-35), although violability of constraints was not possible within this framework. Furthermore, although violability is a feature of the Minimalist framework (Chomsky 1995), this is only the case to a certain extent

(see e.g. the discussion in Vikner 1997a:12-13), although certain implementations of minimalism differ from Chomsky (1995) precisely in assigning a more central role to violability, see e.g. Bobaljik (1995:351) and Holmberg (1997:214).

However, violability is not the only central feature of OT. Following Prince & Smolensky (1993) and Grimshaw (1997:373), I take the following four ideas to sum up the essence of OT:

- (2) a. **Constraints may be violated.**
- b. **Constraints are ranked in a hierarchy**
 (a grammar is a particular ranking of constraints).
- c. **Constraints are universal,**
 i.e. in all languages, the same constraints apply, except that they are ranked differently from language to language (language variation is variation in the constraint hierarchy).
- d. **Only the optimal version of a sentence is grammatical,**
 all non-optimal versions are ungrammatical (the optimal candidate of two is the one with the least violations of the highest constraint on which the two candidates differ).

By way of illustration of (2a-d), consider the variation in position of the finite verb in English and French, as originally discussed in Emonds (1978) and Pollock (1989). This particular difference is one of the central differences to be discussed in this paper. The same difference, although restricted to embedded clauses, is found between Danish and Icelandic, as will be shown in more detail later.

- (3) En. a. The actor really saw the film
- b. *The actor saw really the film

- (4) Fr. a. *L'acteur vraiment voit le film
 b. L'acteur voit vraiment le film
The actor (sees) really (sees) the film

I am assuming (at least) two kinds of movement to take place here: The finite inflectional affix of the verb moves from I° to V° in (3a)/(4a), whereas the finite verb moves from V° to I° in (3b)/(4b).

This variation can be seen as resulting from the interaction between two constraints, which will be discussed in more detail in connection with (13) in section 3 below: **Lx-Mv** (No Movement of a Lexical Head: violated when a verb leaves V°) and **Pr-Bd** (Proper Binding: violated when a trace - here: in I° - c-commands its antecedent - here: in V°). In English, the price to be paid when a trace c-commands its antecedent, i.e. when something is moved downwards in the tree, is smaller than the price to be paid when a lexical category (here the main verb), is moved, because the ranking is Lx-Mv >> Pr-Bd. In French, on the other hand, the ranking is Pr-Bd >> Lx-Mv, and so the cost of moving the verb out of VP is smaller than the cost of moving an element downwards in the tree.

- | | | |
|--|-----------------------------|---------|
| (5) <u>English (and embedded clauses in Danish)</u> | <u>Lx-Mv >> Pr-Bd</u> | |
| a. The actor v _j really saw _j the film | <i>violates Pr-Bd</i> | = (10d) |
| b. *The actor saw _j really v _j the film | <i>violates Lx-Mv</i> | = (10e) |
| (6) <u>French (and embedded clauses in Icelandic)</u> | <u>Pr-Bd >> Lx-Mv</u> | |
| a. *L'acteur v _j vraiment voit _t le film | <i>violates Pr-Bd</i> | = (18a) |
| b. L'acteur voit _t vraiment v _j le film | <i>violates Lx-Mv</i> | = (18b) |

As we see, constraints are violable, (2a), cf. that both (5a) and (6b) violate a constraint and yet they are grammatical. Constraints are ranked in a hierarchy, (2b), and they are universal, (2c), cf. that the same constraints are found in English and French, though in a different ranking. Finally, only the optimal version of a sentence is grammatical, (2d). Both (5b) and (6a) are ungrammatical, because they are less optimal than their competitors (5a) and (6b). On the highest ranked constraint on which there is a difference, (5a) has less violations than its competitor (5b), and (6b) has less violations than its competitor (6a).

2. Preliminary assumptions

Before I begin the discussion of the analysis itself, I want to set out a few preliminary assumptions:

The competing versions of a sentence (the "candidates") are different realisations of the same input. I will follow Grimshaw (1997:375-376) and assume that "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 aspect".

A typical set of competing candidates, in this case for an embedded negated clause, contains at least the following:

- (7) En. a. * Although Peter saw not the film ...
b. * Although Peter not saw the film ...
c. Although Peter did not see the film ...
d. * Although Peter not did see the film ...

Now it is up to the optimality theory linguist to figure out which constraints are relevant here and how they are ranked, so that only (7c) is optimal in

Finally, I will be using the following notational conventions:

- (9) a. \Rightarrow marks the optimal candidate(s), i.e. the grammatical version(s).
- b. ! marks the fatal violation in all non-optimal candidates.
- c. t is a trace of a maximal projection.
- d. v is a trace in the verb chain (a trace of the verb or of the inflection).
- e. - is an empty maximal projection.
- f. e is an empty X° (either only abstract features or radically empty).
- g. Su means "subject".
- h. Ob means "object".
- i. A means "adverbial".

Please notice that regardless of whether the example in question is English or not, the lexical items in the corresponding tableau are always English.


3. The six basic constraints

The basic phenomena to be discussed are the various possible positions of the finite verb in Danish, English, French, and Icelandic, as frequently discussed in the literature, e.g. Emonds (1978), Pollock (1989), Holmberg & Platzack (1995), Vikner (1995, 1997b).

I will begin the analysis with the six constraints that I take to be involved in the derivation of embedded declaratives. (10) shows an analysis valid for both Danish and English. In both languages the optimal candidate is (10d). (10a-u) make up a rather extensive set of candidates, which vary with respect to three different properties: presence or absence of *do*, the position of the subject (VP-spec, IP-spec or CP-spec), and the position of the finite verb (V° , I° or C°):


(10) α. Da. Jeg tror (at) skuespilleren virkelig så filmen

β. En. I think (that) the actor actually saw the film

	Case	Ob-Hd	Lx-Mv	Pr-Bd	Stay
a. * C [IP- v A [VP _{Su} saw Ob	*!			*	*
b. * C [IP- saw A [VP _{Su} v Ob	*!		*		*
c. * C [CP- saw [IP- v A [VP _{Su} v Ob	*!		*		**
d.  C [IP _{Su} v A [VP _t saw Ob				*	**
e. * C [IP _{Su} saw A [VP _t v Ob			*!		**
f. * C [CP- saw [IP _{Su} v A [VP _t v Ob			*!		***
g. * C [CP _{Su} e [IP _t v A [VP _t saw Ob		*!		*	***
h. * C [CP _{Su} e [IP _t saw A [VP _t v Ob		*!	*		***
i. * C [CP _{Su} saw [IP _t v A [VP _t v Ob			*!		****
j. * C [IP- v A [VP- did [VP _{Su} see Ob	*!			*	*
k. * C [IP- did A [VP- v [VP _{Su} see Ob	*!		*		*
l. * C [CP- did [IP- v A [VP- v [VP _{Su} see Ob	*!		*		**
m. * C [IP- v A [VP _{Su} did [VP _t see Ob	*!			*	**
n. * C [IP- did A [VP _{Su} v [VP _t see Ob	*!		*		**
o. * C [CP- did [IP- v A [VP _{Su} v [VP _t see Ob	*!		*		***
p. * C [IP _{Su} v A [VP _t did [VP _t see Ob				*	***!
q. * C [IP _{Su} did A [VP _t v [VP _t see Ob			*!		***
r. * C [CP- did [IP _{Su} v A [VP _t v [VP _t see Ob			*!		****
s. * C [CP _{Su} e [IP _t v A [VP _t did [VP _t see Ob		*!		*	****
t. * C [CP _{Su} e [IP _t did A [VP _t v [VP _t see Ob		*!	*		****
u. * C [CP _{Su} did [IP _t v A [VP _t v [VP _t see Ob			*!		*****


The first of the five constraints in (10) is **Case** (every DP or DP-chain must have case, see Grimshaw 1997:390 and the case filter in Chomsky 1981:49, and see also Costa, this volume, Samek-Lodovici 1996:169, and Woolford, this volume, for more elaborate OT-implementations of the case filter), which makes the subject leave VP-spec and move to IP-spec, on the assumption that nominative is assigned to IP-spec. Compare (10a) to (10d):

(11) (repeated from (10a,d) above)

	Case	Ob-Hd	Lx-Mv	Pr-Bd	Stay
a. * C [IP - v A [VP Su saw Ob	*!			*	*
d.  C [IP Su v A [VP t saw Ob				*	**

Ob-Hd (Obligatory Heads, an X^0 may not be empty, see Grimshaw 1997:377, Haider 1988:101) makes sure there are no radically empty heads. A head containing a trace is not taken to be empty. Compare (10g) to (10d):


(12) (repeated from (10d,g) above)

	Case	Ob-Hd	Lx-Mv	Pr-Bd	Stay
d.  C [IP Su v A [VP t saw Ob				*	**
g. * C [CP Su e [IP t v A [VP t saw Ob		*!		*	***

I am assuming that X-bar theory forces the existence of a head (C^0 in (10g)) whenever the relevant specifier position is filled (CP-spec in (10g), filled by the subject).

Lx-Mv (No Movement of a Lexical Head, see Grimshaw 1997:385) punishes *så/saw* for leaving V^0 . Compare (10e) to (10d):

(13) (repeated from (10d,e) above)


	Case	Ob-Hd	Lx-Mv	Pr-Bd	Stay
d.  C [IP Su v A [VP t saw Ob				*	**
e. * C [IP Su saw A [VP t v Ob			*!		**

Pr-Bd (Proper Binding, traces must not c-command their antecedent, an updated version of Fiengo 1977:45,53) is the (low) cost of having an "upside-down" chain, where the trace in I^0 c-commands the antecedent in V^0 . This constraint is not actually necessary for verb movement in Danish and English,

as there are no candidates which are ruled out by it, but it will be necessary for verb movement in French and Icelandic, see (18) below. Pr-Bd is probably also relevant for other kinds of movement in Danish and English, which like verb movement in French and Icelandic go upwards, not downwards.

Stay (traces should be avoided, see Grimshaw 1997:375, and also "***trace**" in Legendre et al. 1995:611 and economy of derivation in Chomsky 1991:432, 1995:145) punishes every step of every movement ("movement as last resort"). Compare (10p) to (10d):

(14) (repeated from (10d,p) above)


	Case	Ob-Hd	Lx-Mv	Pr-Bd	Stay
d.  C [IP Su v A [VP t saw Ob				*	**
p. *C [IP Su v A [VP t did [VP t see Ob				*	***!

Here *do* is inserted under V° (for an alternative possibility, direct insertion of *do* under I°, see section 5 below).

The ranking of the five constraints in (10) thus causes (10d) to be optimal, i.e. better than its competitors (10a-c) and (10e-u), and this is why (10d) gives rise to a grammatical sentence, whereas the competitors either result in ungrammatical sentences or in a sentence which might also be analysed as (10d) (i.e. (10g)).

However, (10d) is not the only possible grammatical version of the sentence in question: It is also possible to leave out the complementiser *that*. The following, (15a-u), are all further candidates competing with (and parallel to) the ones already encountered above, (10a-u), except that they do not contain the complementiser *that*:


(15) (continues and completes (10) above)

	Case	Pr-Pr	Ob-Hd	Lx-Mv	Pr-Bd	Stay
a. * [IP- v A [VP _{Su} saw Ob	*!				*	*
b. * [IP- saw A [VP _{Su} v Ob	*!	*		*		*
c. * [CP- saw [IP- v A [VP _{Su} v Ob	*!	*		*		**
d.  [IP _{Su} v A [VP _t saw Ob					*	**
e. * [IP _{Su} saw A [VP _t v Ob		*!		*		**
f. * [CP- saw [IP _{Su} v A [VP _t v Ob		*!		*		***
g. * [CP _{Su} e [IP _t v A [VP _t saw Ob			*!		*	***
h. * [CP _{Su} e [IP _t saw A [VP _t v Ob			*!	*		***
i. * [CP _{Su} saw [IP _t v A [VP _t v Ob		*!		*		****
j. * [IP- v A [VP- did [VP _{Su} see Ob	*!				*	*
k. * [IP- did A [VP- v [VP _{Su} see Ob	*!	*		*		*
l. * [CP- did [IP- v A [VP- v [VP _{Su} see Ob	*!	*		*		**
m. * [IP- v A [VP _{Su} did [VP _t see Ob	*!				*	**
n. * [IP- did A [VP _{Su} v [VP _t see Ob	*!	*		*		**
o. * [CP- did [IP- v A [VP _{Su} v [VP _t see Ob	*!	*		*		***
p. * [IP _{Su} v A [VP _t did [VP _t see Ob					*	***!
q. * [IP _{Su} did A [VP _t v [VP _t see Ob		*!		*		***
r. * [CP- did [IP _{Su} v A [VP _t v [VP _t see Ob		*!		*		****
s. * [CP _{Su} e [IP _t v A [VP _t did [VP _t see Ob			*!		*	****
t. * [CP _{Su} e [IP _t did A [VP _t v [VP _t see Ob			*!	*		****
u. * [CP _{Su} did [IP _t v A [VP _t v [VP _t see Ob		*!		*		*****

This illustrates one way of allowing for optionality in OT: The two winning candidates, (10d) and (15d), have the same constraint profile (as suggested by Grimshaw 1997:410, see Baković & Keer, this volume, and Müller 1999 for further discussion of various ways of deriving optionality). The only difference between the two tableaux containing candidates from the same competition, (10a-u) and (15a-u), is the absence vs. presence of *that*. This is relevant for **Pr-Pr** (Projection Principle), which is violated when there is movement into the



highest X° of an embedded clause (following Grimshaw's 1997:393 suggestion, based on Rizzi & Roberts 1996:106, see also McCloskey 1992, and Müller 1997:262, and notice that in Grimshaw 1997:393 and Baković 1998:40, Pr-Pr is called "Pure-EP"). Pr-Pr was not shown in tableau (10), as it was not violated by any of the candidates there. In order for Pr-Pr to be violated, there has to be movement from V° to I° , compare (16b) to (16c), and I° has to be the highest X° of the embedded clause, i.e. C° has to be absent, compare (16a) to (16c):

(16) (repeated from (10e) and (15d,e) above)

	Case	Pr- Pr	Ob- Hd	Lx- Mv	Pr- Bd	Stay	
a. * C [IP Su saw A [VP t v Ob				*!		**	= (10e)
b.  [IP Su v A [VP t saw Ob					*	**	= (15d)
c. * [IP Su saw A [VP t v Ob		*!		*		**	= (15e)

The introduction of Pr-Pr has no consequences in this Danish/English example because the optimal candidate does not involve V° -to- I° movement: It does not affect (17b), the version of (17a) without the complementiser *at/that*, which thus has exactly the same constraint profile as (17a), and given that this constraint profile is the optimal one, both candidates are optimal and thus grammatical:

(17) (repeated from (10d) and (15d) above)

	Case	Pr- Pr	Ob- Hd	Lx- Mv	Pr- Bd	Stay	
a.  C [IP Su v A [VP t saw Ob					*	**	= (10d)
b.  [IP Su v A [VP t saw Ob					*	**	= (15d)

In fact, none of (10a-u) violate Pr-Pr, quite irrespectively of whether they have movement of the finite verb, due to the presence of the complementiser *at/that*, which will always count as the highest head of the embedded clause.

This whole argumentation presupposes that embedded clauses with and without *that* may both be winning candidates in the same competition ("candidates realising the same input"), although they do not contain exactly the same lexical elements (in minimalist terms, they do not have the same "numeration", see Chomsky 1995:225). For an alternative OT analysis that takes clauses with and without *that* to be winners of different competitions, see Legendre et al. (1995, 1998) and Baković & Keer (this volume).


4. Optional vs. obligatory complementisers

Consider now embedded declaratives in French and Icelandic, which are derived by means of the same six constraints. In these two languages, the finite verb precedes the adverbial instead of following it, and there is also no complementiser optionality such as the one found in Danish and English: *que/að*, 'that', is obligatory.

(18) a. Fr. Je crois *(que) l'acteur voit vraiment le film

β. Ic. Ég tel *(að) leikarinn sjái áreiðanlega myndina

I think (that) the actor sees really the film

	Case	Pr-Pr	Ob-Hd	Pr-Bd	Lx-Mv	Stay
a. * C [IP Su v A [VP t saw Ob				*!		**
b.  C [IP Su saw A [VP t v Ob					*	**
c. * [IP Su v A [VP t saw Ob				*!		**
d. * [IP Su saw A [VP t v Ob		*!			*	**

As briefly discussed in section 1, I propose a difference in constraint ranking between Danish/English and French/Icelandic: Danish and English have Lx-Mv >> Pr-Bd, French and Icelandic have Pr-Bd >> Lx-Mv:

(19) Danish and English:	Case	Pr-Pr	Ob-Hd		Lx-Mv	Pr-Bd	Stay
French and Icelandic:	Case	Pr-Pr	Ob-Hd	Pr-Bd	Lx-Mv		Stay

In French/Icelandic, violating Lx-Mv (moving the verb out of VP) is thus less costly than violating Pr-Bd (leaving the verb in V° where it is commanded by a coindexed trace in I°), exactly the opposite of Danish/English. Hence (18b) is optimal and (18a) is not, whereas (10d) (the Danish/English parallel to (18a)) was optimal and (10e) (the Danish/English parallel to (18b)) was not.

As I originally observed some years ago, cf. Grimshaw (1997:413), it now follows that the French/Icelandic complementiser *que/að*, 'that', is obligatory: Leaving the complementiser out of the optimal (18b) results in (18d) which violates Pr-Pr, as it has movement into the highest X° of an embedded clause (parallel to a suggestion in Deprez 1994:124).

5. Insertion of auxiliaries in I° and economy of projection

In Danish, French, and Icelandic, finite auxiliary verbs behave exactly like their non-auxiliary counterparts. All finite verbs follow the adverbial in Danish, and all finite verbs precede the adverbial in French and Icelandic:

- (20) a. Da. Jeg tror (at) skuespilleren virkelig ser filmen
 b. Fr. Je crois que l'acteur voit vraiment le film
 c. Ic. Ég tel að leikarinn sjái áreiðanlega myndina
I think that the actor (sees) really (sees) the film

- (21) a. Da. Jeg tror (at) skuespilleren virkelig har set filmen
 b. Fr. Je crois que l'acteur a vraiment vu le film
 c. Ic. Ég tel að leikarinn hafi áreiðanlega séð myndina
I think that the actor (has) really (has) seen the film

This is not the case in English, however: Finite main verbs follow the adverbial (like Danish), but finite auxiliaries precede the adverbial (like French and Icelandic):

- (22) En. a. I think (that) the actor actually saw the film
 b. I think (that) the actor has actually seen the film

English even has one more complication, namely that main verb *be* behaves like an auxiliary rather than like a main verb.

In order to capture this distributional difference inside one and the same language, the means used in sections 3 and 4 above, reranking of constraints, is not available: Constraint ranking is language-specific, not construction-specific.

The solution to be suggested here (which is an attempt to derive the analysis stipulated in Grimshaw 1997:382, see also Emonds 1994:157-164) is that the English finite auxiliary is inserted directly under I° , in contrast to English finite main verbs and to both main and auxiliary finite verbs in Danish, French, and Icelandic, all of which are inserted under V° .

I would like to suggest that two conflicting constraints exist: **Ec-VP** and **V-V $^\circ$** . Depending on which one is more highly ranked, (auxiliary) verbs are inserted in V° or in I° :

Ec-VP (Economy of VPs) punishes every occurrence of a VP, i.e. the fewer VPs used in a derivation, the better. If this constraint is the crucial one,

insertion of a verb directly under I° is preferred, because it uses one less VP than when a verb is inserted under V°. Compare (23b) to (23c). I formulate this constraint as Ec-VP for expository reasons only, it might as well have been formulated as a more general Ec-XP (Economy of maximal projections).

V-V° (Verb in V°) punishes chains which contain a verb but which do not include a V° (all verbs should be inserted into their own V° and project their own VP), so if this constraint is the crucial one, insertion of a verb under V° is preferred. Compare again (23b) to (23c). Notice that I take both main verbs and auxiliary verbs to be verbs in the sense of this constraint.

The difference between (23a,b) and (23c) below is that (23c) has insertion of *has* directly under I° and therefore has one VP less than the corresponding candidate with *has* inserted under V°, (23b).

(23) En. I think (that) the actor has actually seen the film

		θ-Fl	Case	Pr- Pr	Ob- Hd	Ec- VP	Lx- Mv	V- V°	Pr- Bd	Stay
a.	* C [IP Su v A [VP t has [VP t seen Ob					**!			*	***
b.	* C [IP Su has A [VP t v [VP t seen Ob					**!	*			***
c.	☞ C [IP Su has A [VP t seen Ob					*		*		*
d.	* [IP Su v A [VP t has [VP t seen Ob					**!			*	***
e.	* [IP Su has A [VP t v [VP t seen Ob			*!		**	*			***
f.	☞ [IP Su has A [VP t seen Ob					*		*		*


The candidates (23d-f) correspond to (23a-c), but lack the complementiser *that*. Presence or absence of *that* makes no difference here, the constraint profiles of (23c) and (23f) are identical. Therefore both are optimal, accounting for the optionality of *that*. The reason why there is no violation of Pr-Pr in (23f) is that although the finite verb is in the highest X° of the embedded clause, it was not moved there. If the finite auxiliary had been moved to I°, the relevant

candidates would be (23b,e), and as (23e) will always be less optimal than (23b), the optionality of *that* would have been lost.

Tableau (23) also contains a third new constraint, which is ranked very highly:

θ-FI (Theta-Filter) punishes every argument which is not assigned a thematic role. I take thematic roles to be assigned by V° and P° inside their maximal projections (following Chomsky 1981 and Koopman & Sportiche 1991) and so, if a main verb is inserted directly under I°, then its arguments will not be assigned their thematic roles (indeed, the sentence will not contain a VP at all). To see how this actually works, we have to consider a case where the finite verb is a thematic verb (one that assigns one or more thematic roles). The example corresponding to (23c), i.e. which contains a finite main verb inserted under I°, (24c), will thus lose out to its rival candidate (24a), due to the effects of θ-FI (neither the subject nor the object is assigned a thematic role):

(24) En. I think that the actor (actually) saw the film

	θ-FI	Case	Pr-Pr	Ob-Hd	Ec-VP	Lx-Mv	V-V°	Pr-Bd	Stay	
a.  C _{[IP Su v} A _{[VP t saw Ob}					*			*	**	= (10d)
b. *C _{[IP Su saw A} [VP t v Ob					*	*!			**	= (10e)
c. *C _{[IP Su saw A} Ob	*!*						*			

Notice that θ-FI only punishes arguments which do not receive a thematic role, and not thematic roles which are not assigned to an argument. I think that the latter, unassigned thematic roles, may indeed exist in all four languages under discussion, e.g. when *have* or *do* are used as auxiliary verbs. Therefore a constraint punishing unassigned thematic roles may have to be ranked lower than θ-FI, which punishes the opposite, arguments which do not receive a thematic role.

Until now, the constraints discussed were part of the evaluation component of the grammar, EVAL (see section two of the introductory chapter), and so whether or not the optimal candidate violates a given constraint depends on the ranking of that particular constraint. Even if a given constraint is unviolated in a particular grammar, there is no guarantee that this is also the case in a different grammar. Another possibility is that a particular constraint is part of the candidate-generating component of the grammar, GEN. In this case it cannot possibly be violated by the optimal candidate, because GEN will not generate any candidates that violate it. If a constraint (e.g. θ -Fl) is part of GEN, it is never violated in any languages, because GEN does not vary between languages; language variation results from variation in the constraint ranking in EVAL.

θ -Fl prefers main (thematic) verbs to be inserted under V° , but says nothing about auxiliary verbs. $V-V^\circ$ prefers all verbs to be inserted under V° , and $Ec-VP$ prefers all verbs to be inserted under I° . If it is assumed that θ -Fl is unviolable (i.e. part of GEN), then main verbs are generated in V° in all languages, and only the point of insertion of auxiliary verbs (i.e. verbs that do not assign a thematic role) is subject to language variation: Either $V-V^\circ$ is ranked above $Ec-VP$, in which case all verbs are inserted under V° (Danish, French and Icelandic), or $Ec-VP$ is ranked above $V-V^\circ$, in which case non-thematic verbs are inserted under I° (English).

Only if non-thematic verbs may be inserted directly under I° (i.e. only under the English ranking) is *do*-support possible, because under the Danish/French/Icelandic ranking, also non-thematic verbs like dummy *do* would have to be inserted under V° , in which case such an insertion would not achieve anything: The reason why English has *do*-support with negation is that both V° -to- I° movement and I° -to- V° movement across negation may be avoided, see sections 6 and 7 below (*do*-support with movement to C° behaves in a parallel fashion, see Vikner 1998).

A reason for linking the auxiliary/main verb difference in English to thematic roles may be found in the following difference between auxiliary *have/do* and main verb *have/do*, (25) and (26): Only the auxiliaries occur in I° (as can be seen from their appearance left of negation, left of a sentential adverbial, left of a subject quantifier, in tag questions, and in VP-ellipses, Scholten 1988:3-7).

- (25) En. a. He actually did the dishes (main verb in V°)
 b. *He did actually the dishes (main verb in I°)
 c. He did not do the dishes (main verb in V°)
 d. *He did not the dishes (main verb in I°)
 e. *He did not do see the film (auxiliary verb in V°)
 f. He did not see the film (auxiliary verb in I°)

- (26) En. a. She actually had a lot of money (main verb in V°)
 b. *She had actually a lot of money (main verb in I°)
 c. She did not have a lot of money (main verb in V°, US)
 d. *She had not a lot of money (main verb in I°)

("somewhat uncommon", Quirk et al. 1985:131)

- e. She had not got a lot of money (verb: "get", UK, informal)
 f. ?She actually had earned a lot of money (auxiliary verb in V°)
 g. She had actually earned a lot of money (auxiliary verb in I°)
 h. *She did not have earned a lot of money (auxiliary verb in V°)
 i. She had not earned a lot of money (auxiliary verb in I°)

There are no such differences between auxiliary *be* and main verb *be*, both occur in I°:



- (27) En. a. ?She actually is very nice (*main verb in V°*, not stressed)
- b. She is actually very nice (*main verb in I°*)
- c. *She does not be very nice (*main verb in V°*)
- d. She is not very nice (*main verb in I°*)
- e. ?He actually is working (*auxiliary verb in V°*, not stressed)
- f. He is actually working (*auxiliary verb in I°*)
- g. *He does not be working (*auxiliary verb in V°*)
- h. He is not working (*auxiliary verb in I°*)

Roberts (1985:30) and Scholten (1988:160) suggest that only verbs that do not assign thematic roles may be inserted under I°. This gives the right prediction concerning main verb *be*, which presumably does not assign a thematic role (Haegeman 1994:68, Parsons 1995:645, Williams 1994:44), unlike main verb *have/do*, but like auxiliary *have/be/do*. This is captured here by taking θ -Fl to be inviolable (i.e. part of GEN), which will ensure that verbs that assign thematic roles are universally generated in V°.

Consider now what happens in Danish, where V-V° is ranked above Ec-VP:

(28) Da. Jeg tror (at) skuespilleren virkelig har set filmen

I think (that) actor-the really has seen film-the

	θ-Fl	Case	Pr-Pr	Ob-Hd	V-V°	Ec-VP	Lx-Mv	Pr-Bd	Stay
a.  C [IP Su v A [VP t has [VP t seen Ob						**		*	***
b. *C [IP Su has A [VP t v [VP t seen Ob						**	*!		***
c. *C [IP Su has A [VP t seen Ob					*!	*			*
d.  [IP Su v A [VP t has [VP t seen Ob						**		*	***
e. * [IP Su has A [VP t v [VP t seen Ob			*!			**	*		***
f. * [IP Su has A [VP t seen Ob					*!	*			*

(28c), which was optimal in English, loses out to (28a), because it is more important in Danish that every verb chain contains a V° (V-V°) than it is to keep the number of VPs to an absolute minimum (Ec-VP).


Pr-Pr is not relevant here: The optimal candidate without *at*, 'that', does not violate it, as there is no movement into the highest X° of an embedded clause. (28a) and (28d) are thus both optimal.

Consider finally what happens in French and Icelandic in (29). Here V-V° is ranked above Ec-VP (like Danish but unlike English):

(29) α. Fr. Je crois *(que) l'acteur a vraiment vu le film

β. Ic. Ég tel *(að) leikarinn hafi áreiðanlega séð myndina

I think (that) the actor has really seen the film

	θ- Fl	Case	Pr- Pr	Ob- Hd	V- V°	Ec- VP	Pr- Bd	Lx- Mv	Stay
a. * C _[IP Su v] A _{[VP t has [VP t seen Ob]}						**	*!		***
b.  C _{[IP Su has A [VP t v [VP t seen Ob]}						**		*	***
c. * C _{[IP Su has A [VP t seen Ob]}					*!	*			*
d. * [IP Su v A [VP t has [VP t seen Ob]						**	*!		***
e. * [IP Su has A [VP t v [VP t seen Ob]			*!			**		*	***
f. * [IP Su has A [VP t seen Ob]					*!	*			*

(29c), which was optimal in English, loses out to (29b), because it is more important in French and Icelandic (as in Danish) that every verb chain contains a V° (V-V°) than it is to keep the number of VPs as low as possible (Ec-VP).

Pr-Pr is relevant in (29), any version of the optimal candidate which lacks *que/að*, 'that', would violate Pr-Pr, as there would be movement into the highest X° of an embedded clause. Compare (29e) to (29b).

As already shown in (18), in French and Icelandic, Pr-Bd is ranked above Lx-Mv, whereas in Danish and English, Lx-Mv is ranked above Pr-Bd. When Pr-Bd is ranked above Lx-Mv, verbs inserted under V° move from V° to I° rather than stay in V°. Compare (29a) to (29b).

So far I have discussed the ranking between Ec-VP and V-V° independently of the ranking between Lx-Mv and Pr-Bd. The interaction between all four constraints is as follows:

If V-V° outranks Ec-VP, i.e. if all verbs are inserted under V°, then there are two relevant possibilities: The result is French/Icelandic (all finite verbs undergo V°-to-I° movement) if Pr-Bd outranks Lx-Mv, and it is Danish (no finite verbs undergo V°-to-I° movement) if Lx-Mv outranks Pr-Bd.

If Ec-VP outranks V-V°, i.e. if non-thematic verbs are inserted under I° and thematic verbs under V°, no variation is found when the input contains a non-thematic verb, as this verb will occur in I° regardless of the ranking of the other constraints.

Variation is found with Ec-VP outranking V-V° when there is no non-thematic verb in the input. There are three relevant possibilities: The result is English if Pr-Bd is outranked by the other three constraints (i.e. the finite thematic verb occurs in V°). The result would be a variant of English with V°-to-I° movement of thematic verbs if Lx-Mv were outranked by the other three constraints. Finally, the result would be another variant of English with *do*-insertion with thematic verbs in all contexts if V-V° were outranked by the other three constraints.

Until now two differences in constraint ranking have been assumed:

1. English/Danish differ from French/Icelandic as far as Pr-Bd is concerned (absence vs. presence of V°-to-I° movement).
2. English differs from the other three as far as V-V° is concerned (difference vs. no difference between thematic and non-thematic verbs, and also presence vs. absence of *do*-insertion):

(30) English:

θ-	Case	Pr-	Ob-		Ec-		Lx-V-	Pr-	Stay
Fl		Pr	Hd		VP		MvV°	Bd	

Danish:

θ-	Case	Pr-	Ob-	V-	Ec-		Lx-		Pr-	Stay
Fl		Pr	Hd	V°	VP		Mv		Bd	

French/Icelandic:

θ-	Case	Pr-	Ob-	V-	Ec-	Pr-	Lx-			Stay
Fl		Pr	Hd	V°	VP	Bd	Mv			



6. Negation, the HMC, and insertion of empty "do" in I°

For negated clauses, no further ranking differences between the languages are necessary, just one additional constraint, the Head Movement Constraint (HMC), which may be taken to be ranked immediately below Ec-VP in all four languages.

In earlier versions of this paper, I attempted to explain the different effects of negation in the four languages by saying that Neg° could not be part of the verb chain in English (*not* being an X° in Neg°), but the verb could move through Neg° in the other three (Danish *ikke*, French *pas*, Icelandic *ekki* being XPs in NegP-Spec). This could be achieved by two constraints "Neg=X°" and "Neg=XP" which would be ranked one way in English and the other way in the other three languages, but such an approach does not seem particularly insightful.

I now want to show that it is possible to use only constraints that are independently necessary: **V-V°** (as seen above) and the **HMC** (the Head Movement Constraint, see Travis 1984:131, Baker 1988:53, Rizzi 1990:11). I would like to suggest that the HMC is violated whenever Neg° intervenes in the verb chain in any of the four languages (i.e. when I° c-commands Neg° and Neg° c-commands V°). This is least controversial for English, where *not* has often been taken to be in Neg°, and therefore to block the formation of a chain between V° and I° (Pollock 1989:397, Roberts 1993:338, n21).

(31) En. I think (that) the actor did not see the film

		θ -Fl	Case	Pr-Pr	Ob-Hd	Ec-VP	HM-C	Lx-Mv	V-V $^{\circ}$	Pr-Bd	Stay
a.	*C [IP Su v not [VP t saw Ob					*	*!			*	**
b.	*C [IP Su saw not [VP t v Ob					*	*!	*			**
c.	*C [IP Su v not [VP t did [VP t see Ob					**!	*			*	***
d.	*C [IP Su did not [VP t v [VP t see Ob					**!	*	*			***
e.	 C [IP Su did not [VP t see Ob					*			*		*
f.	* [IP Su v not [VP t saw Ob					*	*!			*	**
g.	* [IP Su saw not [VP t v Ob			*!		*	*	*			**
h.	* [IP Su v not [VP t did [VP t see Ob					**!	*			*	***
i.	* [IP Su did not [VP t v [VP t see Ob			*!		**	*	*			***
j.	 [IP Su did not [VP t see Ob					*			*		*

HMC (Head Movement Constraint) punishes every X° which intervenes in an X° chain (i.e. which is c-commanded by a link of an X° chain and itself c-commands another link in the same X° chain) unless it is bound by the antecedent (the lexical element) of the chain. Neg° is taken to be such a potential intervening element when the verb moves across it. It does not matter whether the movement would have been downwards, compare (31a) to (31e), or upwards, compare (31b) to (31e). If (31) had not contained a negation, there would have been no *do*-insertion and (31a) would have been more optimal than (31e).

Ec-VP (Economy of VPs) means the fewer VPs, the better, hence the preference for a *do* which is inserted under I° rather than under V° . Compare (31d) to (31e).

I assume that *did* is in I° in (31e,j), an assumption based on the fact that it precedes *not*, and I also assume that it has been inserted directly under I° , as otherwise its presence in I° would violate Pr-Pr in (31j). This is how the optionality of *that* is achieved: With *did*



inserted under I° , the two candidates with and without *that* have the same constraint profile, see (31e) and (31j). If *did* had been inserted under V° , the two candidates with and without *that* would not have had the same constraint profile because of Pr-Pr, see (31d) and (31i).

Insertion of *do* under I° violates V-in V° , just like insertion of *have* or *be* under I° does. In all three cases, the verb will neither be in V° nor have a trace in V° (see also section 8 below). In English, however, this violation of V- V° has no consequences: The candidate that violates it, (31e), is still optimal, because the competing candidates all have more violations of higher ranking constraints than (31e).

Also for Danish, French and Icelandic, I take HMC to be violated when Neg° intervenes in the verb chain. However, in these three languages, such a violation is preferable to a violation of V- V° , i.e. a verb chain across Neg° is preferable to *do*-insertion. This is due to the independently motivated higher ranking of V- V° in these three languages than in English, see section 5 above. The HMC itself is ranked the same in all four languages, below Ec-VP and above Lx-Mv. This analysis presupposes that something like the stray affix filter (see Baker 1988:140) applies and is either very highly ranked or unviolable, i.e. that the verbal affix in I° has to be affixed to a verb.

(32) Da. Jeg tror (at) skuespilleren ikke så filmen

I think (that) actor-the not sawfilm-the

	θ-Fl	Case	Pr-Pr	Ob-Hd	V-V°	Ec-VP	HM-C	Lx-Mv	Pr-Bd	Stay
a.  C [IP Su v not [VP t saw Ob						*	*		*	**
b. * C [IP Su saw not [VP t v Ob						*	*	*!		**
c. * C [IP Su v not [VP t did [VP t see Ob						**!	*		*	***
d. * C [IP Su did not [VP t v [VP t see Ob						**!	*	*		***
e. * C [IP Su did not [VP t see Ob					*!	*				*
f.  [IP Su v not [VP t saw Ob						*	*		*	**
g. * [IP Su saw not [VP t v Ob			*!			*	*	*		**
h. * [IP Su v not [VP t did [VP t see Ob						**!	*		*	***
i. * [IP Su did not [VP t v [VP t see Ob			*!			**	*	*		***
j. * [IP Su did not [VP t see Ob					*!	*				*

V-V° (Verb in V°) punishes verb chains which do not include a V°.

Compare (32e), which was optimal in English, to (32a), which lost in English because it violates HMC.

Lx-Mv (No Movement of a Lexical Head) punishes *så*, 'saw', for leaving V°. Compare (32b) to (32a).

Ec-VP (Economy of VPs) means the fewer VPs, the better. Compare (32c) to (32a).



Pr-Pr is not relevant here, as the version of the optimal candidate which lacks *at*, 'that', does not violate it, as there is no movement into the highest X° of an embedded clause. (32a) and (32f) are thus both optimal.

In French and Icelandic, the picture is only slightly different, due to the different ranking of Pr-Bd, which also entails a difference with respect to the complementiserless version of the optimal candidate:

(33) α. Fr. Je crois *(que) l'acteur ne voit pas le film

β. Ic. Ég tel *(að) leikarinn sjái ekki myndina

I think (that) the actor sees not the film

	θ-Fl	Case	Pr-Pr	Ob-	V-V°	Ec-VP	HM-C	Pr-Bd	Lx-Mv	Stay
a. * C [IP Su v not [VP t saw Ob						*	*	*!		**
b.  C [IP Su saw not [VP t v Ob						*	*		*	**
c. * C [IP Su v not [VP t did [VP t see Ob						**!	*	*		***
d. * C [IP Su did not [VP t v [VP t see Ob						**!	*		*	***
e. * C [IP Su did not [VP t see Ob					*!	*				*
f. * [IP Su v not [VP t saw Ob						*	*	*!		**
g.  [IP Su saw not [VP t v Ob			*!			*	*		*	**
h. * [IP Su v not [VP t did [VP t see Ob						**!	*	*		***
i. * [IP Su did not [VP t v [VP t see Ob			*!			**	*		*	***
j. * [IP Su did not [VP t see Ob					*!	*				*

V-V° (Verb in V°) punishes verb chains which do not include a V°.

Compare (33e) (which was optimal in English) to (33b) (which lost in English because it violates HMC and Lx-Mv).

Pr-Bd (Proper Binding) punishes *voit/sjái*, 'sees', for staying in V°, which forces V°-to-I° movement. Compare (33a) to (33b).

Ec-VP (Economy of VPs) means the fewer VPs, the better. Compare (33d) to (33b).

Pr-Pr is relevant here, any version of the optimal candidate which lacks *que/að*, 'that', would violate it, as there would be movement into the highest X° of an embedded clause. Compare (33g) to (33b).

Let me finally point out that I do not necessarily want to suggest that the negative elements *pas*, *ikke*, and *ekki* have to be X° elements that c-command the VP. Rather, I only want to suggest that irrespective of the actual XP/X°-status


of the lexical element of negation, Neg° (like NegP-spec) may block the formation of a verb chain across it in any of the four languages.

This has often been assumed about NegP-spec, i.e. that irrespective of whether NegP-spec is lexically filled by a negation, it cannot be part of an A-bar-chain, cf. that English has negative islands even though *not/n't* is (or could be) an X° (Cinque 1990:80, Rizzi 1990:15). I want to say the same about Neg°, it cannot form part of the verb chain. Where exactly *not*, *pas*, *ikke*, and *ekki* actually are, in NegP-spec, Neg°, or somewhere else, is a different question (for a discussion of the possibilities with respect to Italian and related languages, see Zanuttini 1997:23, 101). This means that the above analysis is compatible with Roberts' (1999:fn 12) suggestion that X° can block XP-movements and vice versa as long as they have the same type of features (in this case operator-features).

7. Negation and insertion under I° of auxiliaries or of empty "do"

To account for clauses with negation and auxiliaries, no further constraints and no further differences in constraint ranking are necessary. It follows from the already suggested rankings that neither in English, nor in the other three languages, is *do*-insertion possible. Consider first English:

(34) En. I think (that) the actor has not see the film

	θ- Fl	Case	Pr- Pr	Ob- Hd	Ec- VP	HM C	Lx- Mv	V- V°	Pr- Bd	Stay
a. *C _{[IP Su v not [VP t has [VP t seen Ob}					**!	*			*	***
b. *C _{[IP Su has not [VP t v [VP t seen Ob}					**!	*	*			***
c. *C _{[IP Su v not [VP t did [VP t have [VP t seen Ob}					**!*	*			*	****
d.  C _{[IP Su has not [VP t seen Ob}					*			*		*
e. *C _{[IP Su did not [VP t have [VP t seen Ob}					**!			*		**

Adding *do* serves no purpose here (following Grimshaw 1997:384), as *have* may itself be base-generated in I° , so that the HMC is not violated anyway. This means that adding *do* will merely cost an extra VP, namely the VP of non-finite *have*, compare (34e) to (34d). Notice that this presupposes that auxiliaries like *have* in (34) are part of the input, as mentioned in section 2.


The complementiser is optional as any version of the optimal candidate without *that* would not violate Pr-Pr, because there is no movement into the highest X° of an embedded clause. (34d) and a version of (34d) without *that* would thus have the same (optimal) constraint profile.

Ec-VP punishes (34a,b) (optimal in Danish and in French/Icelandic respectively) which have more VPs than (34d). Because *have* does not assign any thematic roles, base-generating it in I° , as in (34d), will not violate θ -Fl, but only V- V° .

Consider next Danish, where there are no relevant changes compared to the negated finite main verb in (32):

(35) Da. Jeg tror (at) skuespilleren ikke har set filmen

I think (that) actor-the not has seen film-the


	θ -Fl	Case	Pr-Pr	Ob-Hd	V- V°	Ec-VP	HM C	Lx-Mv	Pr-Bd	Stay
a.  C _[IP Su v] not [VP t has] _[VP t seen Ob]						**	*		*	***
b. *C _[IP Su] has not [VP t v] [VP t seen Ob]						**	*	*!		***
c. *C _[IP Su v] not [VP t did] _{[VP t have [VP t seen Ob]}						***!	*		*	****
d. *C _[IP Su] has not [VP t seen Ob]					*!	*				*
e. *C _[IP Su] did not [VP t have [VP t seen Ob]					*!	**				**

V- V° punishes (35d) (optimal in English) because *har*, 'has', has no trace in V° , as opposed to (35a). Lx-Mv punishes (35b) (optimal in French/Icelandic) because *har*, 'has', leaves V° , as opposed to (35a). Ec-VP punishes (35c), which has more VPs than (35a). The complementiser is optional, as any version of the

optimal candidate which lacks *at*, 'that', would not violate Pr-Pr, because there is no movement into the highest X° of an embedded clause. (35a) and a version of (35a) without *at* would have the same (optimal) constraint profile.

Consider finally French and Icelandic. Also here there are no relevant changes compared to the negated finite main verb in (33):

- (36) α. Fr. Je crois *(que) l'acteur n'a pas vu le film
 β. Ic. Ég tel *(að) leikarinn hafi ekki séð myndina
I think (that) the actor has not seen the film

	θ-Fl	Case	Pr-Pr	Ob-Hd	V-V°	Ec-VP	HM-C	Pr-Bd	Lx-Mv	Stay
a. *C _{[IP Su v not [VP t has [VP t seen Ob}						**	*	*!		***
b.  C _{[IP Su has not [VP t v [VP t seen Ob}						**	*		*	***
c. *C _{[IP Su v not [VP t did [VP t have [VP t seen Ob}						***!	*	*		****
d. *C _{[IP Su has not [VP t seen Ob}					*!	*				*
e. *C _{[IP Su did not [VP t have [VP t seen Ob}					*!	**				**


V-V° punishes (36d) (optimal in English) because *a/hafi*, 'has', has no trace in V°, as opposed to (36b). Pr-Bd punishes (36a) (optimal in Danish), because *a/hafi*, 'has', stays in V°, as opposed to (36b). Ec-VP punishes (36c), which has more VPs than (36b). The complementiser is obligatory, any version of the optimal candidate without *que/að*, 'that', would violate Pr-Pr, as there would be movement into the highest X° of an embedded clause.

8. Constraints are never switched off

An important difference between OT and the "principles and parameters" framework (Chomsky 1986, Rizzi 1990) is that nothing in OT directly corresponds to parameters (see also section 3 in the introductory paper). In a certain sense, one might say that ranking V-V° above Ec-VP corresponds to a

particular setting of a particular parameter, namely one that makes thematic and non-thematic verbs behave alike in Danish, French and Icelandic. Correspondingly, one might expect that ranking V-V° below Ec-VP corresponds to a different setting of the same parameter, namely the one that forces a difference between thematic and non-thematic verbs in English. But whereas in principles and parameters terms, this might have meant switching off the parameter that forces a difference between thematic and non-thematic verbs, there is no switching off constraints in OT: Even in English where V-V° is ranked below Ec-VP, V-V° is still active.

(37) En. I think that the actor (actually) saw the film

	θ-Fl	Case	Pr-Pr	Ob-Hd	Ec-VP	HM-C	Lx-Mv	V-V°	Pr-Bd	Stay	
a.  C _{[IP} Su v A _{[VP} t saw Ob					*				*	**	= (10d)/(24a)
b. *C _{[IP} Su saw A _{[VP} t v Ob					*		*!			**	= (10e)/(24b)
c. *C _{[IP} Su did A _{[VP} t see Ob					*			*!		*	

V-V° must be ranked below Ec-VP to allow insertion directly under I° (of auxiliaries or dummy *do*), see e.g. (23c) or (31e) above. Nevertheless, V-V° is still active, in that it serves to exclude (37c), as nothing is gained by inserting *do* in I° here: It would not mean using one less VP, nor would it prevent movement out of V°, as (37a) also has only one violation of Ec-VP and no violation of Lx-Mv.

In addition to punishing the insertion of verbs directly under I°, V-V° thus also does the work here that Grimshaw's (1997:381) "Full Interpretation" (FI) does in her analysis, in that it punishes the use of dummy *do* (which she calls "light verb" *do*).

Corresponding to Grimshaw's (1997:386-387) suggestion for FI, V-V° may also have to be a gradient constraint, so that it would be violated to a lesser extent by dummy *do* than by dummy *divulge* or dummy *domesticate*:

Dummy *do*, dummy *divulge*, and dummy *domesticate* would all violate V-V° because (some of) their lexical/categorial properties would be ignored if they were not inserted under V°. However, the violation incurred by dummy *do* would be smaller than the violations incurred by other verbs, because *do* has fewer lexical/categorial properties than other verbs, and so when *do* is used as a dummy or light verb, there are fewer lexical/categorial properties which have to be ignored (exactly how to quantify this is anything but straightforward). This view would also necessitate that *do* be seen as having fewer lexical/categorial properties than even e.g. *be* and *have*, to ensure that the light verb used with negation (and in questions and other subject-auxiliary inversion contexts) is *do* rather than *be* or *have*.

The reason for the well-formedness of a string corresponding to (37c) with emphatic/ contrastive stress on *do* (i.e.... *that the actor DID see the film*) is precisely that something IS gained by inserting *do* here (i.e. by allowing a violation of V-V°), namely that a violation of the HMC is avoided, under the assumption (see Trachtenberg 1996) that emphasis is an independent functional head between I° and V°, much like Neg°, see (31a,e) in section 6. In e.g. Danish, where V-V° is ranked higher than HMC, the analysis correctly predicts that emphasis, like negation, does not make any difference for the position of the verb, (the example is... *at skuespilleren virkelig SÅ filmen*, `... that the actor really SAW the film'), see (32a,e) in section 6.

Finally another shortcoming of a putative "dummy *do*-parameter" should be noted, namely that it would lead us to expect that languages either do or do not have a dummy *do*. This is not the case: Although Danish, French and Icelandic do not have *do*-insertion the way English does, they all do have a so-called "verbum vicarium", i.e. a verb that substitutes for other verbs under certain circumstances. Furthermore, these verbs are the straight-forward translations of *do*: Icelandic *gera*, Danish *gøre*, and French *faire*.

- (38) a. Ic. Haltu á blýantinum eins og ég geri
 b. Da. Hold på blyanten ligesom jeg gør
 c. En. Hold the pencil as I do (it)
 d. Fr. Tiens le crayon comme je le fais

Exactly how to analyse these examples is not completely clear, however, as all of them would seem to involve optionality, in that "full" versions are also possible, e.g. *Hold the pencil as I hold it*. Presumably it would be possible to take the price of using *do* as a light verb (i.e. the price of ignoring some of its lexical/categorial properties) to be more than offset by the gain of not having to use a full VP-structure containing an object coreferent with an object in the matrix clause.

9. A brief sketch of the analysis of main clauses

Until this point, only two differences in constraint ranking have been assumed (viz. the two that were already shown in (30). In fact, the only difference compared to (30) is the introduction of HMC between Ec-VP and Lx-Mv in all four languages):

(39) English:

θ-	Case	Pr-	Ob-		Ec-	HM		Lx-	V-	Pr-	Stay
Fl		Pr	Hd		VP	C		Mv	V°	Bd	

Danish:

θ-	Case	Pr-	Ob-	V-	Ec-	HM		Lx-		Pr-	Stay
Fl		Pr	Hd	V°	VP	C		Mv		Bd	

French/Icelandic:

θ-	Case	Pr-	Ob-	V-	Ec-	HM	Pr-	Lx-			Stay
Fl		Pr	Hd	V°	VP	C	Bd	Mv			

Unfortunately, space does not permit a close examination of main clauses, which is also the reason why we will not see any syntactic differences between French and Icelandic. Such a close examination (see Vikner 1998) would reveal two important differences with respect to verb second: Whereas all four languages have V2 main clauses when the initial element is a *wh*-element, only Danish and Icelandic have V2 in non-*wh*-initial main clauses. A different difference concerns subject-initial main clauses, where English never has V2 (*Peter actually saw the film*), not even in questions (*Who actually saw the film?*).

To account for these different variations concerning V2, three more constraints would have to be introduced (following Grimshaw 1997 and Baković 1998): **Op-Sp** - operators in specifier position, **Wh-Sp** - *wh*-operators in specifier position, and **Op-Sc** - operators in scope position. Operators here include *wh*-operators and topicalised constituents. A scope position is any position which c-commands the IP.

V2 in main clause questions is driven by Wh-Sp, which forces non-subject *wh*-elements to move to CP-spec, see also the *wh*-criterion (Rizzi 1996:64, Müller 1997:263). V2 in non-*wh* main clauses is driven by Op-Sp, which forces non-subject operators (including both *wh*-operators and topicalised constituents) to move to CP-spec. Wh-Sp would thus be ranked highly in all four languages, whereas Op-Sp is ranked highly in Danish and Icelandic but low in English and French.

The movement of subjects to a position outside IP is forced in Danish, French, and Icelandic by the high ranking of Op-Sc, and avoided in English by a low ranking of the same constraint.

The V2 verb movement itself is driven by a different constraint, Ob-Hd, which also drives many other types of X^o-movement, see e.g. Grimshaw (1997) and Baković (1998). This is different from the *wh*-criterion, which drives both *wh*-movement and V^o-movement. Whether the verb undergoing V2 is *do* or a

verb also present in the input depends on the ranking, so that only English (with $Lx-Mv \gg V-V^\circ$) will insert *do* in such contexts. In the other three languages (with $V-V^\circ \gg Lx-Mv$), it is less expensive to move a verb out of VP than it is to insert a dummy verb.

In accounting for V2 in OT by means of the two constraints Wh-Sp and Op-Sp, it is possible to avoid the trap of binary parameters, whereby languages either have to have a particular property (e.g. V2) or not to have it. Languages may have a lot of V2, i.e. "real" V2 (Danish/Icelandic), or have just a little V2, i.e. "residual" V2 (English/French). Furthermore, by having Wh-Sp be a more restricted version of (i.e. a subset of) Op-Sp, the (hopefully correct) prediction is made that there are no languages which have V2 in topicalisations but not in questions.

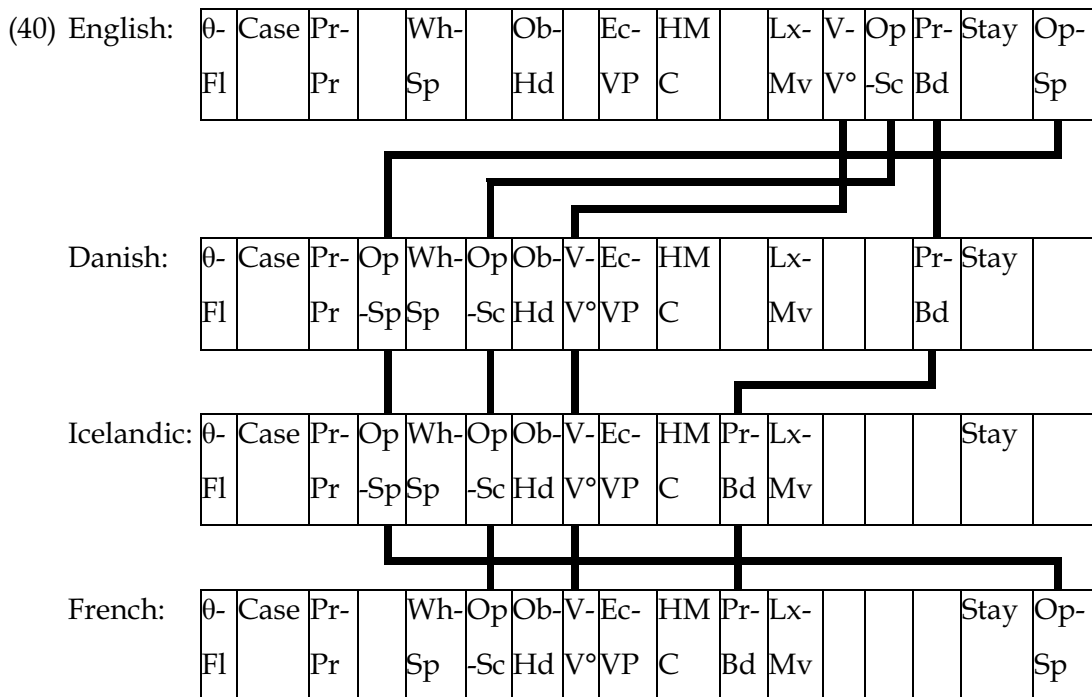
A more complete picture would thus have four different variations among the languages:

1. The variation concerning **Pr-Bd** accounts for the variation in V° -to- I° movement.

2. The variation concerning **V-V[°]** accounts for the variation concerning *do*-insertion and also for whether non-thematic verbs behave like thematic ones or not.

3. The variation concerning **Op-Sp** accounts for the variation in non-*wh* V2.

4. The variation concerning **Op-Sc** accounts for the variation in subject questions.



10. Conclusions

Optimality Theory Syntax allows the existence of constraints which are violated by the optimal candidate and nevertheless still rule out other candidates. Examples are **Ec-VP** (an implementation of economy of projection), **Stay** (an implementation of economy of derivation), and **V-V°**.

Optimality Theory Syntax also allows the existence of constraints which are both violated by the optimal candidate in some language and nevertheless still rule out other candidates in another language. Examples are **HMC**, **Lx-Mv**, and **Pr-Bd**.

This treatment of the **HMC** makes possible a unified treatment of NegP across the languages, in particular of the blocking effects of Neg° and NegP-spec.

Also made possible are principled accounts of the position of the finite verb in embedded clauses, of why *do* and other non-thematic verbs are inserted under I° in English, though not in the other languages, of why there is no *do*-

insertion with auxiliaries in English, and of when the complementiser *that/que/að/at* is optional and when it is obligatory.

The interaction between the four crucial constraints is as follows:

If $V-V^\circ$ outranks Ec-VP, i.e. if all verbs are inserted under V° , then there are two relevant possibilities: The result is French/Icelandic (all finite verbs undergo V° -to- I° movement) if Pr-Bd outranks Lx-Mv, and it is Danish (no finite verbs undergo V° -to- I° movement) if Lx-Mv outranks Pr-Bd.

If Ec-VP outranks $V-V^\circ$, i.e. if non-thematic verbs are inserted under I° and thematic verbs under V° , no variation is found when the input contains a non-thematic verb, as this verb will occur in I° regardless of the ranking of the other constraints.

Variation is found with Ec-VP outranking $V-V^\circ$ when there is no non-thematic verb in the input. There are three relevant possibilities: The result is English if Pr-Bd is outranked by the other three constraints (i.e. the finite thematic verb occurs in V°). The result would be a variant of English with V° -to- I° movement of thematic verbs if Lx-Mv were outranked by the other three constraints. Finally, the result would be another variant of English with *do*-insertion with thematic verbs in all contexts if $V-V^\circ$ were outranked by the other three constraints.

Because the facts discussed in this paper (the position of finite main verbs and of finite auxiliary verbs in embedded clauses with and without negation in Danish, English, French, and Icelandic) may now be seen as connected and deriving from common sources, and because their derivation in earlier (principles and parameters) accounts, e.g. Vikner (1995), involved a number of stipulations which are no longer necessary in the present analysis, I hope to have shown that Optimality Theory Syntax allows us to take a significant step forward in the analysis of verb movement.

Bibliography

- Alexiadou, Artemis: 1997, *Adverb Placement*, Johns Benjamins, Amsterdam.
- Baker, Mark: 1988, *Incorporation*, University of Chicago Press, Chicago.
- Baković, Eric: 1998, "Optimality and Inversion in Spanish" in Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds.), *Is the best good enough?*, MIT Press, Cambridge MA. pp. 35-58.
- Bobaljik, Jonathan: 1995, "Morphosyntax: The Syntax of Verbal Inflection", Ph.D., M.I.T.
- Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky, Noam: 1986, *Barriers*, MIT Press, Cambridge MA.
- Chomsky, Noam: 1991, "Some Notes on Economy of Derivation and Representation" in Robert Freidin (ed.), *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge MA, pp. 417-454.
- Chomsky, Noam: 1995, *The Minimalist Program*, MIT Press, Cambridge MA.
- Cinque, Guglielmo: 1990, *Types of A-bar-Dependencies*, MIT Press, Cambridge MA.
- Cinque, Guglielmo: 1999, *Adverbs and Functional Heads, A Cross-Linguistic perspective*, Oxford University Press, New York
- Deprez, Viviane: 1994, "A Minimal Account of the That-t Effect" in Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi, & Raffaella Zanuttini (eds.): *Paths Towards Universal Grammar, Studies in Honor of Richard Kayne*, Georgetown University Press, Washington DC, pp. 121-135.
- Emonds, Joseph: 1978, "The Verbal Complex of V'-V in French" in *Linguistic Inquiry* 9, 151-175.
- Emonds, Joseph: 1994, "Two Principles of Economy", in Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi, & Raffaella Zanuttini (eds.): *Paths Towards Universal Grammar, Studies in Honor of*

- Richard Kayne*, Georgetown University Press, Washington DC, pp. 155-172.
- Fiengo, Robert: 1977, "On Trace Theory" in *Linguistic Inquiry* **8.1**, 35-61.
- Grimshaw, Jane: 1991, "Extended Projection", ms, Brandeis University.
- Grimshaw, Jane: 1997, "Projection, Heads, and Optimality" in *Linguistic Inquiry*, **28.3**, 373-422.
- Haegeman, Liliane: 1994, *Introduction to Government and Binding Theory*, 2nd edition, Blackwell, Oxford.
- Haider, Hubert: 1988, "Matching Projections" in Anna Cardinaletti, Guglielmo Cinque, & Giuliana Giusti (eds.), *Constituent Structure: Papers from the 1987 GLOW Conference*, Foris, Dordrecht, pp. 101-121.
- Holmberg, Anders: 1997, "The True Nature of Holmberg's Generalization" in *Proceedings of NELS 27, GLSA*, Amherst MA, pp. 203-217.
- Holmberg, Anders & Christer Platzack: 1995, *The Role of Inflection in Scandinavian Syntax*, Oxford University Press, New York.
- Koopman, Hilda & Dominique Sportiche: 1991, "The Position of Subjects" in *Lingua* **85**, 211-258.
- Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer, & William Raymond: 1995, "Optimality and *Wh*-Extraction" in Jill Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers 18, GLSA, Amherst, pp. 607-636.
- Legendre, Géraldine, Paul Smolensky, & Colin Wilson: 1998, "When is Less More? Faithfulness and Minimal Links in *wh*-chains" in Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, & David Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge MA, pp. 249-289.
- McCloskey, James: 1992, "Adjunction, Selection and Embedded Verb Second", *Linguistics Research Report LRC-92-07*, University of California at Santa Cruz.

- Müller, Gereon: 1997, "Partial *wh*-movement and optimality theory" in *The Linguistic Review* **14**, 249-306.
- Müller, Gereon: 1999, "Optionality in Optimality-Theoretic Syntax" in *Glott International*, **4.5**, 3-8.
- Parsons, Terence: 1995, "Thematic Relations and Arguments" in *Linguistic Inquiry* **26.4**, 635-662.
- Pollock, Jean-Yves: 1989, "Verb Movement, Universal Grammar, and the Structure of IP" in *Linguistic Inquiry* **20.3**, 365-424.
- Prince, Alan & Paul Smolensky: 1993, "Optimality Theory", Technical Report 2, Rutgers University Center for Cognitive Science.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik: 1985, *A Comprehensive Grammar of the English Language*, Longman, London.
- Rizzi, Luigi: 1990, *Relativized Minimality*, MIT Press, Cambridge MA.
- Rizzi, Luigi: 1996, "Residual Verb Second and the *Wh*-Criterion" in Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads*, Oxford University Press, New York, pp. 63-90.
- Rizzi, Luigi & Ian Roberts: 1996, "Complex Inversion in French" in Adriana Belletti & Luigi Rizzi (eds.), 1996, *Parameters and Functional Heads*, Oxford University Press, New York, pp. 91-116. Reprinted from 1989, *Probus* **1.1**, 1-30.
- Roberts, Ian: 1985, "Agreement Parameters and the Development of English Modal Auxiliaries" in *Natural Language & Linguistic Theory* **3.1**, 21-58.
- Roberts, Ian: 1993, *Verbs and Diachronic Syntax*, Kluwer, Dordrecht.
- Roberts, Ian: 1999, "Head Movement", ms, University of Stuttgart, to appear in Mark Baltin & Chris Collins (eds.), *Handbook of Syntactic Theory*, Blackwell, Oxford.
- Samek-Lodovici, Vieri: 1996, "Constraints on Subjects: An Optimality Theoretic Analysis", Ph.D., Rutgers University.

- Scholten, Christina: 1988, "Principles of Universal Grammar and the Auxiliary Verb Phenomenon", Ph.D., U.C.L.A.
- Trachtenberg, Melissa: 1996, "Emphatic Affirmation in English and Danish", ms, Rutgers University.
- Travis, Lisa: 1984, "Parameters and Effects of Word Order Variation", Ph.D., M.I.T.
- Vikner, Sten: 1995, *Verb Movement and Expletive Subjects in the Germanic Languages*, Oxford University Press, New York.
- Vikner, Sten: 1997a, "The Interpretation of Object Shift, Optimality Theory, and Minimalism" in *Working Papers in Scandinavian Syntax* 60, 1-24.
- Vikner, Sten: 1997b, "V^o-to-I^o Movement and Inflection for Person in All Tenses" in Liliane Haegeman (ed.): *The New Comparative Syntax*, Longman, London. pp. 189-213.
- Vikner, Sten: 1998, "V2, V^o-to-I^o Movement, *do*-Insertion, and Negation in Optimality Theory", ms, University of Stuttgart.
- Williams, Edwin: 1994, *Thematic Structure in Syntax*, MIT Press, Cambridge MA.
- Zanuttini, Raffaella: 1997, *Negation and Clausal Structure*, Oxford University Press, New York.