On the relation between V-to-I and inflectional features

Peter Ackema

(University of Groningen)

1. Introduction[1]

Recent research in the realm of V-to-I movement (verb movement out of VP, across a VP-initial adverb, cf. Emonds 1978, Pollock 1989) has given a firm basis to the idea that this movement is related to the presence of "rich" verbal agreement inflection in the language in question (see for instance Platzack & Holmberg 1989, Roberts 1993, Rohrbacher 1994, Koeneman 1997, Vikner 1997a). Vikner (1997a:189) supplies the following data, showing that in English, Danish and Faroese the verb remains in situ whereas in Icelandic, Yiddish and French it undergoes V-to-I. When these data are compared with the paradigms of the regular inflection for finite verbs in these languages (given in (3), from Vikner 1997a:191), it becomes clear that there is some relation between V-to-I and rich inflection, at least within this group of languages.

(1)	a.	That John often eats tomatoes (surprises most people)	(English)
	b.	At Johan ofte spiser tomater (overrasker de fleste)	(Danish)
	c.	At Jón ofta etur tomatir (kemur óvart á tey flestu)	(Faroese)
	d.	*Að Jón oft borðar tómata (kemur flestum á óvart)	(Icelandic)
	e.	*Az Jonas oft est pomidorn (is a xidesh far alemen)	(Yiddish)
	f.	*Que Jean souvent mange des tomates (surprend tout le monde)	(French)

(2)	a.	*That John eats often tomatoes (surprises most people)	(English)
	b.	*At Johan spiser ofte tomater (overrasker de fleste)	(Danish)
	c.	(*)At Jón etur ofta tomatir (kemur óvart á tey flestu)	(Faroese)
	d.	Að Jón borðar oft tómata (kemur flestum á óvart)	(Icelandic)
	e.	Az Jonas est oft pomidorn (is a xidesh far alemen)	(Yiddish)
	f.	Que Jean mange souvent des tomates (surprend tout le monde)	(French)

(3) Paradigms of the present tense of 'hear'

Eng	Dan	Far	Ice	Yid	Fre
hear	høre	hoyra	heyra	hern	écouter
I hear	jeg hører	eg hoyri	ég heyri	ikh her	j'écoute
you hear	du hører	tú hoyrir	þú heyrir	du herst	tu écoutes
he hears	han hører	hann hoyrir	hann heyrir	er hert	il écoute
we hear	vi hører	vit hoyra	við heyrum	mir hern	nous écoutons
you hear	I hører	tit hoyra	þið heyrið	ir hert	vous écoutez
they hear	de hører	tey hoyra	þeir heyra	zey hern	ils écoutent

Nevertheless, some problems with the idea that rich inflection triggers V-to-I movement keep recurring as well. First, it appears to be difficult to give a decisive definition of "richness" of inflectional paradigms, in the sense that all languages with such a rich inflectional paradigm have V-to-I and all languages that do not have such a paradigm lack V-to-I. Various proposals have been made as to what the defining characteristic of a rich inflectional paradigm is (see the references quoted above). Although these definitions cover a lot of empirical ground, there remain some languages which have paradigms that should not be rich enough to trigger V-to-I but have it nevertheless (see for example Roberts 1993 on Middle Scots and Platzack & Holmberg 1989 or Vikner 1995 on Kronoby Swedish).[2] Moreover, there are two languages, namely the two dialect of Faroese as described in Jonas 1996, that have an identical verbal inflectional paradigm but differ with respect to whether they allow V-to-I or not (hence the bracketed asterisk in (2c)). Also, English poses a problem in this respect. Although this language is often taken as a model of a non-V-to-I language, it does have V-to-I for auxiliaries (cf. Pollock 1989, Roberts 1998):

(4) a. John has completely lost his mind
 b. ?*John completely has lost his mind

This is something of a puzzle, given that an auxiliary like *have* in (4) does not have a richer inflectional paradigm than main verbs.

The second question that keeps recurring concerns the trigger for V-to-I. Several accounts have been given that establish a relation between rich inflection and an independent I-position in syntax to which the verb must move, but difficulties remain with respect to the explanation of why such a relation exists. Rohrbacher (1994), for example, argues that the distinction between rich and poor inflection is that the former is referential but the latter is not. He further argues that only referential inflection is listed in the lexicon and inserted separately from the verb in a distinct I-node. Nonreferential inflection is generated directly on the verb. Neeleman (1997) notes that this approach necessitates a model of grammar in which inflectional affixes can be added in the lexicon (as in strong lexicalist models, cf. Lapointe 1980) but can also occur on the verb as a result of postsyntactic spell-out rules (as in "split morphology" models, cf. Anderson 1992). This is an unattractive state of affairs, since inflection is handled by two different mechanisms which are usually regarded as being incompatible with one another. Moreover, it leads to the problem why referential (rich) inflection cannot possibly be added by the lexical inflectional rules operating before syntax, which would preclude V-to-I after all.

In this paper I will give an account of the trigger for V-to-I from which it follows that relatively rich inflection involves a distinct inflectional node whereas relatively poor inflection is base-generated on the verb, as in Rohrbacher's (1994) account, without having to assume that there are two fundamentally different types of inflection, one referential and the other nonreferential.[3] (In fact, it will turn out that rich inflection partly involves base-generation too). I will argue that the relation between V-to-I and richness of inflection is caused by the fact that having V-to-I and *not* having rich inflection are alternative ways of avoiding a violation of the same constraint, namely a general constraint that disfavours branching structure below the word (X^0) level. This constraint is a ranked, violable constraint, as in the optimality-theoretic conception of grammar (cf. Prince & Smolensky 1993). Its interaction with two other relevant constraints (concerning economy of movement and realization of the input in the output) will explain that languages can vary in their tolerance level of the amount of inflection on unmoved verbs. The analysis will also straightforwardly account for the English "V-to-I for auxiliaries only" puzzle.

2. V-to-I as a means to avoid branching under X^0

Grimshaw (1997:382) suggests that the difference between a V-to-I language like French and a non-V-to-I language like English results from a different ranking of the following two constraints:

(5) No-Lex-Mvt
A lexical head cannot move

(6) No-Morphology
All syntactic heads contain exactly one morpheme

In English, (5) outranks (6), with the result that the lexical verb remains in its base position. This means that the inflection must be generated directly on the verb in its basic V-position. Thus (6) is violated in order to make satisfaction of higher-ranked (5) possible.[4] In contrast, (6) outranks (5) in French. A violation of (6) then is avoided by inserting the inflectional morpheme in a separate head. The lexical verb must move to this I-position to pick up its inflection, resulting in a violation of lower-ranked (5). Grimshaw does not discuss what happens when there is more than one inflectional morpheme; following the logic of her argument, it seems these must be generated in a separate head each in a language with ranking (6) >> (5), like French.

The proposal to be made below at its core is a development of this account, but there are some problems with Grimshaw's proposal as it stands.

First, given its formulation, the constraint in (6) appears to be based on the idea that actual morphemes (instead of just features) are inserted under functional heads in syntax. However, there a number of disadvantages to generating actual affixes in functional heads (see Stump 1998:37-38 and references cited there). Below I will assume that there is a radical distinction between syntactic positions containing features and their realization by morphemes at PF (see for instance Halle & Marantz 1993 for discussion).

Second, it seems that after V-to-I movement has taken place in a language like French, (6) is violated after all. That would mean (6) must be valid at D-structure but not at S-structure. This is a rather unattractive state of affairs, however, given that the status of D-structure is unclear in the usual conception of optimality theoretic syntax, where constraints operate on surface representations.

Third, and most important for the present purpose, it is not clear why there is a relation between V-to-I (ranking (6) above (5)) and richness of inflection. Grimshaw remarks that "this should now be understood as reflecting a relationship between properties of inflection and existence as an independently projecting head", but does not indicate what causes this relationship.

I will argue that an answer to this problem becomes possible if the constraint in (6) is replaced by (7). (The constraint in (7) is related, though not identical, to Neeleman's (1994) "Complexity Constraint" and Roberts's (1997) *[$_{\text{Word}}$ Word Word] constraint).

(7) No-Branching
*Dominate(X^0, Sisters(Y,Z)), where Y and Z are distinct from X^0

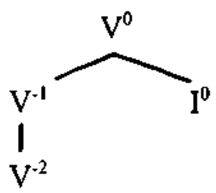
The constraint in (7) is violated by every triple in a tree structure of a zero-level node A and a pair of sister nodes B and C distinct from A when A dominates B and C.

Consider now the structure of an inflected verb when the inflection is base-generated on the verb, as in (8).

(8)

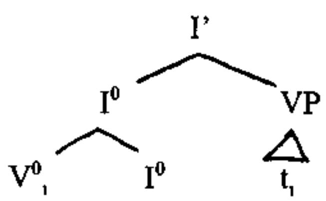
This structure violates (7) if the lower V is a distinct category from the higher V. Elsewhere I have argued that the I-node in (8) is not adjoined to V, but is a specifier within a subzero verbal projection, as in (10) (see Ackema 1998).

(9)



V^{-1} and I^0 are distinct from V^0, which dominates these nodes. Hence, (7) is violated.[5] As in Grimshaw's proposal, the trigger for V-to-I therefore must be that it avoids a violation of (7) by generating the inflection in a distinct head position. As noted, it must be ensured that after V-to-I the relevant condition is not violated after all. Indeed, in the structure that results after V-to-I, as in (10), (7) is not violated, since the two I-segments that are created as a result of adjunction are standardly considered to be instances of the same category, not distinct categories (Chomsky 1986, 1995, cf. also footnote 4).

(10)



The constraint in (7), which in essence penalizes structure building by base-generation (or Merge in current terms), may seem odd. It can be regarded as an "economy of merger" counterpart to economy of movement. Just like economy of movement says movement should be minimized, the constraint in (7) in effect says merger should be minimized. The only special thing about (7) is that it specifically mentions merger below the X^0-level. Of course, (7) may just be a special instance of a more general "do not merge" constraint (compare the general *STRUC ("no structure") constraint of Prince & Smolensky 1993:25), but as the implications of assuming such a general constraint go well beyond the scope of this paper I will stick to the definition in (7).[6]

We now come to the main issue, concerning the question why there is a correlation between V-to-I and rich inflection.

3. The relation between V-to-I and rich inflection

In the previous section it was argued, basically following Grimshaw (1997), that V-to-I avoids the violation of (7) that a verb with base-generated inflection causes. However, there is an alternative way of avoiding such a violation: leaving out the inflection. If V^0 does not need to branch in order to accommodate an inflectional node, a violation of (7) is avoided as well. Although things cannot be as simple as this (languages without V-to-I do not show total absence of inflection), this is the basic idea behind the account to be proposed here: V-to-I and leaving out inflection are two alternative solutions to exactly the same problem, hence the fact that languages without V-to-I usually have poorer inflection than languages with V-to-I automatically follows.

Before I come to the complication that non-V-to-I languages do have some inflection, another issue must be discussed first, namely why one of the options to avoid violation of (7) is not universally preferred over the other. As discussed in the previous section, a V-to-I structure like (10) satisfies (7). However, it violates No-Lex-Mvt in (5) (hence, as in Grimshaw's account, it must be assumed that in V-to-I languages No-Lex-Mvt is ranked below the constraint with which it interacts with respect to V-to-I, here assumed to be the No-Branching constraint in (7)). If the inflection is left out and the verb is not moved, both No-Branching in (7) and No-Lex-Mvt in (5) are satisfied. So why is this latter option not always preferred over the first one?

Obviously, this must be because leaving out some element from the structure is penalized as well. If we suppose, reasonably enough, that person, number and tense features are fully specified in the input, then not expressing such features in syntax violates the well-known Parse constraint of optimality theory, which applied to this particular case reads as in (11).

(11) Parse-Phi
Parse the phi-features in the input

This constraint is violated when the phi-features in the input are not fully parsed (see also Vikner 1997b).[7]

The basic account then is as follows. Languages in which Parse-Phi is ranked above No-Lex-Mvt have rich inflection and avoid a violation of (7) via V-to-I movement. Languages in which No-Lex-Mvt is ranked above Parse-Phi do not have V-to-I and avoid a violation of (7) by not parsing inflection.

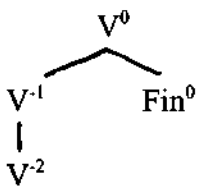
So, completely leaving out inflection is a means to satisfy No-Branching. However, this cannot be the end of the story. In languages that do not have V-to-I some inflection can appear on the verb. Yet any base-generated inflection under V^0 leads to a violation of No-Branching, as discussed in section 2. Also, another problem that was mentioned in the introduction still exists, concerning the fact that not all languages with V-to-I have the same number of distinctions in their inflectional paradigm; in fact, there is a language with the bare minimum that does have this movement (Kronoby Swedish).

In order to make solutions for these problems possible, it must first be noted that No-Branching is not an absolute constraint but a gradual one. It is not a constraint that is either violated or not, but a constraint that can be violated to different degrees. As stated in section 2, it is violated once by every triple in a tree structure of a zero-level node A and a pair of sister nodes B and C distinct from A when A dominates B and C. Informally, more (base-generated) branching structure leads to more violations of (7).

I assume that for every opposition that is present in the inflectional paradigm of a language an inflectional node must be present in the structure that expresses this opposition. Furthermore, I assume, as is fairly standard, that inflectional features express binary oppositions only (so there cannot be monolithic [1], [2] and [3] person features). For concreteness, I will adopt a feature system based on that of Kerstens 1993, including the binary inflectional features [+/-ad(dressee)] and [+/-sp(eaker] besides [+/-num(ber)] and [+/- fin(ite)]. Then 1st person is [+sp], 2nd person is [-sp, +ad] and 3rd person is [-sp, -ad]). Note that 1st person is not specified for [ad] since its specification as [+sp] already discriminates it from 2nd and 3d person. (This will become important in the discussion on English, see section 4).

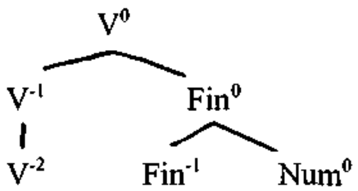
To illustrate, if in a paradigm the finite verb forms differ from the infinitive, the tree in (12) occurs in case inflection is base-generated on the verb (cf. (9)).

(12)



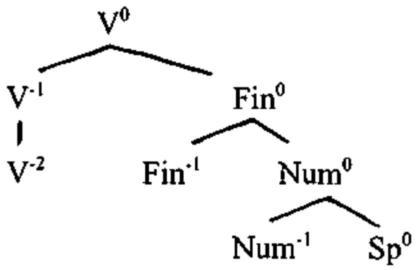
If in addition singular endings differ from plural ones we get (13) instead of (12).

(13)



If also, for example, 1st person differs from the other persons we must have (14).

(14)



Every terminal inflectional node has a + or a - value. The string of terminal nodes in the tree corresponds at PF to one (in case of fusion) or more (in case of (partial) agglutination) morphemes from the inflectional paradigm. Clearly, (14) is worse with respect to No-Branching than is (13), which in turn is worse than (12). To be precise, whereas (12) violates No-Branching once, (13) violates it three times (once for each of the triples (V^0, V^{-1}, Fin^0), (V^0, Fin^{-1}, Num^0) and (Fin^0, Fin^{-1}, Num^0)) and (14) violates No-Branching six times (once for each of the triples (V^0, V^{-1}, Fin^0), (V^0, Fin^{-1}, Num^0), (V^0, Num^{-1}, Sp^0), (Fin^0, Fin^{-1}, Num^0), (Fin^0, Num^{-1}, Sp^0), and (Num^0, Num^{-1}, Sp^0)).

As yet, the fact that No-Branching is a gradual constraint does not help in explaining the possible presence of inflection in non-V-to-I languages. If No-Branching is ranked higher than the other relevant constraints, as assumed in the account of V-to-I versus non-V-to-I above, then any violation of it must be avoided when possible, given the absolute priority of higher constraints over lower ones in optimality theory. Therefore (12), simple as it is, should already be banned. Instead of creating this structure, there should either be V-to-I or all inflection should be left out. Hence, it is still predicted (wrongly) that V-to-I is absent only in languages without any inflection.

I will argue now that both the fact that absence of V-to-I does not imply complete absence of inflection and the fact that languages differ in precisely how much inflection they tolerate on verbs in situ fall out naturally if constraint conjunction as proposed by Smolensky (1995), is applied to the constraint in (7).

Smolensky argues that, given two constraints A and B, a new constraint can be created by conjoining A and B. The resulting constraint A&B is violated in case both A and B are violated but not otherwise. Crucially (for our purposes), a constraint can also be conjoined with itself (cf. also Legendre et al. 1998). If a constraint C is conjoined with itself, the result is a new constraint, call it C^2, which is violated whenever C is violated twice. But C^2 is not violated when C is violated only once (or when C is not violated at all). C^2 can be conjoined with C again. This results in constraint C^3 which is violated whenever C is violated three times, but not when C is violated less than three times. And so on. Crucially, C^N can be in a different position in the constraint hierarchy than C is. Specifically, C^N can be in a higher position than C^{N-1}, with crucially interacting constraints in between the two.[8]

If "auto-conjunction" like this is applied to the No-Branching constraint in (7), the result is that the constraints No-Branching^2 (violated whenever No-Branching is violated twice), No-Branching^3 (violated when No-Branching is violated three times), etc., can be part of the constraint hierarchy as well.

Now, as noted, the structure in (12), which can occur in non-V-to-I languages, violates No-Branching only once. Hence, it does not violate No-Branching^{1+N} (N>0). That V-to-I need not be triggered in this case then is explained if it is not a violation of No-Branching itself but of a higher order No-Branching constraint that is to be avoided.

Consider for instance the following constraint ranking:

(15) No-Branching^3 >> No-Lex-Mvt >> Parse-Phi >> No-Branching

Because both No-Lex-Mvt and Parse-Phi outrank No-Branching, a single inflectional Fin-node can be base-generated on the verb (violating No-Branching but satisfying Parse-Fin), without there being V-to-I (so that No-Lex-Mvt is also satisfied). But if a second phi-feature is parsed, the resulting structure either violates No-Branching^3 or No-

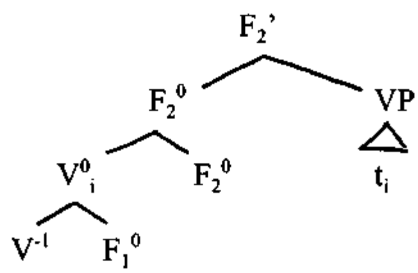
Lex-Mvt. If the second feature is base-generated on the verb as well, No-Branching^3 is violated (since No-Branching is violated three times in that case, see for instance (13) and the discussion below it). If the second feature is generated in a distinct head and there is V-movement to this head No-Lex-Mvt is violated. Given that No-Branching^3 and No-Lex-Mvt outrank Parse-Phi in (15), the second feature will not be parsed at all given this ranking. Hence only one feature is parsed, with the result that the verb can be specified as finite (distinct from the infinitive) but is underspecified as to person and number.

Compare this with the ranking in (16).

(16) No-Branching^3 >> Parse-Phi >> No-Lex-Mvt >> No-Branching

As with the ranking in (15), one inflectional node can be base-generated on the verb, but again it is impossible to let this node branch into nodes for two distinct phi-features, because this would incur a fatal violation of NoBranching^3. But since in (16) Parse-Phi outranks No-Lex-Mvt it is now more important to parse a second phi-feature than it is not to have verb movement. Hence, the second feature is generated in a distinct head and there will be verb movement to this head. So we get a representation as in (17) (where I have left the particular nature of the phi-features unspecified), which satisfies No-Branching^3 (No-Branching is violated twice, by the triplets (F_2^0, V^{-1}, F_1^0) and (V^0, V^{-1}, F_1^0); recall that the two F_2^0-nodes created by V-adjunction are nondistinct, so that the triple (F_2^0, V^0, F_2^0) does not violate (7)).

(17)

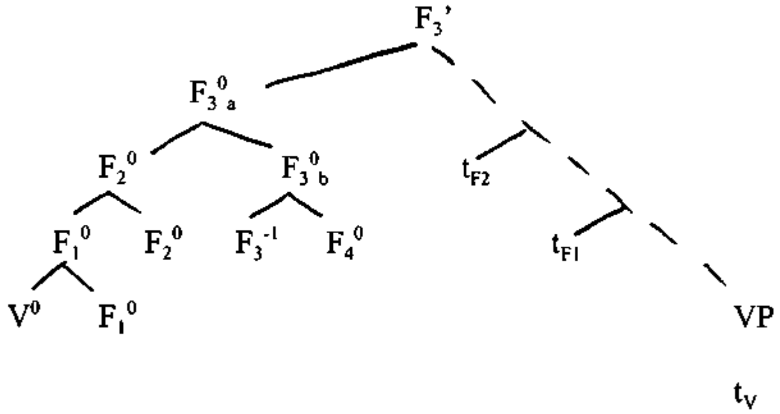


If, in contrast to (15)-(16), No-Branching^3 is ranked below Parse-Phi and No-Lex-Mvt, two distinct phi-features are allowed on a verb in situ. Parsing a third feature, however, leads to six instead of three No-Branching violations (see (14) and the discussion there), so that this will be impossible again if No-Branching^6 is ranked high, as in (18).

(18) No-Branching^6 >> Parse-Phi >> No-Lex-Mvt >> No-Branching^3

The ranking in (18) has an interesting consequence. As just explained, given this ranking two phi-features can be base-generated on the verb. To avoid a violation of No-Branching^6 and Parse-Phi, a third feature is generated in a distinct head to which there will be verb movement. Interestingly, a derivation is possible in this case in which even a fourth phi-feature can be parsed without there being a violation of No-Branching^6. This is possible by generating F_1, F_2 and F_3 in distinct heads, with successive cyclic head movement to these nodes, and base-adjoining F_4 to F_3. The resulting surface representation is given in (19).

(19)



This representation violates No-Branching five times (because of ($F_{3\,a}^0$, F_1^0, F_2^0), ($F_{3\,a}^0$, V^0, F_1^0), (F_2^0, V^0, F_1^0), ($F_{3\,a}^0$, F_3^{-1}, F_4^0) and ($F_{3\,b}^0$, F_3^{-1}, F_4^0); all other triplets involve nondistinct nodes), so No-Branching^6 is satisfied. Given that this derivation exists, the fourth phi-feature is in fact parsed obligatorily, or else a fatal violation of Parse-Phi results (cf. (18)). Hence, although V-to-I is triggered already by the impossibility of base-generating three features on the verb (in languages characterized by (18)), the result is that all four features Fin, Num, Sp and Ad must be parsed obligatorily. This explains why the generalization in (20), from Rohrbacher 1994, is adequate for a large part in the Germanic languages:

(20) A language has V-to-I raising iff in at least one number of one tense of the regular verbs, the person features [1st] and [2nd] are both distinctively marked

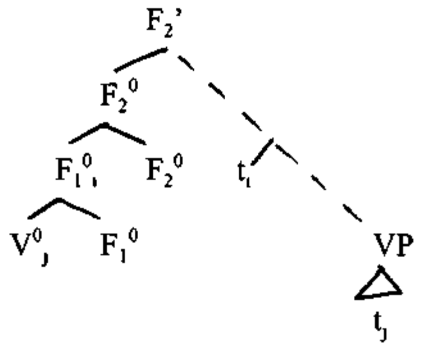
If in at least one number of one tense 1st and 2nd person forms differ from 3rd person, from the infinitive and from each other, this indicates that indeed all phi-features [fin], [num], [sp], and [ad] are parsed.[9]

The fact that, as mentioned in the introduction, there are exceptions to (20) is explained as well. In the grammar of these 'exceptional' languages it is not No-Branching^6 but a lower order No-Branching constraint that is ranked above No-Lex-Mvt and Parse-Phi and hence can trigger V-to-I (and actually does trigger V-to-I if also Parse-Phi >> No-Lex-Mvt) in case of poorer inflection, see the discussion around (15)-(16). In Kronoby Swedish, for instance, already No-Branching itself outranks Parse-Phi and No-Lex-Mvt and the single phi-feature that is parsed must be generated in a distinct head, as in (10). (See also section 4 below.) So, a language with very poor inflection that yet has V-to-I has the partial grammar in (21).

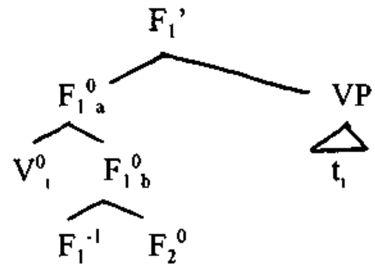
(21) No-Branching >> Parse-Phi >> No-Lex-Mvt

At first sight, we might still expect rich inflection in a language with the constraint ranking in (21), because of the possibilities of generating every feature to be parsed in a distinct head position or letting the additional head(s) branch itself (as discussed in connection with the example in (19)). That would rob us of our explanation of exceptions to (20). However, already if we try to parse a second feature in a language characterized by the ranking in (21), either in a distinct head position (as in (22a)) or as an adjunct to the head position of the first feature (as in (22b)), representations with complex heads result that do in fact violate No-Branching.

(22) a.



b.



In (22a) No-Branching is violated once ((F_2^0, V^0, F_1^0)), in (22b) it is violated twice (($F_{1\ a}^{0}$, F_1^{-1}, F_2^0) and ($F_{1\ b}^{0}$, F_1^{-1}, F_2^0)). Hence, despite the partial Parse-Phi >> No-Lex-Mvt ranking, there is no other option to satisfy No-Branching than by leaving further features unparsed. But since one feature can be parsed without violating No-Branching by having V-to-I, V-to-I is triggered and this one feature is parsed.

Concluding, how much inflection a language tolerates on a verb in situ, i.e. before V-to-I is triggered or before inflection is left out, depends on the lowest No-Branching^N constraint that still dominates either Parse-Phi or No-Lex-Mvt (recall from footnote 7 that universally C^N >> C^{N-1}). If base-generating another phi-feature on the verb leads to a violation of this particular constraint, then either V-to-I is triggered (if Parse-Phi >> No-Lex-Mvt) or the feature is left unparsed (if No-Lex-Mvt >> Parse-Phi). This accounts for the general correlation between V-to-I and a richer inflectional paradigm, while at the same time explaining that there is not a rigid, but a fluid, inflectional demarcation line between V-to-I and non-V-to-I languages.

4. Typology

In this section I will show how the observed language typology with respect to V-to-I is derived by different rankings of the constraints, as discussed in the abstract in the previous section.

Consider the ranking in (18) again, repeated below in (23).

(23) No-Branching^6 >> Parse-Phi >> No-Lex-Mvt >> No-Branching^3

As discussed in section 3, it results in a language which has V-to-I and which parses all phi-features. This is indicated by at least one 1st person form that is distinct from the other persons (+/-sp), at least one 2nd person form that is distinct from 3d person (+/-ad), at least one singular form that is distinct from the corresponding plural (+/-num) and by an infinitive ending (+/-fin) (if the infinitive differs from the stem an SS-PF correspondence rule mentioning [-fin] is necessary; note that if there are different person and number endings but no distinct infinitive ending, no correspondence rule mentioning the [fin] feature is necessary, since +finite is implied by the presence of [num], [sp] and/or [ad] features).[10]

Languages conforming to this pattern are for instance Icelandic and French. Both have V-to-I (cf. Vikner 1995, Emonds 1978) and they have the following paradigms:

(24)

	Icelandic		French	
	inf. -a		inf. -er	
	sing	plur	sing	plur
1	-a	-um	-e	-ons
2	-ar	-id	-e(s)	-ez
3	-ar	-a	-e	-e(nt)

The SS-PF correspondence rules that are minimally necessary for the data in (24) are given in (25).

(25) a. (Ice) [-num,+sp] → -a , [-num,-sp] → -ar, [+num,+sp] → -um, [+num,-sp,+ad] → -id, [+num,-sp,-ad] → -a , [-fin] → -a
b. (Fr) [-num] → -e , [+num,+sp] → -ons , [+num,-sp,+ad] → -ez , [+num,-sp,-ad] → -e , [-fin] → -er

Indeed all phi-features are referred to in these minimally necessary rules, which means these features must have been parsed in syntax. This implies that the paradigms in (24) show instances of accidental homophony (instead of syncretism that results from not parsing phi-features). For example, given that all phi-features are parsed, 2sg and 3sg in Icelandic must be [-num,-sp,-ad] and [-num,-sp,+ad] respectively, so that there are in fact two correspondence rules (instead of one that mentions just [-num,-sp]) that happen to relate these features to the same affix *-ar* at PF.

Consider next languages that also have V-to-I but have poorer inflection than French and Icelandic. As discussed in section 3, such languages are characterized by a ranking that is like the one in (23), but with a lower order No-Branching constraint on top. The most extreme case is instantiated by Kronoby Swedish. This language has V-to-I but the inflectional endings for finite forms, although distinct from the infinitive, do not differ from one another (Platzack & Holmberg 1989). Hence, the correspondence rules need to mention just one feature, namely [+/-fin]. This language has the partial grammar in (21), repeated here in (26).

(26) No-Branching >> Parse-Phi >> No-Lex-Mvt

This grammar leads to one inflectional [+/-fin]-node being generated in a distinct head plus V-movement. The resulting structure satisfies No-Branching, see (10) and the discussion there. However, parsing more features will result in No-Branching being violated, even if these features were also generated in distinct heads, as discussed in connection with (22).

Note that, given this discussion, in V-to-I languages parsing at least one ([+/-fin]) feature must always be possible, hence obligatory. On the other hand, in some non-V-to-I languages, namely those characterized by the ranking in (27), inflectional endings are banned altogether.

(27) No-Branching[1] >> No-Lex-Mvt >> Parse-Phi

So there is in fact a minimum amount of inflection for V-to-I languages, namely a distinct form for finite verbs. If a language has only one verbal form overall it should not have V-to-I. According to Sten Vikner (p.c.) such a language is indeed unattested.

Let us now turn to non-V-to-I languages more generally. Consider first a non-V-to-I language with poor inflection, like Danish (cf. Vikner 1995):

(28) Danish
inf. -e

	sing	plur
1	-er	-er
2	-er	-er
3	-er	-er

The following SS-PF correspondence rules suffice to account for this paradigm:

(29) [+fin] → -er , [-fin] → -e

We see that (as in Kronoby Swedish) only the [+/-fin] feature is mentioned. Absence of V-to-I with this paradigm follows if the Danish grammar contains the ranking in (30):

(30) No-Branching^3 >> No-Lex-Mvt >> Parse-Phi >> No-Branching

One feature can be base-generated on the verb, but further features are left unparsed in order to avoid violations of either No-Branching^3 or No-Lex-Mvt. If this analysis is correct the Danish paradigm shows real syncretism, in the sense that the single *-er* affix just represents [+fin], not num, sp or ad features.

Consider next English, which also lacks V-to-I, but allows for one more inflectional feature to be parsed on the in-situ verb. In comparison to (30), this language also has a partial ranking No-Lex-Mvt >> Parse-Phi, but higher up with respect to the hierarchy of No-Branching constraints. In particular, English has the ranking in (31).

(31) No-Branching^6 >> No-Lex-Mvt >> Parse-Phi >> No-Branching^3

The English paradigm is given in (32), its SS-PF correspondence rules in (33).

(32) English
inf. 0

	sing	plur
1	0	0
2	0	0
3	-s	0

(33) [-num,-ad] → -s[11]

The one correspondence rule in (33) suffices; note that no rule mentioning [+/-fin] is needed since the infinitive has no distinct ending.[12]

English does give rise to some extra complications. One, V-to-I for auxiliaries, will be discussed in the next section. The other concerns the occurrence of *do*-support.

In cases of *do*-support, as in cases of V-to-I, there is an extra head position in the structure. However, the reason for generating this extra head is not the same as the reason for generating an extra head position in cases of V-to-I. Exactly what reason this is is extensively discussed by Grimshaw (1997) and I will not go into it here. The potential problem for the account of V-to-I proposed above concerns the reason why, in case there *is* such an extra head position in English, *do* is inserted there instead of there being verb movement. Grimshaw argues that this is a consequence of the interaction between No-Lex-Mvt and the constraint Full-Int(erpretation). Insertion of dummy *do* violates Full-Int, but makes movement of the main verb unnecessary. Therefore, if No-Lex-Mvt >> Full-Int we get *do*-support in case a higher head position must be filled.

This raises the question why, in case the higher head position is generated in order to avoid a violation of a No-Branching constraint (in a V-to-I language, in contrast to English), the *do*-insertion option is never chosen over the option of moving the main verb. (If *do*-support would be an alternative for V-to-I an unattested language would be possible in which main verbs cannot appear in finite form).[13] The answer is that, if avoidance of a No-Branching violation is the trigger for generating the extra head, base-generating fully inflected *do* in that head position does not help, since the relevant No-Branching constraint is violated all the same then. Only by moving a less inflected verb to a distinct inflectional node is No-Branching violated less, as argued in section 3. This explains why there are no languages that have obligatory *do*-support in all finite clauses.

Faroese constitutes an interesting case for the typology developed in this section. As noted above, there are two dialects of this language, which have the same verbal paradigm but differ with respect to allowing V-to-I or not (Jonas 1996, Bobaljik & Thraínsson 1997). For obvious reasons, such a situation is hard to account for within theories in which a given inflectional paradigm obligatorily leads either to V-to-I or to its absence. Bobaljik & Thraínsson (1997:17) argue that an alternative is necessary in which "it is not the "richness" of agreement morphology that is important, but rather the more general structure of the verbal morphology. In particular it is the (im)possibility of multiple inflectional morphemes on the finite verb which is important, and not the features which these morphemes express". The analysis proposed here is of course of this type, and the distinction between the two dialects of Faroese can indeed be made to follow.

The verbal paradigm for Faroese is given in (34). The SS-PF correspondence rules necessary to account for this paradigm are given in (35).

(34) Faroese
inf. -a

	sing	plur
1	-i	-a
2	-ir	-a
3	-ir	-a

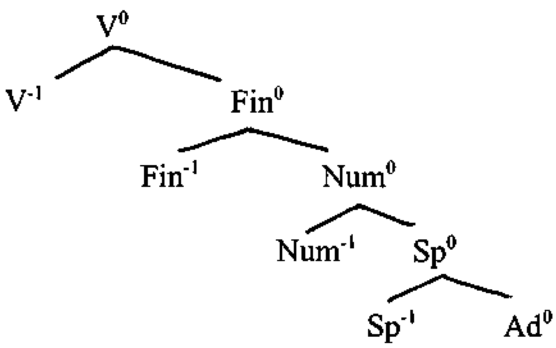
(35) [+sp, -num] → -i, [-sp, -num] → -ir, [+num] → -a, [-fin] → -a

Faroese I, which has V-to-I, has a grammar in which both Parse-Phi and No-Branching^{6} outrank No-Lex-Mvt, as in (36) for instance. This ranking forces three[14] features to be parsed, which at the same time cannot be base-generated on the verb (see section 3). Hence V-to-I occurs. Faroese II, which lacks V-to-I, differs only in that here No-Lex-Mvt >> No-Branching^{6}, as in (37), which makes it better to base-generate the three inflectional features on the verb than to apply verb movement.

(36) Parse-Phi >> No-Branching^{6} >> No-Lex-Mvt (Faroese I)
(37) Parse-Phi >> No-Lex-Mvt >> No-Branching^{6} (Faroese II)

A final prediction with respect to typology that follows from the analysis is that the correlation between absence of V-to-I and poor inflection should break down in languages or language families that are characterized by a grammar in which Parse-Phi and No-Lex-Mvt both outrank a high order No-Branching constraint. In particular, if Parse-Phi and No-Lex-Mvt both outrank No-Branching^{10} (and consequently all lower No-Branching constraints), the structure in (38), with all phi-features base-generated on the verb, will occur.

(38)



So, in principle languages can exist which have rich inflection but which nevertheless lack V-to-I. It is unclear to me whether this is a desirable prediction or not. There appear to be some languages that might confirm it. The Bantu language Kirundi (VO), for example, has a rich agreement system, but the object must be adjacent to the verb when unfocused (Ndayiragije 1996:273-274), adverbs do not intervene. Similarly, Bresnan & Mchombo (1987) note that in Chichewa, another Bantu language with rich subject-verb agreement, "the object immediately follows the verb" when the object is not an incorporated pronoun. This might indicate absence of V-to-I, but of course only if the syntax of adverbs in these languages is such that some adverbs can be taken to be in a VP-peripheral position and there are indeed no V-Adv-Obj orders with these adverbs.

If languages with rich inflection on in situ verbs are rare (but as noted this is unclear to me), this may be related to the fact that in the ranking giving rise to this a high order No-Branching constraint must crucially be present, namely No-Branching^{10} (as noted, it must be the case that No-Lex-Mvt>>Parse-Phi>>No-Branching ^{10}, since in (38) No-Branching is violated ten times). Now, it does not seem unreasonable to suppose that there is a certain cost attached to adopting (self-) conjoined constraints in a ranking. In contrast to the constraints themselves, conjoined constraints cannot be given by UG, for the simple reason that there is an infinity of them. Hence, whereas in case of simple constraints the only task for the language learning child is to rank them so that the data are accounted for (see for instance Tesar & Smolensky 1998), in case of data that involve crucial interaction of constraints with a conjoined constraint the child must also adopt the conjoined constraint itself. This means every application of constraint conjunction brings with it some cost in terms of acquisitional load, in that an extra constraint must be learned. This in turn means that (38), which only arises under a ranking mentioning a constraint that results from applying constraint conjunction ten times, should be a marked structure.

If this is correct, then the structures that should be most unmarked are those that arise in languages with grammars containing only No-Branching itself, and no conjunctions of it. As discussed in section 3, such languages have very poor inflection or no inflection at all (either there is only a distinction between finite and infinite verbs plus V-to-I, or there is just one verb form and no V-to-I). If the contention that creole languages show unmarked characteristics (cf. Bickerton 1984) makes some sense then the prediction that poor inflection on the verb should be unmarked is correct, given that creole languages often display poor inflection.[15]

5. Further evidence for the No-Branching constraint: V-to-I and auxiliaries

In this section I will supply some additional evidence for the No-Branching constraint, which I argued to be the trigger of V-to-I. I will argue that the adoption of this constraint makes possible a straightforward account of why V-to-I applies to auxiliaries but not main verbs in English.

Auxiliaries undergo V-to-I in English (cf. Pollock 1989, Roberts 1998), in contrast to the mainland Scandinavian languages, which are non-V-to-I for main verbs and auxiliaries alike (cf. Vikner 1995). This fact is of course unexpected for accounts that regard V-to-I as a necessary consequence of rich inflection, since the inflection of a verb like *have* is not richer than that of main verbs (for modals it is even poorer). Let us see if the analysis proposed here can account for this, and for the difference between English and Scandinavian in this respect.

In Ackema 1998 it is argued that auxiliaries and the main verbs they select are base-generated as a complex predicate (in order to explain the thematic properties of such structures), as in (39).

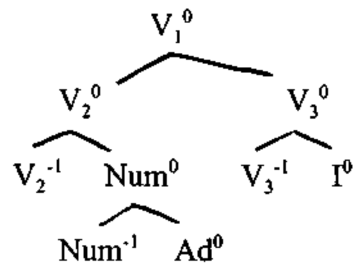
(39) V^0

$V_{aux}{}^0$ $V_{main}{}^0$

Empirical evidence can be given for the assumption that auxiliary and main verb do not head distinct projections in the base, based on facts from embedded clauses in Dutch as discussed in Reuland 1990 and Ackema et al. 1993. It is further argued in Ackema 1998 that the auxiliary can excorporate from a structure as in (39), but no trigger for V-to-I of auxiliaries in English is given there.

Given the proposal made above, a trigger clearly presents itself in the form of the No-Branching constraint. If the inflection were base-generated on the verbs in a verbal complex like (39), as it is base-generated on single main verbs in English, then a structure as in (40) results (where I^0 stands for the inflectional features of the main verb, for instance a participial feature (like [+perf]) if the auxiliary is perfect *have*; the subscripts on the V-nodes serve to distinguish these nodes and have no theoretical significance).

(40)



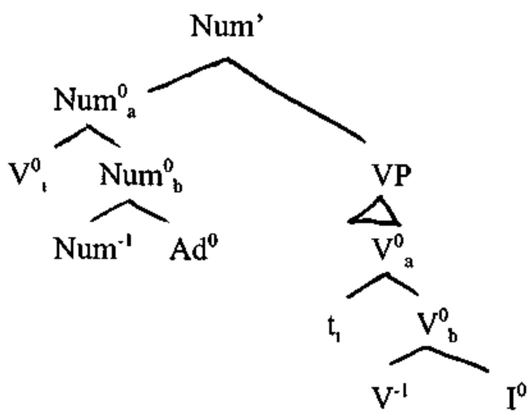
In this structure No-Branching is violated eight times: (V_1^0 ,V_2^0 ,V_3^0), (V_1^0, V_2^{-1}, Num^0), (V_1^0, Num^{-1}, Ad^0), (V_2^0, V_2^{-1}, Num^0), (V_2^0, Num^{-1}, Ad^0), (Num^0, Num^{-1}, Ad^0), (V_1^0, V_3^{-1}, I^0) and (V_3^0, V_3^{-1}, I^0) violate it. The constraint ranking that was established for English in section 4 (cf. (31)) is repeated below in (41).

(41) No-Branching^6 >> No-Lex-Mvt >> Parse-Phi >> No-Branching^3

Given the eight violations of No-Branching just mentioned, the structure in (40) violates the highest constraint in (41), so if a better alternative exists in this respect this alternative should be chosen. Because No-Lex-Mvt outranks Parse-Phi the option that must normally be chosen in English to avoid violation of No-Branching^6 is underparsing of phi-features (hence its relatively poor inflection). But that is no option to improve on (40), since in that case we would have to leave either the auxiliary or the main verb unparsed.[16] That is ruled out, however, as the resulting structure will differ in meaning from the one in which the relevant verb is parsed, hence the two alternatives do not compete in this case (this in contrast to structures which only differ in how many phi-features are parsed, see footnote 6). A (nonperfect) structure with only a main verb cannot block a periphrastic perfect, for example.[17]

In case underparsing is no option, the next best thing with respect to the ranking in (41) is to avoid a fatal violation of No-Branching^6 by V-to-I of the auxiliary. This results in the structure in (42). In this structure No-Branching is violated only four times[18] (by Num^0_a, Num^{-1}, Ad^0), (Num^0_b, Num^{-1}, Ad^0), (V^0_a, V^{-1}, I^0) and (V^0_b, V^{-1}, I^0)), so No-Branching^6 is satisfied, as desired.[19]

(42)



Hence, the fact that V-to-I can be triggered for auxiliaries but not main verbs in a language is a consequence of the fact that periphrastic constructions induce extra violations of No-Branching because of the complex predicate structure of main verb and auxiliary.

This account has the advantage that the constraint No-Lex-Mvt, which specifically mentions movement of lexical heads, can be generalized to No-Head-Mvt (or perhaps even to Stay), valid for lexical verbs and auxiliaries alike. No-Head-Mvt is violated by movement of the auxiliary in (42), but as argued, this is the only option to save a structure with an auxiliary from fatally violating No-Branching^6.

One may wonder whether a distinct No-Lex-Mvt constraint, which does not penalize movement of auxiliaries, is necessary anyway to account for the non-occurrence of *do*-support in sentences with an auxiliary:

(43) a. John did not go to school today
b. *John did not have gone to school today
c. John has not gone to school today

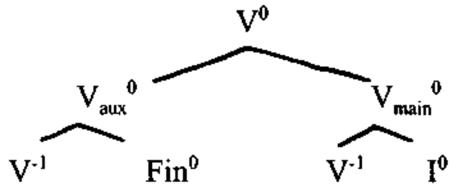
However, this fact can also be explained without a movement constraint that specifically mentions lexical verbs as opposed to auxiliaries. As just explained, in case there is an auxiliary present in English, this auxiliary must be excorporated and moved to a distinct head. If such a distinct head must be generated anyway in case of auxiliaries, than in those cases in which an extra head would be necessary when there is no auxiliary (i.e. those cases triggering *do*-support with main verbs) there is no need to generate a second extra head. The one extra functional head position suffices to avoid a No-Branching^6 violation (see above) and at the same time can act as head of the extra projection that is needed in English under the circumstances otherwise triggering *do*-support, e.g. negation (see Grimshaw 1997 for discussion). Therefore, adding *do* does not improve things when there is an auxiliary. In particular, it does not avoid a violation of No-Head-Mvt, since No-Head-Mvt must be violated in any case when there is an auxiliary in order to avoid violation of higher-ranked No-Branching^6. Since *do* violates Full-Int it will therefore never be inserted when there is an auxiliary. Only in case there is just a main verb, i.e. in case No-Branching^6 is satisfied without V-movement being necessary, *do*-insertion has the function of avoiding violation of No-Head-Mvt.[20]

This leaves unexplained as yet why the mainland Scandinavian languages do not have V-to-I for auxiliaries.[21] Let us consider the ranking that was established for Danish in section 4 again (cf. (30)), repeated here in (44) (where I have replaced No-Lex-Mvt by general No-Head-Mvt, in accordance with the argument above).

(44) No-Branching^3 >> No-Head-Mvt >> Parse-Phi >> No-Branching

A base-generated complex of main verb plus auxiliary, with the single inflectional node corresponding to *-er* on the finite auxiliary (cf. section 4) and a single inflectional (participial or infinitival) node on the main verb has the structure in (45).

(45)



This structure violates No-Branching four times. So the highest constraint in (44) is violated and, since underparsing of auxiliary or main verb is not an option, this should be avoided by generating the inflection in a distinct node and V-to-I of the auxiliary to this node, as in English.

This indicates that the ranking in (44) for mainland Scandinavian is in fact not quite correct. Instead, No-Head-Mvt must be the highest constraint:

(46) No-Head-Mvt >> No-Branching^3 >> Parse-Phi >> No-Branching^1

This has no consequences for the analysis of section 3-4. As long as both No-Branching^3 and No-Head-Mvt are ranked above Parse-Phi, underparsing of all but one phi-feature is the result; the mutual ranking between No-Branching^3 and No-Head-Mvt is irrelevant in this respect. But the absence of V-to-I for auxiliaries in Danish is now explained: it is more important not to move the auxiliary than it is to satisfy No-Branching^3.[22] Therefore, the structure in (45) is more optimal than one involving verb movement, despite its four No-Branching violations. The difference between non-V-to-I languages that do have V-to-I for auxiliaries (English) and total non-V-to-I languages (mainland Scandinavian) thus falls out as a consequence of the mutual ranking between general No-Head-Mvt and the first No-Branching constraint above Parse-Phi.

6. Conclusions

Constraints about faithfulness to the input (Parse) and about economy of representation (No-Branching) and economy of movement (No-Head-Mvt) play a role in various parts of grammar. I hope to have shown above that from the interplay of these three general types of constraint the fact that "rich" inflection can trigger V-to-I follows. However, rather than give a short restatement of the above analysis, let me point out in this final section that, if it is anywhere near correct, the following general points about aspects of grammar must hold.

6.1 Modularity

The analysis is based on the idea that phi-features are fully specified in the input for the syntactic generator, but can be left unparsed in the optimal syntactic structure. Between the syntactic module and the phonological module a similar mismatch can occur: even when some phi-feature *is* represented in the syntactic structure, this does not mean that a distinct phonological affix that corresponds exactly to this feature must occur in the phonological structure.[23] This picture fits best in a theory of grammar which assumes that the semantic/conceptual, syntactic and phonological modules are radically autonomous, in the sense that they each generate their own structures that are subject to their own wellformedness principles and that are in principle independent of one another. Correspondence between the various structures must be established via linking rules or principles, which can establish regular or idiosyncratic correspondences between semantics, syntax and phonology. Various models of grammar have been proposed that are partly or completely compatible with this; see Jackendoff 1997 for a recent example.

6.2 Syncretism versus homophony

Directly related to the modularity issue is the issue of syncretism versus homophony of affixes. As we have seen above, in some languages the use of the same ending for various persons and numbers is an indication of the fact that not all person (sp and ad) and number features are parsed in syntax, so that the phonological affix that occurs simply corresponds to a less rich feature bundle. In that case we have "real" syncretism (see for instance Plank 1991, Blevins 1995). But it is also possible that distinct feature bundles happen to be spelled out by the same affix (like the 2sg and 3sg in Icelandic, see the discussion below (25)), a case of accidental homophony. The two can be distinguished on the basis of the following two criteria. First, "real" syncretism must necessarily concern a natural class: the distinction that is lost must concern a particular feature num, ad or sp or a combination of these. Accidental homophony can possibly concern affixes that do not constitute a natural class (as with the 1sg, 2sg and all pl forms in English, which happen to be the same despite the fact that ad and num features are parsed in this language, see section 4). Second, in case some feature is left unparsed in syntax no syntactic principle can refer to that particular feature in the relevant language. In case two different feature bundles happen to be spelled out by an identical affix, syntactic principles can refer to those features. An example involving case-features of this difference between underparsing in syntax and "under-spelling-out" at PF is discussed in Ackema 1997.

6.3 Flexible structure

The above analysis is crucially based on the idea that there is no fixed functional clause structure in a language. A functional head position and its projection are only generated if this results in a structure that outperforms the competitors without the functional head with respect to the grammar of the relevant language. If no V-to-I is necessary no IP is present. That such flexibility must be allowed for has been argued before by for Haider 1989, Ackema et al. 1993, Grimshaw 1997 and Bobaljik & Thráinsson 1997, among others. In Ackema et al. 1993 it is assumed that even in case the extra projection is necessary it is not-base generated, but derived via verb movement and reprojection of the moved verb. This differs from the analysis above, for which it is crucial that in a V-to-I language part of the verb's inflectional features are directly generated in a distinct head.

Notes

1. For comments and discussion I would like to thank Johan Kerstens, Olaf Koeneman, Sten Vikner, Fred Weerman, and the audience at the 1998 *Tabu*-day at Groningen University.

2. The reverse of this problem, namely languages with rich inflection but without V-to-I movement is potentially even more problematic, since most of the accounts of the phenomenon given so far imply that failure to move the verb when the inflection is rich leads to ungrammaticality. Nevertheless, some Germanic OV-languages with rich inflection (German, Frisian) might in fact be non-V-to-I. Several arguments have been given against V-to-I raising in embedded clauses in Germanic OV-languages (cf. Reuland 1990, Haider 1993, Ackema et al. 1993, Koopman 1995), but the issue is difficult to decide, cf. Vikner (1995:152 ff.). See also section 4 below.

3. The difference between the rich inflection observed in pro-drop languages versus the poor inflection in non-pro-drop languages sometimes is also described in terms of referential versus nonreferential inflection (see for instance Rizzi 1982). If this is correct, this also pleads against treating the distinction between V-to-I-triggering inflection and non-V-to-I-triggering inflection in terms of referentiality, since languages with V-to-I can be non-pro-drop (e.g. French). If only in pro-drop languages the inflection is referential, this means that in non-pro-drop V-to-I languages the inflection would have to be "a bit but not quite referential", if referentiality of inflection is what is at stake in V-to-I versus non-V-to-I. Such a conclusion would not make sense, since something is referential or not.

4. This presupposes I-to-V lowering is impossible (pace Vikner 1997b), as I will assume throughout the paper. (Overt lowering presumably is ruled out altogether by independent principles of grammar, see Bošković & Takahashi 1998 for recent discussion). Hence, in a simple affirmative declarative sentence in English no IP is generated at all (cf. also Ackema et al. 1993).

5. But even when base-generated inflection would be adjoined to V, the lower V and higher V arguably are distinct categories. Coopmans (1988) supplies evidence, based on extraction facts within the *Barriers* theory, that a structure $[_X\ Y_i\ [_X\ ...t_i...]]$ that results from movement must involve two segments of a single X category, whereas a base-generated $[_X\ Y\ [_X\ ...]]$ structure must involve two distinct X categories.

6. Of course, in every language structure is built, so constraints on economy of merger must be counteracted by constraints that require realization of the input, namely the family of Parse constraints. The interaction of (7) and other constraints with Parse constraints allows for the situation that in some instances non-realization of the input is optimal. This accounts for some instances of "absolute ungrammaticality" (inputs for which no output at all exists in some language); see Ackema & Neeleman 1998 for discussion of this issue within syntax. (Note that if a general "do not merge" constraint outranks every other constraint, no input at all will be realized; for obvious functional reasons such a ranking does not correspond to the grammar of an actual natural language).

7. In Ackema & Neeleman 1998 it is argued that a structure in which not all semantically contentful words are parsed usually has a different semantics than one in which everything in the input is parsed. As a result, these two structures will not be in the same candidate set (under the assumption that candidate sets are defined in terms of nondistinct semantics) and hence will not compete. However, since no specification for a feature is nondistinct from any particular specification of that feature, in case of (phi-) features the structure with underparsing will compete with the structure in which all features are parsed. (Ackema & Neeleman in fact also discuss cases of competition between structures with different function words and inflection).

8. A ranking in which C is ranked higher than C^2, does not have different effects from a ranking with C only. This is because in all remaining candidates that survive evaluation of C, C is violated the same minimal number of times, so that lower-ranked C^2 cannot discriminate between these candidates. The reverse is not true: if C^2 is ranked above C, then if there are candidates in which C is violated either once

or not at all, these equally survive C^2; obviously, C can discriminate between such candidates. In general, adding C^N to a ranking can only have an empirical effect if it is ranked higher than C^{N-1} and some constraint(s) with which C interacts intervenes between the two.

9. Note that this does not necessarily mean that there is a distinct phonological form for each and every cell in the paradigm. Next to the real syncretism that results from underparsing of phi-features, as happens in non-V-to-I languages (and in some V-to-I languages, as explained in the text), there can be accidental syncretism in the form of homophonous affixes (cf. section 6.2).

10. The fact that infinitives are not specified for num, ad or sp might explain that they do not seem to undergo V-to-I even in V-to-I languages, given that the trigger for V-to-I proposed here is related to the parsing of these features. (Infinitives do move in some languages under some circumstances, see Thráinsson 1993 on Icelandic and Belletti 1990 on Italian, but this might involve V-to-C rather than V-to-I, cf. Johnson & Vikner 1994. I will not discuss this here).

11. Since this rule explicitly mentions [-ad] in its specification, 1st person is excluded from it. 1st person is not specified for the [ad] feature at all (cf. section 3). Its specification as [+speaker] implies that it must be -addressee, but [-ad] is not in its feature make up.

12. Despite the distinct *-ed* ending of past tense forms the rule in (32) need not mention [-past] either, given the plausible assumption that present tense is default so that a morpheme need not be specified as such (cf. Jensen 1990:141, Lieber 1992:101, Noyer 1993:6). Note that for the past tense a correspondence rule referring only to [+past] suffices, so No-Branching^6 is in no threat of being violated there either. Only in languages where the past tense shows richer inflection than the present tense may correspondence rules mentioning more features in the past tense be necessary - but none of the languages discussed here fall in that class (if there are, indeed, any such languages).

13. In nonstandard Dutch and German a *do*-periphrasis can occur instead of a simple finite main verb, but, apart from the fact that it is doubtful that this has anything to do with V-to-I (given that it is doubtful that Dutch and German are V-to-I languages, cf. footnote 1), this is always optional. For data and discussion see Tieken-Boon van Ostade 1990 (Dutch dialects, child language and historical data) and Erb 1995 (German dialects).

14. In fact, all four (see section 3), so in the grammar of speakers of Faroese I the forms in (34) are also specified for [+/-ad], although the distinctions in this respect are not expressed phonologically (a case of partial homonymy).

15. Though see Kegl et al. 1998 for a counterexample.

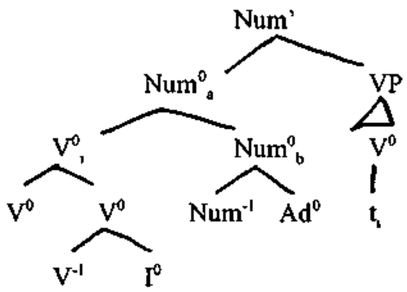
16. In fact there seems to be another option, namely parsing less inflectional features with finite auxiliaries than with finite main verbs. However, it is not possible to parse different amounts of features with different finite verbs. There is only one set of SS-PF correspondence rules for inflectional features in a language, which for finite verbs in English consists of the rule in (33). Such rules cannot be sensitive to whether the inflectional node(s) happen(s) to be adjacent to a main verb or to an auxiliary. Hence, once the learner has established (on the basis of data involving finite main verbs) that the rule in (33), which mentions Ad and Num features, is part of the English grammar, he must assume that Ad and Num features are present in syntax in English, as in (40).

17. In those cases where a single main verb does express the same thing as a periphrastic construction, the periphrastic construction is indeed blocked; see Ackema & Neeleman 1998 for data and discussion.

18. If traces are taken into account, there would be one extra violation, resulting in five violations, still not fatal for No-Branching^6.

19. Note that the complex predicate consisting of main verb and auxiliary cannot undergo V-to-I as a whole (instead of excorporation of the auxiliary):

(i) *



This is impossible for two reasons. First, it leads to the main verb carrying both participial and finite inflection, which is morphologically impossible in the modern Germanic languages (see Drijkoningen 1989, Ackema 1998 for discussion). Second, the structure resulting from it in (i) leads to more No-Branching violations than occur in (42) (since Num^0_a dominates two extra pairs of sister nodes), whereas with respect to the other constraints it is no better. This means (i) is "harmonically bound" by (42) (cf. Prince & Smolensky 1993), meaning that it is worse under any possible constraint ranking.

20. I have no account for raising of main verb *be* and main verb *have* in some dialects (cf. Pollock 1989). Also, the analysis predicts that languages could exist in which V-to-I is not triggered with clusters consisting of two verbs but is triggered with clusters consisting of more than two verbs, i.e. it predicts that having V-to-I or not can depend on the number of auxiliaries being present. I do not know if such a language exists. Note, however, that this only occurs under a ranking crucially mentioning a high order No-Branching constraint again (since in such a hypothetical language a two-verb cluster with base-generated inflection on both verbs does not trigger V-to-I yet, No-Head-Mvt must be ranked above a high order No-Branching constraint), which makes it a marked option for the same reason that having rich inflection on an in situ verb is marked, see section 4.

21. Roberts (1998) relates the difference between English and mainland Scandinavian to the latter being V2 languages, which in his account makes V-to-I for auxiliaries superfluous and hence impossible. In fact, given Roberts's account, it seems having V2 would always preclude having independent V-to-I, for any verb. Icelandic, however, is a V2 language that independently has V-to-I (see Vikner 1995).

22. Note that this supplies further evidence that No-Lex-Mvt should be replaced by more general No-Head-Mvt.

23. Various different theories adopting a radical separation between overt affixes and the syntactic representations of the features these affixes correspond to are also proposed in Anderson 1992, Halle & Marantz 1993 and Beard 1995, among others.

References

Ackema, Peter (1997). Case-features and changes in auxiliary selection in unaccusative perfects. Ms. University of Groningen.

Ackema, Peter (1998). *Issues in morphosyntax*. Ms. University of Groningen. To appear at John Benjamins.

Ackema, Peter & Ad Neeleman (1998). Absolute ungrammaticality. To appear in J. Dekkers et al. (eds.), *Optimality theory: phonology, syntax and acquisition*, Oxford University Press.

Ackema, Peter, Ad Neeleman & Fred Weerman (1993). Deriving functional projections. In: A. Schafer (ed.), *Proceedings of NELS 23*, 17-31. Amherst: GLSA.

Anderson, Stephen (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.

Beard, Robert (1995). *Lexeme morpheme base morphology*. New York: SUNY Press.

Belletti, Adriana (1990). *Generalized verb movement*. Torino: Rosenborg and Sellier.

Bickerton, Derek (1984). The Language Bioprogram Hypothesis. *The Behavioral and Brain Sciences* 7, 173-221.

Blevins, James (1995). Syncretism and paradigmatic opposition. *Linguistics and Philosophy* 18, 113-152.

Bobaljik, Jonathan David & Höskuldur Thráinsson (1997). Two heads aren't always better than one. Ms. Harvard/McGill/University of Iceland.

Bošković, Željko & Daiko Takahashi (1998). Scrambling and last resort. *Linguistic Inquiry* 29, 347-366.

Bresnan, Joan & Sam Mchombo (1987). Topic, pronoun and agreement in Chichewa. *Language* 63, 741-782.

Chomsky, Noam (1986). *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Coopmans, Peter (1988). On extraction from adjuncts in VP. In: H. Borer (ed.), *Proceedings of WCCFL 7*, 53-65.

Drijkoningen, Frank (1989). *The syntax of verbal affixation*. Tübingen: Niemeyer.

Emonds, Joseph (1978). The verbal complex V'-V in French. *Linguistic Inquiry* 9, 151-175.

Erb, Christine (1995). *Zur Theorie expletiver Verben*. MA thesis, Johann Wolfgang von Goethe Universität, Frankfurt a.M.

Grimshaw, Jane (1997). Projection, heads, and optimality. *Linguistic Inquiry* 28, 373-422.

Haider, Hubert (1989). Matching projections. In: A. Cardinaletti, G. Cinque & G. Giusti (eds.), *Constituent structure*, 101-122. Dordrecht: Foris.

Haider, Hubert (1993). *Deutsche Syntax generativ*. Tübingen: Günter Narr.

Halle, Morris & Alec Marantz (1993). Distributed morphology and the pieces of inflection. In: K. Hale & S.J. Keyser (eds.), *The view from building 20*, 111-176. Cambridge, Mass.: MIT Press.

Jackendoff, Ray (1997). *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.

Jensen, John (1990). *Morphology*. Amsterdam: John Benjamins.

Johnson, Kyle & Sten Vikner (1994). The position of the verb in Scandinavian infinitives: in V^0 or in C^0 but not in I^0. *Working Papers in Scandinavian Syntax* 53, 61-84.

Jonas, Dianne (1996). *Clause structure and verb syntax in Scandinavian and English.* Ph.D. dissertation, Harvard University.

Kegl, Judy, Annie Senghas & Marie Coppola (1998). Creation through contact: sign language emergence and sign language change in Nicaragua. To appear in: M. DeGraff(ed.), *Creoles, language change and language acquisition*, MIT Press.

Kerstens, Johan (1993). *The syntax of person, number and gender.* Berlin: Mouton de Gruyter.

Koeneman, Olaf (1997). On V to I movement and flexible syntax. In: T. Cambier-Langeveld et al. (eds.), *Proceedings of Console 5.* Den Haag: Holland Academic Graphics.

Koopman, Hilda (1995). On verbs that fail to undergo V-second. *Linguistic Inquiry* 26, 137-163.

Lapointe, Stephen (1980). *A theory of grammatical agreement.* Ph.D. dissertation, University of Massachusetts at Amherst.

Legendre, Géraldine, Paul Smolensky & Colin Wilson (1998). When is less more? In: P. Barbosa, D. Fox, P. Hagstrom, M.J. McGinnis & D. Pesetsky (eds.), *Is the best good enough?*, 249-289. Cambridge, Mass.: MIT Press / MITWPL.

Lieber, Rochelle (1992). *Deconstructing morphology.* Chicago: University of Chicago Press.

Ndayiragije, Juvénal (1996). Case checking and OVS in Kirundi. In: A.-M. Di Sciullo (ed.), *Configurations.* Somerville, Mass.: Cascadilla Press.

Neeleman, Ad (1994). *Complex predicates.* Ph.D. dissertation, Utrecht University.

Neeleman, Ad (1997). Review of Rohrbacher 1994. *Glot International* 2.3.

Noyer, Ralf (1993). Optimal words: towards a declarative theory of word-formation. Ms. Princeton University.

Plank, Frans (1991). Rasmus Rask's dilemma. In: F. Plank (ed.), *Paradigms: The economy of inflection.* Berlin: Mouton de Gruyter.

Platzack, Christer & Anders Holmberg (1989). The role of AGR and finiteness. *Working papers in Scandinavian syntax* 44, 101-117.

Pollock, Jean-Yves (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365-424.

Prince, Alan & Paul Smolensky (1993). *Optimality theory.* Ms. Rutgers University / University of Colorado at Boulder. To appear at MIT Press.

Reuland, Eric (1990). Head movement and the relation between morphology and syntax. In: G. Booij & J. van Marle (eds.), *Yearbook of morphology 1990*, 129-161. Dordrecht: Foris.

Rizzi, Luigi (1982). *Issues in Italian syntax.* Dordrecht: Foris.

Roberts, Ian (1993). *Verbs and diachronic syntax.* Dordrecht: Kluwer.

Roberts, Ian (1997). Restructuring, head movement, and locality. *Linguistic Inquiry* 28, 423-460.

Roberts, Ian (1998). *Have/be* raising, Move F, and Procrastinate. *Linguistic Inquiry* 29, 113-125.

Rohrbacher, Bernard (1994). *The Germanic languages and the full paradigm: a theory of V to I raising*. Ph.D. dissertation, University of Massachusetts at Amherst.

Smolensky, Paul (1995). On the internal structure of *Con*, the constraint component of UG. Handout of talk given at UCLA.

Stump, Gregory (1998). Inflection. In: A. Spencer & A. Zwicky (eds.), *The handbook of morphology*, 13-43.

Tesar, Bruce & Paul Smolensky (1998). Learnability in Optimality Theory. *Linguistic Inquiry* 29, 229-268.

Thráinsson, Höskuldur (1993). On the structure of infinitival complements. *Harvard Working Papers in Linguistics* 3, 181-213.

Tieken-Boon van Ostade, Ingrid (1990). The origin and development of periphrastic auxiliary *do*: a case of destigmatisation. *North-Western European Language Evolution* 16, 3-52.

Vikner, Sten (1995). *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Vikner, Sten (1997a). V^0-to-I^0 movement and inflection for person in all tenses. In: L. Haegeman (ed.), *The new comparative syntax*, 189-213. London: Longman.

Vikner, Sten (1997b). V^0-to-I^0 movement, the person feature, and optimality theory. Handout of talk, Stuttgart University.