Last Resorts: ATypology of *Do*-Support[1]
Jane Grimshaw
11/14/2010

1 Introduction

The notion of a 'last resort' figures prominently in linguistic discussion within the last 25 years.[2] The present investigation suggests that it is both fundamentally right and fundamentally incoherent. It can be right and coherent only within a theory of grammatical optimization, such as that provided by Optimality Theory (Prince and Smolensky 2004). In this theory, every grammatical candidate is a last resort, in a sense which is both theoretically and empirically definable. This paper explores the phenomenon of *do*-support, a frequently cited example of a last resort grammatical device, and argues that it is indeed a last resort, but so is every grammatical alternative. *Do*-support is nothing more than an option made available by universal grammar and selected in a particular configuration by universal constraints as ranked in a grammar. It is no more language-particular than, say, moving V to T. It is simply the best grammatical structure that the language can construct.

When some grammatical structure is said to be a 'last resort', what does it mean? It is a meaningful claim in cases where some definition allows alternatives to be systematically assessed. This is the case where a universal markedness hierarchy governs a phenomenon: the most marked member is guaranteed never to occur unless requirements such as faithfulness or conflicting markedness constraints force them to occur. (See the discussion of markedness in Prince and Smolensky 2004, for example.) However, no markedness hierarchy lies behind the last resort status of *do*-support. Perhaps a universal hierarchy in which *do*-support appears as the most marked member could be posited. This doesn't seem very promising, though, given that much current research holds that movement is a last resort.

Perhaps it is possible to state for each grammar what its last resort is. The problem is that a language can choose different 'last resorts' in different situations. Monnese, for example, as analyzed in Benincà and Poletto (2004), prefers V-raising to *do*-support when it is movement to T that is at issue, but prefers *do*-support to V-raising when movement to C is at stake. German (Bader and Schmid 2006) chooses V-raising to both T and C, but *do*-support with VP-preposing. English, I will argue, chooses T stranding in the future and *do*-support in the present and past. This is not a good fit with the concept of a last resort: a given choice either is, or is not, last. We do not want a theory that lists the choice of grammatical devices for each configuration or 'construction': English past tense interrogatives use *do*-support, Korean negatives use *do*-support, German negatives use V-raising, etc. The choices are not independent but are systematically related to each other, as explored in Anttila and Andrus 2006 and Prince 2006, and here. In fact the constraints investigated here, numbering 9 if we collapse the present and future tense constraints, generate a typology which includes only 28 languages, because most combinations of optima for negation, *wh* questions and VP displacement require inconsistent rankings and thus cannot co-occur in a single language. I will show, for example, that 'Anti-Monnese', a language which is just like Monnese except for having V-raising to C and *do*-support in T, cannot be derived given the universal constraints proposed here. Neither can 'Anti-German', a language with *do*-support winning in the negation and *wh* question c(andidate)-sets but V-raising in the VP displacement competition.

Within this perspective, *do*-support is no more language-particular than any other grammatical option. This contradicts the frequently found characterization of it as language-particular or unique to English. Lying behind the last resort theme is the assumption that *do*-support is in fact an odd phenomenon from some perspective. It is not found in positive declarative main clauses in neutral focus contexts, for example. However, far from being language-particular, *do*-support is in fact widespread. It is found in many languages and language families in one situation or another. In addition to the cases analyzed below, which come from Germanic, Romance and Panoan languages, many others have been discussed. French has documented cases of pleonastic *faire* (see Meinunger (2001), Miller (1997)). Celtic languages exhibit it also: see Newton (2009) on *do*-support in Old Irish complementizers, and Legendre (2001) and references therein on *do*-support in Breton. Korean exhibits *do*-support with the negative *ani* (see Hagstrom (1996)). Hagstrom also shows that the 'nominalizer' *ci* appears with *do*-support. Japanese 'emphatics' trigger *do*-support (Kuroda (1965), Miyagawa (2001)). Korean and Japanese offer a further case, rather different in structure from the ones considered here, analyzed e.g. in Grimshaw and Mester (1988), Saito and Hoshi (2000). These are composed from a noun which bears thematic argument structure, and an outer framework consisting of tense and the light verb *do* (*suru* in Japanese and *ha* in Korean). Abstracting away from many properties of these structures discussed in the literature, we can see that in this case, *do*-support again provides a verb for tense to associate with.

Within English there are numerous instances of *do*-support beyond its use in interrogatives and negatives. It appears in positive imperatives, alternating with a form without *do*: *Do come in*; *Come in.* It appears in VP ellipsis. It appears emphatically, e.g. in contradictions: *They said I hadn't paid but I DID pay*. It occurs in evasive answers to questions: *Is he a good doctor? Well he does have a lot of patients.* Its presence is a response to a cluster of prosodic and syntactic factors, as well as polarity and focus. It appears to be a favourite resource of the language, rather than a last resort. (It is interesting in this context to note that children learning English use *do*-support with enthusiasm: see Hollebrandse and Roeper 1996.)

Of course *do*-support is irreducibly language-particular in one sense: it involves a particular morpheme of a particular language. But it is language-particular only in the most shallow way: the choice of a morpheme is predictable given the vocabulary of the language. In all cases discussed here the verb used in support is the same morpheme as the verb translating *do*. Setting the morpheme itself aside, there is no evidence I can see to favour the view that *do*-support has a special status vis à vis other grammatical options like verb movement. Both are language-particular choices: they depend on constraint ranking. Both are universal: they follow from the interaction of universal constraints. The last resort status of *do*-support is both genuine and spurious: a *do*-support candidate will always have competitor(s) which satisfy constraints violated by the presence of *do*. In this sense *do*-support is a genuine last resort: the candidate wins if and only if all the alternatives lose. On the other hand, the same is true for any other winning candidate. They are all last resorts.

In sum, the notion 'last resort' is entirely central to grammatical theory. Every optimal candidate is a last resort, the best in the circumstances, the last candidate standing. The set of last resorts is identical to the set of potential winners – candidates which are not harmonically bounded. (A candidate is harmonically bounded if for every ranking of the constraints, there is a better candidate; see Samek-Lodovici and Prince 2005.)[3] The fact that every language does the best it can with what it has available is a result of the very way that grammaticality is decided. Grammaticality itself is irreducibly comparative and a theory of how candidates are compared is required to render the concept coherent.

OT does not include 'last resort' principles or constraints, or lists of last resort 'devices'. The last resort property of optima is a direct consequence of the theory of grammar, and does not need to be stipulated by auxiliary external statements. The work of giving last resort effects, of which economy effects are one example, is done by the theory of constraints and constraint interaction. This is the argument of Grimshaw (2001, 2002) with respect to economy of structure and Grimshaw (2006a) with respect to chains, i.e. movement.

The massive literature which appeals to last resorts contains several points that are quite striking and support the idea that the concept of a last resort lacks theoretical content in the absence of a theory of constraint conflict, candidate sets and optimization. First of all, the frequent use of scare quotes in the punctuation of 'last resort' suggests a certain lack of commitment on the part of the writers. (In Lasnik, Uriagareka and Boeckx (2005: 269) the term appears in this form five times in the penultimate paragraph.) Second, while last resort analyses are extremely common, they generally appeal to the last resort idea by word and not by deed. For instance, in many variants of 'movement as a last resort' the theory says that if a particular kind of feature is present movement is required. If it is absent, movement is disallowed. If we take away the 'last resort' concept from this, we get exactly the same results. In fact, if we claim that no-movement is a last resort and is possible only when a particular (type of) feature is absent, we get exactly the same result, at least in the absence of a substantive theory of features. The words 'last resort' are employed but the concept plays no role in the analysis. An example that is directly relevant here is the Benincà & Poletto (2004) account of *do*-support in Monnese, in which the presence of a feature on C forces a head to be filled and the head cannot be filled by a lexical V. The theory simply imposes these two requirements. It in no way relies upon or uses a special status for *do*-support, however encoded.

Finally, I would point out that researchers discussing the last resort 'flavor' of an analysis often resort to language which is precisely that of OT, and refers to concepts defined within OT. Benincà & Poletto (2004: 79) comment of English and Monnese: 'Hence, both languages react to the same tension between two conflicting constraints, the necessity to check a strong feature and the fact that the position hosting the feature has become thematically opaque in Pollock's terms; for both languages the problem is solved by inserting a dummy verb.' The language of constraint conflict and competition between candidates is appealing in conceptualizing the problem, despite the fact that the proposed analysis involves neither. A similar point is made in Samek-Lodovici (to appear): constraint conflict is proposed in theories which do not in fact have a way to accommodate it.

In sum:

- The notion of a last resort is explicated only by a theory of optimization with constraint interaction.
- The last resort character of grammatical structures is attributable to the determination of well-formedness itself, not to additional principles.
- Every grammatically possible structure is a last resort. The set of last resorts is the same as the set of possible winning candidates (i.e. grammatical sentences).
- *Do*-support is a last resort in exactly the same sense as any other grammatical strategy is. It's the best that can be done in the circumstances.

2 The cost and benefit of *do*-support

Certain grammatical demands can be satisfied by the presence of *do*: it can carry tense and it can fill a head position. The cost of *do* is that is has no meaning. Like other expletives such as *it* and *of* it appears to be the least-marked member of its category. (See Halle and Marantz (1993), Grimshaw (1997), Schütze (2001).) Many theories of verb meaning (e.g. Dowty 1979) posit a semantic prime which corresponds closely to the meaning of *do* and corresponding predicates in other languages. As mentioned above, in all cases to be analyzed here it is a verb with the same morphological realization as the verb translated as *do* that is recruited for support. The idea that some languages have the right kind of *do* to allow *do*-support, while others do not, plays no role in this explanation of the distribution of *do*, as I argued in Grimshaw (1997). The presence or absence of *do* in grammatical sentences is a consequence of the grammar of the language, not a matter of lexical stipulation.

The constraints which evaluate *do*-support and its competitors are in (1).

(1)	FULLINT	No meaningless element
	V+T	T is a bound affix[4]
	OBHD	A phrase has a head
	*LINF	No lexical head within a functional head[5]
	*LINA′	No lexical head within an A′ functional head

CCOM	The overt member of a chain c-commands null copies
UNIQUE	An element in the input has no more than one correspondent in the output

I assume here that the constraints FULLINT, V+T, *LINF, and *LINA′ are violated at most once per chain, by the head of the chain. Thus, for example, a chain containing *do*+T and a copy receives one violation of FULLINT and not two.

Two additional constraints divide the set of VP displacements into those in which VP occupies a specifier projection and those in which it is adjoined. This is discussed further in Section 4.

(2) VPSPEC A VP is a specifier
VPADJ A VP is an adjunct

Inputs are represented in a highly simplified form, specifying that they contain a verb, a subject and a tense. An internal argument ('int') is included in the inputs for VP displacement, because its position shows the location of the VP if the V has moved outside it. Since all candidates under consideration are faithful (aside from the unfaithful character of chains, which violate UNIQUE) no faithfulness constraints are at play, so such inputs suffice for present purposes. The order in which elements appear in the input is not significant, since it is the constraints themselves that determine the order of outputs, although the constraints responsible are not discussed here.

These constraints are chosen in part because their properties have been explored in earlier work. The constraint *LINA′ is important in distinguishing a lexical head raised to C, which violates it, from a lexical head raised to T, which does not. Quite possibly this constraint and *LINF are members of a set of such constraints which together prefer the lexical head to be as close to the lexical projection as possible. OBHD was posited in Grimshaw (1997) to force T-to-C movement in interrogatives. For the present purpose it is not crucial that this is the constraint at issue as opposed to one relating to feature specifications as in, for example Pesetsky and Torrego (2001, 2007).UNIQUE is a faithfulness constraint based on McCarthy and Prince (1997/1999). Grimshaw (2006a) argues that this constraint entails economy of movement effects, which thus reduce to the theory of faithfulness. (See Grimshaw (2001, 2002) for an argument that other constraints can entail economy of movement under certain circumstances.)

Finally, it is possible that FULLINT is not required, given the argument in Grimshaw (2001, 2002) regarding economy of structure. The presence of *do* will necessarily induce violations of the set of alignment constraints (HDLFT, SPECLFT and COMPLEFT) and it is possible that these violations are sufficient to restrict the distribution of *do*. However, this has not yet been proved. A different view of the constraint violated in *do*-support is put forward in Vikner (2001). He proposes that *do*-support violates a constraint VV^0, which is violated by any verb that is not inserted in V (the constraint requires verb chains to include a V^0). Since *do* is inserted directly into T when it is present for *do*-support, this constraint is violated. A final point: I do not try to analyze restrictions on *do* as due to a faithfulness constraint. This is because occurrences of *do* would not be penalized by such a constraint if *do* were in the input. Given 'richness of the base' (Prince and Smolensky 2004), there is no way to exclude it from inputs entirely. Nor is there any need to, as far as I can see.

3 Which resort is the last?

Cross-linguistically *do*-support is often associated with negation and interrogatives generated by *wh*-movement. This is because these configurations challenge the normal relations between T and V, and between head positions in the verbal extended projection. The present analysis proposes, following most literature on the topic, that negation can cause *do*-support because it separates the verb from tense.[6] See e.g. Chomsky (1957, 1991), Halle & Marantz (1993), Bobaljik (1995), Grimshaw (1997), Vikner (2001). *Wh*-movement induces the presence of a head position in the structure which needs to be filled.

There are several different ways to resolve such crises. One is to raise a main verb out of its position inside VP and one is to use *do*-support. The third option is to strand the Tense and allow it to be separated from V syntactically. (A fourth option is to have T lower: see Vikner 2001). The possibility that such structures might be grammatically licit was drawn to my attention by Elías-Ulloa (p.c.). It seems likely that structures in which T or other inflectional material is stranded may be widely found, but not analyzed as involving stranding, especially in theories which do not admit constraint violation, in which stranding should be ungrammatical. I propose here that the English future form *will* is an instance of a stranded tense, rather than in any real sense an auxiliary verb. (I assume that similar analyses extend to the other "modals", but do not examine this issue here.) English thus shows a non-uniform distribution in *do*-support, like German and Monnese. The present and past tenses invoke *do*-support, and the future does not, and thus appears as a free morpheme, in violation of the constraint V+T. Note that languages in which Infl heads take complements other than VPs (such as Chamorro as analyzed in Chung (1990), where VP, AP, DP/NP and PP can all occur as complements to Inflection) necessarily violate V+T whenever the complement is not a VP. The analysis of the English future is developed in Section 5. Until that point I will discuss only the patterns of *do*-support found with present and past tenses in English.

The three options, which I will abbreviate for the remainder of the paper as DS (for *do*-support), VR (for verb raising) and TS (for tense stranding), form the core of the candidates in the Negation c-set that I will analyze. Can one of them be identified as a last resort in some meaningful way? This section shows that within a single language, one of the three options can be selected in one structure and another elsewhere. We will see that this follows from the interaction of the constraints in (1). Later, in Section 6, I will show that not every combination of optima is possible. Rather the constraints define a restricted and systematic typology.

The Northern Italian dialect spoken in Monno, reported in Benincà and Poletto (2004),which I refer to as 'B&P' from this point on, prohibits the expletive verb *fa* 'do' from appearing in negated clauses (B&P: 70):

(3) (a) I so mìa
It I know not
(b) *fo mìa savè-l
I do not know-it

B&P hypothesize that main verbs raise to T, observing that V always raises across low adverbs as in (4), (B&P (7a) p 59), and hence across *mìa*, the negation, as shown by (3) above.

(4) l tʃàkola semper
he speaks always
'He always speaks'

In the terms of the present analysis, Monnese thus presents a case in which a constraint ranking allows violation of *LINF, and hence allows V-to-T movement.

German patterns like Monnese, as shown in (5) and (6).[7] (These examples were provided by Fabian Heck.) The main verb precedes negation and low adverbs like *immer*.

(5) (a) Ich weiß es nicht
I know it not
(b) *Ich tu(e) es nicht wissen
I do it not know
'I do not know it'

(6) (a) Er spricht immer
'He speaks always'
(b) *Er immer spricht
'He always speaks'

In contrast, English, in accordance with the hypothesis first put forward in Pollock (1989), does not allow V-to-T movement, and thus disallows a lexical V before a low adverb and requires *do*-support with negation.

(7) (a) *I know it not
(b) I do not know it

(8) (a) He always speaks
(b) *He speaks always

The analysis of the Negation c-set that I present here is based on the analysis of English in Grimshaw (1997), but restricted in scope and limited to very small candidate sets. Vast, in fact infinite, numbers of candidates cannot be generated in this system, because they are harmonically bounded. This argument has been made in earlier relevant literature and I will not repeat it here: see in particular Grimshaw (1997), Vikner (2001). Grimshaw (in prep.a) is a supplement to the present work, which contains the analysis of important additional candidates. This includes those which contain auxiliary *have* or the verb *be*, where *do*-support is never possible because such candidates are harmonically bounded. Similarly, every candidate containing more than one occurrence of *do* is harmonically bounded by an otherwise matching candidate containing only one. This is also demonstrated in the supplement. A similar argument with respect to multiple occurrences of the complementizer *that* in English is developed in Grimshaw (2008).

The table in (9) shows the number of violations that each candidate incurs on the five constraints that are relevant in assessing present and past tense forms. (I include OBHD and *LINA′ in the table because they can be violated in interrogatives, which are discussed later in this section. The constraints CCOM, VPSPEC and VPADJ are unviolated in the negation and interrogative systems, and are omitted.) The first candidate, which is optimal in German and Monnese, raises V to T, in violation of *LINF. The second candidate, which is optimal in English, violates FULLINT by virtue of the presence of *do*. The third candidate leaves tense stranded, violating V+T.

(9) Violation table: Negation

Input: {V Subj T Neg}		V+T	OBHD	FULLINT	*LINA′	*LINF	UNIQUE
a. VR	$[_{TP}$ Subj V-T Neg $[_{VP}$ <V> ...]]					1	1
b. DS	$[_{TP}$ Subj do-T Neg $[_{VP}$ V ...]]			1			
c. TS	$[_{TP}$ Subj T Neg $[_{VP}$ V ...]]	1					

The candidates and the violation profiles are necessarily identical in all three languages. The only difference lies in the ranking and hence the optimum. (I gloss over many important issues such as exactly where negation is located, whether it is a head and so forth.) The three comparative tableaux (Prince 2002) in (10)–(12) show how the optima are selected, and allow us to determine the rankings that generate the languages. All comparative tableaux in this paper have the winning candidate in the first row, bolded, and annotated with ‘W’. When the winner is the VR optimum, *LINF and UNIQUE prefer the loser in both competitions, and they must both be subordinated to both V+T and FULLINT.

(10) VR-Neg winner

Input: {V Subj T Neg}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
a. W [$_{TP}$ Subj V-T Neg [$_{VP}$ <V> ...]]						
b. DS [$_{TP}$ Subj do-T Neg [$_{VP}$ V ...]]			W		L	L
c. TS [$_{TP}$ Subj T Neg [$_{VP}$ V ...]]	W				L	L

In the competition below, between the DS winner and the TS candidate, FullInt prefers the loser. Since only V+T prefers the winner, V+T must dominate FullInt. In the competition between DS and VR, *LinF and Unique prefer the winner. One of these must dominate FullInt, since it prefers the TS candidate to the DS winner.

(11) DS-Neg winner

Input: {V Subj T Neg}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
b. W [$_{TP}$ Subj do-T Neg [$_{VP}$ V ...]]						
a. VR [$_{TP}$ Subj V-T Neg [$_{VP}$ <V> ...]]			L		W	W
c. TS [$_{TP}$ Subj T Neg [$_{VP}$ V ...]]	W		L			

When the optimum has a stranded tense as in (12), V+T prefers both of the losing candidates over the winner. Hence FullInt must dominate V+T in order for TS to beat DS, and either Unique or *LinF must dominate V+T in order for TS to beat VR. The TS candidate is the winner in the English future tense as discussed in Section 5.

(12) TS-Neg winner

Input: {V Subj T Neg}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
c. W [$_{TP}$ Subj -T Neg [$_{VP}$ V ...]]						
a. VR [$_{TP}$ Subj V-T Neg [$_{VP}$ <V> ...]]	L				W	W
b. DS [$_{TP}$ Subj do-T Neg [$_{VP}$ V ...]]	L		W			

The rankings in (13) are thus required to select the optima from the Negation c-set.

(13) VR-Neg FullInt, V+T >> *LinF, Unique
DS-Neg V+T >> FullInt *LinF or Unique >> FullInt
TS-Neg FullInt >> V+T *LinF or Unique >> V+T

The violation table and rankings already suggest that there is no asymmetry in the status of the three candidates. Each is both ‘better’ and ‘worse’ than the others. Each violates at least one of the constraints under discussion. Each option is chosen by several rankings (all rankings that are consistent with those in (13)). There is, as far as I can see, no basis for identifying the solution of *do*-support, as opposed to stranding or V-to-T movement, as a last resort.

This point is further sharpened when we pursue a prediction of this system: the theory predicts the possibility of a language with *do*-support in interrogatives but not with negation. In such a language is *do*-support a last resort or not? The reason for the non-uniform pattern is that there is a stringency relationship (Prince 1997) between two of the constraints posited in (1). *LINA′ is violated by a high lexical head in a projection with A-bar properties, such as CP. However, *LINF is violated there *and* in functional projections with A properties such as those in the IP 'region'. Hence *LINF is violated wherever *LINA′ is violated, but not vice versa. In particular, a V in T violates *LINF but a V in C violates *LINF and *LINA′.

B&P demonstrate that Monnese is a language which has a non-uniform pattern of *do*-support. In interrogatives it patterns with English, instead of with German, having *do*-support in matrix questions. The grammaticality of *fa*-support is illustrated for *wh* interrogatives in (14). (B&P (1a) p 52)

(14)	(a)	ke	fa-l	majà?
		What	does-he	eat
	(b)	*ke	majà-l?	
		What	eats he	

The corresponding German examples in (15) were provided by Fabian Heck.

(15)	(a)	Was	isst	er?	
		what	eats	he	
	(b)	?*Was	tut	er	essen?
		what	does	he	eat

(16) (a) What does he eat?
(b) *What eats he?

Again we begin by examining the violation table for the candidates under consideration. There are two decisions to be made in selecting the optimum within the c-set. The first, the focus of this paper, is the choice between the V-raising, *do*-support and tense-stranding structures. The second is the location of the stranded T, the raised V or the *do*, which might be in T or in some higher head position to the left of the subject. Here I consider the candidates with VR to C as in German, with C filled by *do*-support, and with tense in C, the TS candidate. I have included a fourth candidate as a representative of the class of candidates with an empty C position, which include candidates with and without V-raising, *do*-support and a stranded T. Candidate d. in (17) has V-raising to T, but has no further raising. (Candidates with an empty C are optimal in, for example, subordinate clauses in both English (Grimshaw (1997, 2006c)) and Monnese (B&P).) Since none of the systems we are examining allows the C to be empty in the cases at hand I have not included the comparative tableau and ranking information for this candidate as winner.

(Grimshaw (in prep. a.) includes the candidates with empty C, candidates with *have* or *be*, and those with multiple occurrences of *do*. Since the WhQ c-set there is larger than the one presented here, it establishes some rankings above and beyond what can be determined from (18), (20) and (22)).

The empty C candidates here and in Section 4 motivate the constraint Unique, which is violated (potentially more than once) by every chain. It thus prefers candidates with empty C positions to those in which the C position is filled by movement. Without Unique, candidates with VR, TS and DS in the T position are harmonically bounded by those with VR, TS and DS in the C position, which do not violate ObHd. Since I am making the simplifying assumption that all *wh*-movement candidates have a *wh* chain of the same length, and that the same holds for the VP displacement candidates in Section 4, I do not include violations of Unique caused by these chains in the analysis.

(17) Violation table: *wh* Questions

Input: {V Subj T Wh}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
a. VR [$_{CP}$ Wh V-T [$_{TP}$ Subj <V-T> [$_{VP}$ <V> …]]]				1	1	2
b. DS [$_{CP}$ Wh do-T [$_{TP}$ Subj <do-T> VP]]			1			1
c. TS [$_{CP}$ Wh T [$_{TP}$ Subj <T> VP]]	1					1
d. ØC [$_{CP}$ Wh __ [$_{TP}$ Subj V-T [$_{VP}$ <V> …]]]		1			1	1

The critical point to note is that the VR candidate violates both *LinF and *LinA′, while the DS and TS candidates violate neither. Again, I present the comparative tableaux for the VR, DS and TS candidates as winners.

(18) VR-WhQ winner

Input: {V Subj T Wh}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
a. W [$_{CP}$ Wh V-T [$_{TP}$ Subj <V-T> [$_{VP}$ <V> …]]]						
b. DS [$_{CP}$ Wh do-T [$_{TP}$ Subj <do-T> VP]]			W	L	L	L
c. TS [$_{CP}$ Wh T [$_{TP}$ Subj <T> VP]]	W			L	L	L
d. ØC [$_{CP}$ Wh __ [$_{TP}$ Subj V-T [$_{VP}$ <V> …]]]		W		L		L

The rankings required to select the VR winner are those in (19), which in fact select the VR winner in both c-sets discussed so far, i.e. they correspond to German.

(19) VR-WhQ ObHd >> *LinA′, Unique
FullInt, V+T >> *LinF, *LinA′, Unique

The TS comparative tableau establishes the rankings in (21).

(20) TS-WhQ winner

Input: {V Subj T Wh}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
c. W [$_{CP}$ Wh -T [$_{TP}$ Subj <-T> VP]]						
a. VR [$_{CP}$ Wh V-T [$_{TP}$ Subj <V-T> [$_{VP}$ <V> …]]]	L			W	W	W
b. DS [$_{CP}$ Wh do-T [$_{TP}$ Subj <do-T> VP]]	L		W			
d. ØC [$_{CP}$ Wh __ [$_{TP}$ Subj V-T [$_{VP}$ <V> …]]]	L	W			W	

(21) TS-WhQ FullInt >> V+T
ObHd or *LinF >> V+T
*LinA or *LinF or Unique >> V+T

The comparative tableau for the DS optimum motivates the rankings in (23):

(22) DS-WhQ winner

Input: {V Subj T Wh}	V+T	ObHd	FullInt	*LinA′	*LinF	Unique
b. W [$_{CP}$ Wh do-T [$_{TP}$ Subj <do-T> VP]]						
a. VR [$_{CP}$ Wh V-T [$_{TP}$ Subj <V-T> [$_{VP}$ <V> …]]]			L	W	W	W
c. TS [$_{CP}$ Wh -T [$_{TP}$ Subj <-T> VP]]	W		L			
d. ØC [$_{CP}$ Wh __ [$_{TP}$ Subj V-T [$_{VP}$ <V> …]]]		W	L		W	

(23) DS-WhQ V+T >> FullInt
ObHd or *LinF >> FullInt
***LinA′ or *LinF or Unique** >> FullInt

The rankings which select the DS candidate for negation are repeated in (24):

(24) DS-Neg V+T >> FullInt
***LinF or Unique** >> FullInt

We can see that the DS-WhQ rankings do not enforce the choice of DS as the optimum in the Negation c-set. This is because the Negation c-set ranking requires *LinF or Unique >> FullInt, while the WhQ competition is won by DS if one of these two constraints or *LinA′ dominates FullInt. (The critical parts of (23) and (24) are bolded.) Hence the DS winner in WhQ is compatible with a ranking in which DS does not win in the negation c-set, namely one in which *LinA′ dominates FullInt but neither *LinF or Unique does. This is the Monnese ranking. In Monnese, *do*-support is preferred when it avoids a violation of *LinA′, even though it is not preferred when just a violation of *LinF or Unique is at issue. Monnese does not tolerate a lexical verb in C even though it tolerates one in T.

The case of Monnese demonstrates that it is not possible to declare a 'last resort' for a language. If *do*-support is the last resort for English, and not for German, then what can we usefully say about Monnese? The fact is simply that the Monnese grammar chooses *do*-support for interrogatives but not for negation, because its constraint ranking gives the *do*-candidate as the optimum in interrogatives only. In contrast, English prefers *do* in both cases, and German prefers to raise the lexical verb in both cases. It is interesting to note that the non-uniform distribution of *do*-support in Monnese leads Culicover (2008: 32) to consider that it is not a real case of *do*-support: 'However, as Benincà & Poletto (2004) suggest, *do*-support in Monnese can be seen as a type of light-verb construction, which has generalized to most of the lexicon and which is restricted to questions.'[8] The fact that it is restricted to questions is simply a consequence of the ranking, in the present account. It is not necessary to appeal to a construction-related restriction, or to separate cases into 'real' *do*-support versus *do*-periphrasis, as Culicover does.

Since the existence of non-uniform systems, such as that of Monnese, is predicted by the constraint system, the obvious question is whether all non-uniform systems are possible. What is the typology of choices between DS, VR and TS? Are we predicting the existence of 'Anti-Monnese', a language with *do*-support in T (i.e. with negation) but with a lexical V occurring in the C position? The answer is that Anti-Monnese cannot be generated, as I will show in Section 6. No language can pick *do*-support as the optimum in negation and V-raising as the optimum in *wh* questions. Before moving to the typology, I add another interesting case of non-uniformity to the system, this time from German.

4 VP Displacement

When a VP is displaced, tense is again potentially separated from the verb, and here again we find all three of the familiar resolutions of the crisis, with the ranking of the constraints in (1) and (2) critical in determining the optimal structure for a given language. An example of the *do*-support choice is found in English examples such as (25).

(25) He was supposed to read the book and *read the book he will/did*.

As for *wh*-movement above, I will consider only candidates in which the VP is displaced, and thereby set aside the constraints responsible for the displacement, which presumably refer to the discourse status of the VP. I separate candidates with displaced VPs into those in which the VP is displaced to a specifier position, and those in which it is in an adjoined position. The reason for positing both structures is the existence of languages, such as English, where *wh*-movement (and negative preposing; see Grimshaw (1997) for an analysis) results in the raising of an auxiliary verb across the subject, but VP-preposing (and topicalization) does not.[9] The interrogative and displaced VP in (26) differ precisely in this respect.

(26)	Read the book *he will/did*.	*Read the book *will/did* he.
	Which book *will/did he* read?	*Which book he *will/did* read?

Why is inversion observed in interrogatives and not in VP-preposing? Here I follow the line of reasoning that this reflects a difference in structure. When the displaced phrase is a specifier, the tensed verb raises to the head of the projection which contains that specifier. We will see that this happens in German. In English the VP is adjoined (like a topicalized phrase; Zepter (2000)), and hence has no associated head position. Therefore, OBHD is not violated in the grammatical sentences in (25) and (26). For the sake of the argument, let us assume that the choice between the specifier analysis of the displaced VP and the adjunct analysis is made by the ranking of the two constraints in (2), VPSPEC and VPADJ. Of course, something much more interesting may well be involved, but this analysis will suffice for present purposes. These constraints will partition the VP displacement candidate set into two sub-c-sets, which we can now examine separately.

None of the constraints discussed here is sensitive to whether the displaced VP is on the left or the right, so I will only show leftward displacement in tables and tableaux.[10]

4.1 Adjoined VPs

Let us consider first the sub-c-set which satisfies the constraint VPADJ and violates the constraint VPSPEC. (27) is the violation table for the three candidates of interest. ‘Int’ is the internal argument of the verb: its position makes it possible to detect the VP displacement even if the V has raised outside the VP.

(27) Violation Table: Adjoined VPs

Input: {V int Subj T}	V+T	OBHD	FULLINT	*LINA′	*LINF	CCOM
a. VR: V in T [$_{TP}$ VP [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]					1	1
b. DS: do in T [$_{TP}$ VP [$_{TP}$ Subj do-T <VP>]]			1			
c. TS: T in T [$_{TP}$ VP [$_{TP}$ Subj T <VP>]]	1					

In the VR candidate V raises from within the VP into the tense position, and the VP which contains the original V position moves away. This is a case of remnant movement (see Müller (1998), Hinterhölzl, (2006)). I have not identified any languages with V-raising and leftward movement of the remnant VP, which should exist if the proposal is correct. However, rightward movement of the remnant VP is posited in Bhatt and Dayal (2007). They propose an analysis of Hindi-Urdu in which the V raises rightward and the remnant VP scrambles further to the right. The example in (28) from Bhatt and Dayal (2007) is their (13c), which has the observed order ‘S V Aux IO DO’. The IO+DO is a VP from which the V has been raised. The trace of the VP is between the subject and the verb. (29), which contains no auxiliary verbal material, was provided by V. Dayal, who comments that it has a particularly natural discourse status.

(28) [[Ram-ne t$_j$ dii$_i$ thii] [$_{VP1}$ Sita-ko [$_{VP2}$ kitaab t$_i$]]$_j$]

Ram-ERG give.PFV.F be.PST.FSG Sita-DAT book.F
'Ram had given a book to Sita.'

(29) [[kis-ne t_j dii_i] [$_{VP1}$ Sita-ko [$_{VP2}$ kitaab t_i]]$_j$]
Who-ERG give.PFV.F Sita-DAT book.F
'Who gave a book to Sita?'

Abstracting away from the direction of displacement, as discussed above, Hindi-Urdu is an instance of the VR candidate in (27). It violates *LINF, as all VR structures do. It also violates CCOM if the trace of V, which is inside the displaced VP, is not c-commanded by the raised V. I assume that this is the situation in all V-raising candidates. No other constraint is violated.

(30) VR-AdjVP winner

Input: {V int Subj T}	V+T	OBHD	FULLINT	*LINA′	*LINF	CCOM
a. W [$_{TP}$ VP [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]						
b. DS [$_{TP}$ VP [$_{TP}$ Subj do -T <VP>]]			W		L	L
c. TS [$_{TP}$ VP [$_{TP}$ Subj T <VP>]]	W				L	L

The V-to-T optimum, we see, is generated by any ranking consistent with the rankings in (31).

(31) V+T, FULLINT >> CCOM, *LINF

The DS candidate wins in English, where the VP adjoins to the left of TP. We can see this in the italicized subordinate clause in (32), where the complementizer precedes the preposed VP and the subject follows it. Here *do*+T occurs in the head of TP, following the subject, just as in (25) above.

(32) He was supposed to read the book, and he says *that read the book he did.*

The crucial rankings are revealed by the relevant comparison:

(33) DS-AdjVP winner

Input: {V int Subj T}	V+T	OBHD	FULLINT	*LINA′	*LINF	CCOM
b. **W [$_{TP}$ VP [$_{TP}$ Subj do-T <VP>]]**						
a. VR [$_{TP}$ VP [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]			L		W	W
c. TS [$_{TP}$ VP [$_{TP}$ Subj T <VP>]]	W		L			

(34) V+T >> FULLINT
CCOM or *LINF >> FULLINT

A TS winner is the third possibility, which I exemplify with *will* in English in Section 5. The comparative tableau shows that the ranking required for T-stranding is (36).

(35) TS-AdjVP winner

Input: {V int Subj T}	*LinF	Ccom	FullInt	V+T	ObHd	*LinA′
c. **[$_{TP}$ VP [$_{TP}$ Subj T <VP>]]**						
a. [$_{TP}$ VP [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]	W	W		L		
b. [$_{TP}$ VP [$_{TP}$ Subj do-T <VP>]]			W	L		

(36) Ccom or *LinF >> V+T
FullInt >> V+T

4.2 Specifier VPs

Moving now to the sub-c-set with specifier VPs, i.e. those chosen by the ranking VPSpec >> VPAdj, the relevant candidate set is twice as large. The reason is the presence of an additional head in the structure. This is the head of the projection containing the preposed VP as its specifier. Since the exact identity of the projection varies from case to case, I label it XP here. As for the Wh-Q c-set it is necessary to consider the location of the stranded T, the raised V or the *do*, which might be in T or in X, so there are two candidates for each of the three strategies. All are analyzed here. As the violation table makes clear, just as in the WhQ c-set, the constraint Unique crucially prevents harmonic bounding of the DS and TS candidates where *do* or T is in the T position by those in which it is in X, satisfying ObHd. (There is no such harmonic bounding in the VR pair, since each violates a constraint which the other satisfies.)

(37) Violation Table: Specifier VPs

Input: {V int Subj T}	V+T	ObHd	FullInt	*LinA′	*LinF	Ccom	Unique
a. VR: V in X [$_{XP}$ VP V-T-X [$_{TP}$ Subj <V-T> <[$_{VP}$ <V> int]>]]				1	1	1	2
b. VR: V in T [$_{XP}$ VP __ [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]		1			1	1	1
c. DS: do in X [$_{XP}$ VP do-T-X [$_{TP}$ Subj <do-T> <VP>]]			1				1
d. DS: do in T [$_{XP}$ VP __ [$_{TP}$ Subj do-T <VP>]]		1	1				
e. TS: T in X [$_{XP}$ VP T-X [$_{TP}$ Subj <T> <VP>]]	1						1
f. TS: T in T [$_{XP}$ VP __ [$_{TP}$ Subj T <VP>]]	1	1					

Since the logic of the situation is, I hope, clear by now, I will not show the CTs for the VR candidates in (37), since I do not have examples to put forward. The tableaux are given in Grimshaw (in prep. a.), and the candidates are included in the typology analyzed in Section 6. It has been suggested these candidates are impossible, because a phrase from which a head has been removed cannot undergo remnant movement. See Hinterhölzl (2006) for discussion and references. Of course within the theoretical assumptions of OT the fact that a structure is ungrammatical (as VR is in German VP displacements) does not license the conclusion that it is not a possible optimum. (It may be of interest that only one of the 28 languages generated by the theory presented here and listed in (54) has the lexical V in C with VP displacement.) In the present analysis, this candidate is ungrammatical in German because the rankings of the language eliminate it.

German, a devotee of VR up to this point, switches to the DS structure with VP displacement. It thus exemplifies, as Monnese did earlier, non-uniform distribution of DS, as Bader and Schmid (2006), henceforth 'B&S', show. Dutch also has *do*-support with VP displacement (Reuland, p.c.) In a main clause, the displaced VP occupies the position usually characterized as 'specifier of CP', and *do* raises from T across the subject into C, creating a verb second configuration. This is in essence the analysis of B&S. 'XP' in (37) is CP in the German case. B&S cite (38), their (14a):

(38)	Tanzen	tut	Katja	immer noch	häufig.
	Dance	*does*	Katja	still	often

Elías-Ulloa (2001) demonstrates DS in Capanahua, an SOV Amazonian language belonging to the Pano linguistic family, spoken in Eastern Peru. He proposes that a displaced VP occupies the specifier of an Evidential Phrase. (This matches the related language Shipibo, which has an overt evidential morpheme as the head of the Evidential Phrase.) Information-prominent constituents can front to first position, before the second position clitic. The examples in (39)–(40) are examples (72), (74) and (76) in his paper:

(39)	ʔɨn-taʔ	ʔɨn	yoʔa	βana-ni-ʔ-ki
	I-Cl2	I	yucca	plant-past_tense-1/2p-evidential
	'I planted yucca (long ago)'			

When the subject is null, the object is initial, as in (40). When both the subject and the object are null, the VP fronts and *h*a 'do' appears in X, as in (41). (Adverbial and aspectual suffixes move with the V, only tense is left behind; Elías-Ulloa (p.c.).)

(40)	yoʔa-taʔ	βana-ni-ʃ-ki
	yucca-Cl2	plant-past_tense-3p-evidential
	'(He) planted yucca'	

(41)	βana-taʔ	ha-ni-ʃ-ki

plant-Cl2   ha_support-past_tense-3p-evidential
'(He) planted (it)'

The analysis in (43) is modified from Elías-Ulloa (2001) to facilitate comparison between Capanahua and the other languages under discussion. For example, some constraint names are different. However, the structure of the analysis is retained. As the comparative tableau shows, FullInt and Unique must be dominated by at least one of CCom, *LinA′ and *LinF to choose the DS winner over the VR candidate a, with V in X. It must be dominated by ObHd or CCom or *LinF to eliminate the VR candidate b, with V in T. Unique must be dominated by ObHd, to eliminate candidate d. Finally, FullInt and Unique must be dominated by either V+T or ObHd, to eliminate candidate e, a TS structure. The subordination of FullInt to several other constraints is the key characteristic.

(42) *LinA′ or *LinF or CCom >> FullInt, Unique
ObHd or *LinF or CCom >> FullInt
ObHd >> Unique
V+T >> FullInt
V+T or ObHd >> FullInt, Unique

(43) DS-SpecVP winner

Input: {V int Subj T}	V+T	ObHd	FullInt	*LinA′	*LinF	CCom	Unique
c. W [$_{\text{XP}}$ VP do-T-X [$_{\text{TP}}$ Subj <do-T> <VP>]]							
a. [$_{\text{XP}}$ VP V-T-X [$_{\text{TP}}$ Subj <V-T> <[$_{\text{VP}}$<V> int]>]]			L	W	W	W	L
b. [$_{\text{XP}}$ VP __ [$_{\text{TP}}$ Subj V-T <[$_{\text{VP}}$ <V> int]>]]		W	L		W	W	
d. [$_{\text{XP}}$ VP __ [$_{\text{TP}}$ Subj do-T <VP>]]		W					L
e. [$_{\text{XP}}$ VP T-X [$_{\text{TP}}$ Subj <T> <VP>]]	W		L				
f. [$_{\text{XP}}$ VP __ [$_{\text{TP}}$ Subj T <VP>]]	W	W	L				L

The comparative tableau for a TS winner is given in (44). (Since English is the only T-stranding language considered here, and it chooses to adjoin displaced VPs, I do not provide an example of a TS optimum with a VP preposed to specifier.) The comparative tableau yields the rankings in (45), which are annotated with the comparison that motivates them. The noteworthy property of the ranking is the subordination of V+T.

(44) TS-SpecVP Winner

Input: {V int Subj T}	V+T	ObHd	FullInt	*LinA′	*LinF	CCom	Unique
e. W **[$_{XP}$ VP T-X [$_{TP}$ Subj <T> <VP>]]**							
a. [$_{XP}$ VP V-T-X [$_{TP}$ Subj <V-T> <[$_{VP}$ <V> int]>]]	L			W	W	W	W
b. [$_{XP}$ VP __ [$_{TP}$ Subj V-T <[$_{VP}$ <V> int]>]]	L	W			W	W	
c. [$_{XP}$ VP do-T-X [$_{TP}$ Subj <do-T> <VP>]]	L		W				
d. [$_{XP}$ VP __ [$_{TP}$ Subj *do*-T <VP>]]	L	W	W				L
f. [$_{XP}$ VP __ [$_{TP}$ Subj T <VP>]]		W					L

(45) *LinA′, *LinF, CCom or Unique >> V+T (e ~ a)
*LinF, CCom or ObHd >> V+T (e ~ b)
FullInt >> V+T (e ~ c)
ObHd or FullInt >> V+T, Unique (e ~ d)
ObHd >> Unique (e ~ f)

The optima for VP displacement illustrate again the important point that the choice of whether *do*-support is allowed is not made by the language as a whole. We have already seen this for interrogatives and negation, where Monnese chooses *do*-support in one and not the other, while both English and German are uniform. Now we see it again in the case of German, which chooses *do*-support for VP-preposing but not elsewhere, a point emphasized in Bader and Schmid (2006). No inherent asymmetry is detectable between *do*-support and V-raising or T-stranding.

5 T-stranding: non-uniform *do*-support in English

As I noted earlier, English also shows non-uniform distribution for *do*-support, if *will*, and perhaps the other “modals” which also are non-verb-like in having no person inflection, are analyzed as cases of stranded T. Languages can choose between realizing certain grammatical information as grammatical affixes, like the English, present and past tenses, or as free morphemes, like *will*. The present hypothesis characterizes this as a decision on whether to strand the morpheme or not, governed by constraint ranking, not via a lexical stipulation. In such a proposal, the choice of a free morpheme for the future tense is governed by the entire grammar: the ranking of the critical constraint must be consistent with the grammar as a whole.

When the input is present or past tense, *do*-support is the optimum in English with negation, in *wh* interrogatives and with VP displacement. However when the input is specified as future, *will* appears in the positions that are occupied by *do* in DS optima. When *will* is present, *do*-support is ungrammatical. Compare the examples in (46) with the *do*-support examples cited earlier.

There is no word order which will make the starred versions of these sentences grammatical: *do* and *will* simply cannot co-occur.

(46) (a) I will not know it. *I do will not know it.
(b) What will he eat? *What does will he eat??
(c) Read the book he will. *Read the book he does will

The analysis I propose is that the constraint V+T universally breaks down into (at least) two constraints, V+F and V+P, where 'F' stands for future and 'P' for past (and, conveniently, present). Because these constraints can be ranked separately, a language choose stranding for one and not the other, as English does. Everything I have said in previous sections about the choice of optima holds for inputs containing P. V+F is vacuously satisfied in all the cases presented, so adding it to the system has no effect on the prior analyses.

The violation tables given above for the Negation c-set, the WhQ c-set and the VP displacement c-set are valid for the future tense versions of the candidate sets, if slightly modified: T in the input is replaced with F and the constraint that is violated by the TS candidate in the tables is now V+F. V+P is vacuously satisfied.

The comparative tableaux for the TS candidates discussed above are also valid for the future versions of the candidate sets. Thus if 'T' in (12) is replaced with 'F' we derive the tableau in (47).

(47) FS-Neg winner

Input: {V Subj F Neg}	V+F	OBHD	FULLINT	*LINA′	*LINF	UNIQUE
c. W [$_{TP}$ Subj -F Neg [$_{VP}$ V ...]]						
a. VR [$_{TP}$ Subj V-F Neg [$_{VP}$ <V> ...]]	L				W	W
b. DS [$_{TP}$ Subj do-F Neg [$_{VP}$ V ...]]	L		W			

The rankings in (48) are required to select the English optimum *I will not know it* from the future tense Negation c-set.

(48) FULLINT >> V+F *LINF or UNIQUE >> V+F

The same transformation is valid for the comparative tableaux for the WhQ c-set and the VP displacement c-sets. The FS-WhQ winner (identical to (20) in all but the tense) is in (49) and establishes the rankings in (50), which are those in (21) with F replacing T. This optimum corresponds to the grammatical question in (46).

(49) FS-WhQ winner

Input: {V Subj F Wh}	V+F	ObHd	FullInt	*LinA′	*LinF	Unique
c. W **[$_{CP}$ Wh -F [$_{TP}$ Subj <-F> VP]]**						
a. VR [$_{CP}$ Wh V-F [$_{TP}$ Subj <V-F> [$_{VP}$ <V> …]]]	L			W	W	W
b. DS [$_{CP}$ Wh do-F [$_{TP}$ Subj <do-F> VP]]	L		W			
d. ØC [$_{CP}$ Wh __ [$_{TP}$ Subj V-F [$_{VP}$ <V> …]]]	L	W			W	

(50) FS-WhQ FullInt >> V+F
ObHd or *LinF >> V+F
*LinA or *LinF or Unique >> V+F

The FS AdjVP winner is identical to (35) apart from the tense-related changes, and the rankings are identical to those in (36) with the same proviso. The winning candidate corresponds to the grammatical VP displacement in (46).

(51) FS-AdjVP winner

Input: {V int Subj T}	*LinF	CCom	FullInt	V+F	ObHd	*LinA′
c. **[$_{TP}$ VP [$_{TP}$ Subj F <VP>]]**						
a. [$_{TP}$ VP [$_{TP}$ Subj V-F <[$_{VP}$ <V> int]>]]	W	W		L		
b. [$_{TP}$ VP [$_{TP}$ Subj do-F <VP>]]			W	L		

(52) CCom or *LinF >> V+F
FullInt >> V+F

In this proposal, the case of *will* is simply the missing exemplar of T stranding. What remains to be shown is that the ranking of V+F that we now know is required for English is consistent with the ranking of V+P that has already been established under the guise of the ranking of V+T. That is, is there a total ranking of the (now 10) constraints examined here that chooses DS in the past (and present) Negative, WhQ and VP displacement candidate sets but chooses to strand *will* in the corresponding future candidate sets? The answer is in the affirmative. Any ranking which is consistent with the following stratal relationships will do exactly this.

(53) Top stratum: *LinA′, V+P, CCom, ObHd, *LinF, VPAdj
Second stratum: VPSpec, Unique, FullInt
Third stratum: V+F

This ranking information was arrived at by performing RCD using OTWorkplace (Prince and Tesar (2010)) over all six candidate sets (the three considered earlier plus their future counterparts) and including two additional candidates in the *wh* question c-set, as described below. The result can be found in Grimshaw (in prep. a.). A crucial property of (53) is that V+P and V+F are separated in the ranking, allowing for different optima to be chosen when the violation of V+P is at stake and when V+F is at stake.

Since the separation of V+T into two constraints makes the typology rather unwieldy, and I wish to focus on properties of its structure which are not related to this aspect of the analysis, in the discussion of the typology in the following section I will examine the typology for uniform tense systems, generated by a constraint set containing constraint V+T. However, the core structural properties of this simpler typology are preserved in the larger one (see Grimshaw in prep. a.), and I will note this as relevant in the next section.

6 The typology

We can now combine the three c-sets under scrutiny. The candidates are those analyzed in (9) for negation plus those analyzed in (17) for *wh* interrogatives. Two additional WhQ candidates are included, one with *do*-support in T and an empty C, and one with a stranded T in T and an empty C. (These correspond to candidates d and f in (37), which is the violation table for VPs displaced to specifier position.) There are thus three candidates for negation, six for *wh* questions, and nine for VP displacement. Three of the nine have the VP in an adjoined position and six have the VP in a specifier position. The violation table constructed from merging all the VTs in this paper and adding the extra candidates just mentioned can be found in Grimshaw (in prep. a.).[11]

Since the system under investigation posits nine constraints, there are 9! possible rankings of them. There are 162 logically possible combinations of candidates, involving one from each of the three candidate sets. The factorial typology, calculated using OTWorkplace (Prince and Tesar (2010)) is in (54). It contains 28 distinct languages, i.e. combinations of optima. To make the patterns as easily visible as possible, I have classified the optima as cases of VR, DS or TS and used these labels, rather than representations of the candidates themselves in the table. (The candidates are directly represented in the factorial typology in Grimshaw (in prep. a.)

(54) The Factorial Typology

Lge	Negation	*Wh* Question	VP	Displacement
1	**VR**	**VR** to C	A-VP	**VR**
2	**VR**	**VR** to C	A-VP	***DS***
3	**VR**	**VR** to C	A-VP	TS
4	**VR**	**VR** to C	S-VP	**VR** to C
5	**VR**	**VR** to C	S-VP	***DS*** in C
6	**VR**	**VR** to C	S-VP	TS in C
7	**VR**	**VR** to T	A-VP	**VR**
8	**VR**	**VR** to T	A-VP	***DS***

9	**VR**	**VR** to T	A-VP	TS
10	**VR**	**VR** to T	S-VP	**VR** to T
11	**VR**	**VR** to T	S-VP	***DS*** in C
12	**VR**	**VR** to T	S-VP	***DS*** in T
13	**VR**	**VR** to T	S-VP	TS in C
14	**VR**	**VR** to T	S-VP	TS in T
15	**VR**	***DS*** in C	A-VP	**VR**
16	**VR**	***DS*** in C	A-VP	***DS***
17	**VR**	***DS*** in C	S-VP	***DS*** in C
18	**VR**	TS in C	A-VP	**VR**
19	**VR**	TS in C	A-VP	TS
20	**VR**	TS in C	S-VP	TS in C
21	***DS***	***DS*** in C	A-VP	***DS***
22	***DS***	***DS*** in C	S-VP	***DS*** in C
23	***DS***	***DS*** in T	A-VP	***DS***
24	***DS***	***DS*** in T	S-VP	***DS*** in T
25	TS	TS in C	A-VP	TS
26	TS	TS in C	S-VP	TS in C
27	TS	TS in T	A-VP	TS
28	TS	TS in T	S-VP	TS in T

Half of the languages are uniform with respect to strategy. Six have VR everywhere: 1, 4, 7, 10, 15, 18. Four have DS everywhere (21–24). Four have TS everywhere (25–28). Each strategy is used by several languages. The languages vary in where V, *do* and T occur, as well as in how they treat VP displacement. (Note that English is #21: DS with negation, DS in C in the WhQ winner and an adjoined VP with DS in T in the VP displacement competition.)

Half of the languages are non-uniform, like German and Monnese. It is impossible to say for a language as a whole that it has or does not have DS, or TS or VR. Twelve of the languages have DS somewhere, and twelve have TS somewhere. Twenty have VR somewhere (all of these have it with negation; see below.) The distribution is predicted by rankings (the rankings for each language are reported in Grimshaw (in prep. a.)). There doesn't seem to be any way in which the notion of a last resort sheds light on the patterns.

Non-uniformity follows from the fact that compatible rankings can yield non-uniform results. I pointed out at the end of Section 3 that the rankings which choose VR with negation are compatible with rankings which choose DS in the WhQ c-set, hence Monnese is generated. Monnese could be language 15, 16 or 17, given what we currently know about it.[12]

The same logic holds for German, which is language #5. The rankings established for German are for VP-preposing, negation and interrogatives are all compatible.

(55) German ranking

```
CCOM          V+T
        \       /
        FULLINT        OBHD
        /       \       /
*LINF         *LINA′
                  |
              UNIQUE
```

Any rankings compatible with the stratal relationships in (55) preserve all of these dominance relations.

So both Monnese and German are generated by the theory. (They are also predicted to be possible in the larger typology generated when T+P and T+F are separated.) In contrast, neither Anti-Monnese nor Anti-German is among the generated languages. Anti-Monnese would have a VR winner in the WhQ c-set and a DS winner in the Negation c-set. Anti-German would have DS with negation and in the WhQ c-set and VR with VP displacement. There are no such languages in (54).[13] In fact, all languages with DS winners in the Negation c-set have DS in all c-sets. (The same is true in the larger typology generated by separating V+P and V+T. See Grimshaw in prep. a.)

Why is this so? The answer is that in order to have a DS winner in the Negation c-set, the language must have a ranking in which V+T and either *LINF or UNIQUE dominate FULLINT. Such a ranking can never choose VR in the WhQ c-set, because this choice requires that FULLINT dominate *LINF, *LINA′ and UNIQUE, as (18) showed. Moreover, a language with DS with negation cannot choose VR in the VP displacement c-set. The only language with a VP displaced as a specifier and VR into C is #4. The ranking for 4 has FULLINT dominating *LINF *LINA′, CCOM and UNIQUE. (These calculations are reported in Grimshaw (in prep. a.))

It is important to be clear exactly what the prediction is. It is not predicted that any language which has DS optima with negation must have DS in *wh* questions. Korean, for example, has DS with negation (see Section 1) but not in questions, but this is because it doesn't have the relevant kind of question, i.e. its other rankings are not the same as those for the languages looked at here. The prediction is that any language where the same constraints make the decisions as in the cases studied here must make the same choices. This point is made effectively by Bader and Schmid (2006) in response to a carelessly worded claim made in Grimshaw (1997), that no language with the rankings of German could have DS optima. A more accurate wording would say that no language with the critical rankings of German could have DS optima in the Negation or WhQ c-sets, because the choice between DS and some alternative is not made by the lexicon (i.e. arbitrarily) but by the grammar of the language.

There are several more general points to be made about the typology. Most obviously, the number of languages generated is not 9!. The number of possible combinations of optima is not 162. The reason is that the typology is highly structured. There are many entailments between optima in one candidate set and optima in others. See Anttila and Andrus (2006), Prince (2006).

These patterns hold:

- No language uses all three structures.
- The languages (#21–#28) which choose DS or TS optima in the negation c-set preserve the choice in the other two c-sets.
- It follows from this, given that there are only three alternative strategies, that any language with VR in the WhQ c-set must have VR in the negation c-set, and any language with VR in the VP displacement c-set must have VR in the Negation c-set.
- And that all non-uniform languages have VR in the Negation c-set.
- VR in the Negation c-set is compatible with all three choices in the WhQ and VP displacement c-sets.

In the larger typology generated by splitting the V+T constraint into two, the same structural generalizations are true. There are 162 languages. 48 use all three structures in optima, but in every case this is due to tense split. That is, none of them uses one optimum in the Negation c-sets, a different one in the WhQ c-sets and yet another in the VP displacement c-sets, for a fixed choice of tense. The second bulleted point is preserved also, relative to tense. Thus a language which chooses DS or TS optima in the Negation c-set for a particular tense must choose the same in the other c-sets which involve the same tense. The third point also holds, relative to tense. Thus, for example, if a language has VR in the WhQ c-set with the past tense, it must also have VR in the Negation c-set for the past tense. Finally, we continue to observe the general distribution in which VR is most frequent in the Negation c-sets (it occurs in 118 of the languages in both of the Negation c-sets) and compatible with all three choices in the others. A compressed version of the typology appears in (56). It records which of the three options is optimal for each candidate sets, but not the location of T, V or *do.*

(56) The Split Tense Typology

	Neg Past/pres	Neg Fut	WhQ Past/pres	WhQ Fut	VP Disp Past/pres	VP Disp Fut
Lg#119,120,121,122	DS	DS	DS	DS	DS	DS
Lg#123, 124, 125,126	DS	TS	DS	TS	DS	TS
Lg#127,130,131,134, 137,138, 149,152	TS	VR	TS	VR	TS	VR

Lg#128,132,135,139, 150,153	TS	VR	TS	VR	TS	DS
Lg#129,133,136,140, 151,154	TS	VR	TS	VR	TS	TS
Lg#141,143	TS	VR	TS	DS	TS	VR
Lg#142,144	TS	VR	TS	DS	TS	DS
Lg#145,147	TS	VR	TS	TS	TS	VR
Lg#146,148	TS	VR	TS	TS	TS	TS
Lg#155,156, 157,158	TS	DS	TS	DS	TS	DS
Lg#159, 160, 161, 162	TS	TS	TS	TS	TS	TS

This picture bears in an interesting way on the question of whether DS, or any of the alternative strategies, can be identified as being highly marked or a 'last resort'. It is clear that what might seem like a particularly marked grammatical device in one c-set may seem much less so in another. For example, VR with negation is highly unmarked in the sense that it is optimal in 20 of the 28 languages. However, it is optimal in the WhQ c-set in only 14 languages and in the VP displacement c-set in only 6. So if the number of languages which use a particular device is a measure of its last resort status, there doesn't seem to be a clear answer to the question of whether a particular strategy is a last resort or not.

Let us consider why DS and TS are found so much more prominently in the WhQ and VP displacement c-sets than in the negation c-set. DS is found in 4 languages in the negation c-set, in 7 in the WhQ c-set and in 11 in the VP displacement c-set. TS shows exactly the same numbers. The explanation is as follows. Each of TS and DS violates exactly one constraint within this system, namely V+T or FULLINT. These constraints are relevant in all three c-sets. Two additional constraints are relevant for the WhQ c-set, namely OBHD and *LINA′. While OBHD is insensitive to the difference between a head occupied by a V, a T or a *do*, *LINA′ is not. It prefers T or *do* over V in the C position. Hence its ranking can force the choice of T or *do* in those candidates in which C is filled. Finally, for the VP displacement c-set, CCOM is relevant because it is violated in some candidates. Again it prefers T or *do* over V since VR leads to CCOM violations. Thus as we move from left to right through the candidate sets in (54) we find that VR is increasingly disfavored and hence the two alternatives, TS and DS, are increasingly favored. (Only one language (#4) allows VR with a VP in specifier position, and it has VR optima in all c-sets.) What emerges is that the languages which have VR optima for the VP displacement c-set are a proper subset of those which have VR optima for the WhQ c-set, which in turn are a proper subset of those that have VR for the Negation c-set.

Does the notion of last resort play any role in our understanding of the system? We can identify a strategy or structure that, as the optimum for a given c-set, is compatible with fewer optima in other c-sets. To put it another way, it is the optimum for a smaller set of combinations of optima, i.e. languages. E.g. DS and TS in the Negation c-set are compatible with fewer optima for the WhQ c-set and the VP displacement c-set than VR. We could agree to call DS or TS a last resort for the Negation c-set. However, nothing is gained by doing so. The work is being done by the theory of constraint interaction and comparison among members of the candidate sets, and the label 'last resort' adds nothing.

Acknowledgements:
Veneeta Dayal and José (Beto) Elías-Ulloa provided me with key examples, and much help in understanding how they work. Sabine Mohr, Britta Sauereisen and Susanne Lohrmann explained *do*-support in Swabian to me. Several people helped with the OTWorkplace calculations or constructed paradigms I needed, in particular Fabian Heck, Brian Smith and Alan Prince. Sara O'Neill's outstanding work turned a manuscript into a paper. Many others, including audiences at UMass Amherst, the University of Stuttgart, RuLing II, RORG and St@r, raised questions which sharpened the argument and analysis presented in the paper. Many helpful people discussed *do*-support or the analysis with me: Birgit Alber, Artemis Alexiadou, Daniel Altshuler, Will Bennett, James Bruno, José Camacho, Christa Gordon, Patrick Houghton, Paula Houghton, Bart Hollebrandse, Carlo Linares, Sara O'Neill, Alan Prince, Ljiljana Progovac, and Tanja Schmid. Special thanks to Alan Prince and Vieri Samek-Lodovici, whose input and insights had a major impact on this research. Finally, thanks are due to Hans Broekhuis and Ralf Vogel for making things happen.

grimshaw@ruccs.rutgers.edu

- References -

Ackema, P. (2001) On the relation between V-to-I and the structure of the inflectional paradigm. *The Linguistic Review* 18: 233--263.
Anttila, A. and Andrus, C. (2006) T-orders. ROA-873. http://roa.rutgers.edu
Bader, M. and Schmid, T. (2006) An OT-analysis of *do*-support in Modern German. Rutgers Optimality Archive. ROA 837-0606. Retrieved on March 31 2010 from http://roa.rutgers.edu/searchlist.php3?num=7&detail=&pointer=0&search=bader&sf=Author&ids=1161
Bhatt, R. and Dayal, V. (2007) Rightward scrambling as rightward remnant movement. *Linguistic Inquiry* 38: 287--301.
Benincà, P. and Poletto, C. (2004) A case of *do*-support in Romance. *Natural Language and Linguistic Theory* 22: 51--94.

Bobaljik, J. D. (1995) Morphosyntax: The syntax of verbal inflection. Doctoral dissertation. Cambridge, MA: MIT Working Papers in Linguistics.
Bosković, Z and Lasnik, H., eds. (2006) *Minimalist Syntax: The Essential Readings*. Oxford: Wiley-Blackwell.
Bosković, Z and Takahashi, D. (1998) Scrambling and last resort. *Linguistic Inquiry* 29: 347--366.
Chomsky, N. (1955) The logical structure of linguistic theory. Manuscript, Harvard University and MIT. Published in part 1975, New York: Plenum.
Chomsky, N. (1957) *Syntactic Structures*. Berlin: Mouton.
Chomsky, N. (1991) Some notes on economy of derivation and representation. In R. Freidin (ed.) *Principles and Parameters in Comparative Grammar* 417--54. Cambridge, MA: MIT Press.
Chung, S. (1990) VPs and verb movement in Chamorro. *Natural Language and Linguistic Theory* 8: 559--619.
Collins, C. (2001) Economy conditions in syntax. In M. Baltin and C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory* 45--61. Oxford: Blackwell.
Corver, N. (1997) Much-support as a last resort. *Linguistic Inquiry* 28: 119--164.
Culicover, P. W. (2008) The rise and fall of constructions and the history of English *do*-support. *Journal of Germanic Linguistics* 20: 1--52.
Dowty, D. (1979). *Word Meaning and Montague Grammar*. Dordrecht, Boston: D. Reidel.
Duffield, N. (1999). Adjectival Modifiers and the Specifier-Adjunct Distinction. In D. Adger, S. Pintzuk, B. Plunkett and G. Tsoulas (eds.) *Specifiers: Minimalist Approaches* 126-145. Oxford, New York : Oxford University Press.
Elías-Ulloa, J. A. (2001) Subject doubling in Capanahua. Ms., Rutgers University.
Erb, M. C. (2001) Finite auxiliaries in German. Doctoral dissertation. Katholieke Universiteit Brabant.
Grimshaw, J. (1997) Projection heads and optimality, *Linguistic Inquiry* 28: 373--422.
Grimshaw, J. (2001) Economy of structure in OT. Rutgers Optimality Archive. ROA 434-0601.
Grimshaw, J. (2002) Economy of structure in OT, in Papers in Optimality Theory II. University of Massachusetts Occasional Papers 26. 81--120.
Grimshaw, J. (2006a) Chains as unfaithful optima. ROA 844.-0606. Note: this is a chapter of ROA-844, *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince*.
Grimshaw, J. (2006b) Last resorts and grammaticality. In H. Broekhuis and R. Vogel (eds.) *Optimality Theory and Minimalism: A Possible Convergence?* LiP 25: http://www.ling.uni-potsdam.de/lip/ 33--41. Also available as ROA 892.2-0107.
Grimshaw, J. (2006c) Location specific constraints in matrix and subordinate clauses. ROA 857.
Grimshaw, J. (2008) Syntactic constraints. In Piet van Sterkenburg (ed.) *Unity and Diversity of Languages* 43--57. Amsterdam/Philadelphia: John Benjamins.
Grimshaw, J. (in prep. a.) Last Resorts: A Typology of *Do*-Support. Supplementary material. To be posted at http://rulinguistics101.org/page/grimshaw.html.
Grimshaw, J. (in prep. b.) The grammar of English complementizer omission.
Grimshaw, J. and Mester, A. (1988) Light verbs and theta-marking. *Linguistic Inquiry* 19: 205--232.

Hagstrom, P. (1996) *Do*-support in Korean: Evidence for an interpretive morphology. In H.-D. Ahn et al. (eds.) *Morphosyntax in Generative Grammar: Proceedings of the 1996 Seoul International Conference on Generative Grammar* 169--180. Seoul: Hankuk.

Halle, M. and Marantz, A. (1993) Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds.) *The View From Building 20*, Cambridge, MA: MIT Press.

Harley, H. and Noyer, R. (1998) Mixed nominalizations, short verb movement and object shift in English. In P. N. Tamanji and K. Kusumoto (eds.) *Proceedings of NELS 28*: 143--157.

Hinterhölzl, R. (2006) *Scrambling, Remnant Movement and Restructuring in West Germanic*. Oxford: Oxford University Press.

Hollebrandse, B. and Roeper, T. (1996) The concept of do-insertion and the theory of infl in acquisition. In C. Koster and F. Wijnen (eds.) *Proceedings of the Groningen Assembly on Language Acquisition* 261--270. Netherlands: Center Language and Cognition Groningen.

Kuroda, S-Y. (1965) Generative grammatical studies in the Japanese language. Doctoral disertation, MIT.

Larson, R. (1988) On the double object construction. *Linguistic Inquiry* 19: 335--391.

Lasnik, H., Uriagareka, J. and Boeckx, C. (2005) *A Course in Minimalist Syntax*. Oxford: Blackwell.

Legendre, G. (2001) Masked Second-Position Effects and the Linearization of Functional Features. In G. Legendre, S. Vikner, and J. Grimshaw, (eds.), OT Syntax. MIT Press. 241-277.

Legendre, G., Grimshaw, J. and Vikner, V. (eds.) (2001) *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press.

McCarthy, J. and Prince, A. (1997/1999) Faithfulness and identity in prosodic morphology. In van der Hulst et al. (eds.) *The Morphology-Prosody Interface* 218--309. Cambridge: Cambridge University Press. ROA-216.

Meinunger, A. (2001) Restrictions on verb raising. *Linguistic Inquiry*. 732--740.

Miller, P. H. (1997) Auxiliary verbs in old and middle French: A diachronic study of substitute faire and a comparison with the modern English auxiliaries. In A. van Kemenade and N. Vincent (eds.) *Parameters of Morphosyntactic Change* 119--133. Cambridge: Cambridge University Press.

Miyagawa, S. (2001) The EPP, scrambling, and wh-in-situ. In M. Kenstowicz (ed.) *Ken Hale: A Life in Language* 293--338. Cambridge, MA: MIT Press.

Müller, G. (1998) *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht, Boston: Kluwer Academic Publishers.

Newton, G. (2009) Accounting for *do*-support post-syntactically: Evidence from Old Irish. *University of Pennsylvania Working Papers in Linguistics* 15: Iss.1, Article19. Retrieved on 31 March 2010 from http://repository.upenn.edu/pwpl/vol15/iss1/19

Pesetsky, D. and Torrego, E. (2001) T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.) *Ken Hale: A Life in Language* 355--426. Cambridge, MA: MIT Press.

Pesetsky, D. and Torrego, E. (2007) The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiian and W. Wilkins, (eds.) *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*. Amsterdam: Benjamins.

Pollock, J.-Y. (1989) Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20: 365--424.

Prince, A. (1997) Stringency and anti-Paninian hierarchies. Handout from LSA Institute. Retrieved on 31 March 2010 from http://ling.rutgers.edu/people/faculty/prince.htm
Prince, A. (2002) Arguing optimality. In A. Coetzee, A. Carpenter and P. de Lacy (eds.) *Papers in Optimality Theory II*. GLSA, UMass. Amherst. ROA-536.
Prince, A. (2006) Implication & impossibility in grammatical systems: What it is & how to find it. ROA-880. http://roa.rutgers.edu.
Prince, A. and Smolensky, P. (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Wiley-Blackwell.
Prince, A. and Tesar, B. (2010) OTWorkplace. http://equinox.rutgers.edu/people/faculty/prince.html
Saito, M and Hoshi, H. (2000) The Japanese light verb construction and the Minimalist Program. In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.
Samek-Lodovici, V. (to appear) Optimality Theory and the Minimalist Program. In H. Broekhuis and R. Vogel (eds.) *Linguistic Derivations and Filtering*. Equinox.
Samek-Lodovici, V. and Prince, A. (2005) Fundamental properties of harmonic bounding. ROA 785-1205. http://roa.rutgers.edu
Schmid, T (2005) *Infinitival Syntax: Infinitivus Pro Participio as a repair strategy*. Amsterdam/Philadelphia: John Benjamins.
Schütze, C. (2001) Semantically empty lexical heads as last resorts. In N. Corver and H. van Riemsdijk (eds.) *Semi-Lexical Categories*. Berlin/New York: Mouton de Gruyter.
Schütze, C. (2004) Synchronic and diachronic microvariation in English do. *Lingua* 114: 495--516.
Soh, H. L. (2007) Ellipsis, last resort, and the dummy auxiliary shi 'be' in Mandarin Chinese. *Linguistic Inquiry*, 38: 178--188.
Vikner, S. (2001) V^0-to-I^0 and *do*-insertion in Optimality Theory. In Legendre et al. (eds.) 427--464.
Zepter, A. (2000) Specifiers and adjuncts. Ms. Rutgers University. ROA 413-0900.

---

[1] This paper is a major extension and development of Grimshaw (2006b).

[2] It is impossible to cite all works which rely on or address the notion of a last resort, which has been prominent since Chomsky (1955). Here are a few which represent or review a variety of different applications: Chomsky (1991), Corver (1997), Rizzi (1997), Bosković and Takahashi (1998), Harley and Noyer (1998), Collins (2001), Benincá and Poletto (2004), Schmid (2005) and several papers in Bosković and Lasnik eds. (2006).

[3] See Grimshaw (in prep. a.) for the analysis of some of the harmonically bounded candidates of relevance here.

[4] This constraint will be split into two: one for past and present, and one for future, in Section 5.

[5] Since many analyses of English, largely stemming from Larson (1988), posit movement of a lexical V to a head position below T, I assume that *LINF is not violated by such movements, since the target positions are too low. This needs further technical development.

[6] See Grimshaw (1997), Ackema (2001) and Vikner (2001) for analyses of how V and T combine when negation is not present.

[7] It is widely reported that both German and Dutch have dialects in which *do*-support is optional in positive declarative clauses. See Ackema (2001), Vikner (2001) and Erb (2001). (See also Schütze (2004) on English *do* and Soh (2007) on Chinese *shi*.) Bader and Schmid (2007) analyze the optionality of *do* as due to tied constraints. However, I believe that the apparent optionality has two sources. The first is the discourse status of a clause: as Erb (2001) notes, there seems to be a relationship between the status of a clause as an assertion and the presence of *do*. The second is speakers' control of multiple grammars. See Anttila (2007) for recent discussion. Grimshaw (in prep. b.) presents an analysis of the apparent optionality of *that* omission in complement sentences which argues that it is the result of two register-connected grammars. In this analysis omission is not optional in either grammar.

[8] The only place that Benincà & Poletto identify where *do*-support is broader than in English concerns the verbs *fa* 'do' and *nda* 'go', which optionally occur with *do*-support (p. 73) (i.e., optionally move to C in their analysis), whether they are used as auxiliaries or as main verbs, taking thematic complements. Benincà & Poletto analyze these verbs as 'semi-auxiliaries'. See p. 81–84.

[9] An alternative way to address the difference between VP-preposing and *wh*-movement would posit a specifier position in both cases but treat the heads differently. E.g., the head of the projection containing the *wh* phrase would be governed by an OBHD constraint while that for VP displacement would not. Or both would be associated with constraints but the rankings would differ. In a theory which disallows adjunction some solution along these lines must be posited. See Duffield (1999) for an argument in support of a theory admitting both specifiers and adjunction.

[10] Other constraints, in particular the alignment constraints discussed in Grimshaw (2001, 2002), govern the order of heads and phrases.

[11] While candidates with remnant movement of the VP are in fact in the c-sets for both negation and *wh* questions I do not consider them here.

[12] I assume that Monnese does not have VP displacement with *do*-support. B&P mention that it lacks VP ellipsis after auxiliaries. The typology generated by the constraints under discussion predicts that it could have either DS or VR with VP displacement if it has the structures at all.

[13] Provided that movement through T is a prerequisite for movement to C, the B&P analysis will not generate Anti-Monnese. (I am not sure about Anti-German.) Under this assumption, a language which does not allow V to raise to T could not allow it to raise to C. B&P do not discuss this question, but the theory of features which regulate movement might restrict the

options in this way. Anti-Monnese and Anti-German are clearly generable, under the construction based approach of Culicover (2008), in which ‘every construction in which there is *do*-support must explicitly mention Vaux. This includes not only SAI and negation, but ellipsis and related constructions, and VP-topicalization.’