Clustering Theories[1]

Jonathan David Bobaljik
McGill University

Fall 2002
minor revisions Feb 2003

This paper aims only to make a few brief remarks about what seem to me to be the more salient aspects of the verb cluster puzzles. I will comment briefly and somewhat selectively on points where the verb cluster data appears to bear on differences, and hence choices, among current syntactic theories. However, most questions will be left open, and I make no attempt at exhaustivity. For a far longer and more general overview of the phenomenon within Germanic see Wurmbrand (forthcoming) in addition to the introduction to this volume.

1. Verb Clusters

Observationally, Dutch and German are verb-final languages. Verb-second (V2) and other movements (such as extraposition) can mask this, but when these are controlled for, the verb occurs at the right periphery of its clause, following its complement(s). In the case of verbs selecting VP or clausal complements, the net effect is that the verbs "stack up" at the right, a fact which Mark Twain (1880) commented upon with great amusement. This is illustrated for German in (1), where the subscript numbers on the verbs illustrate their relative (semantic) hierarchical order (1 is the highest), and the verbs occur in the mirror-image order of their English counterparts in the paraphrase.
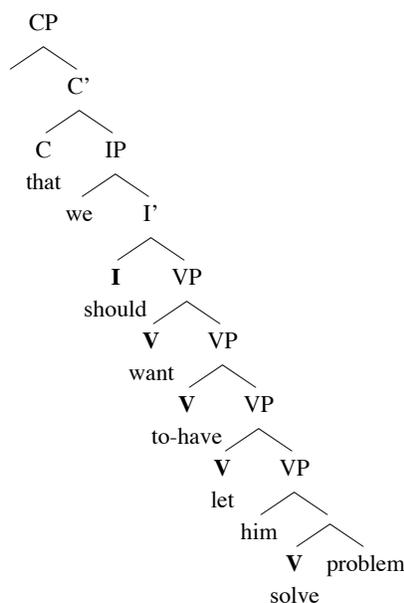
(1) *...dass wir ihn dieses Problem lösen$_5$ lassen$_4$ müssen$_3$ wollen$_2$ sollten$_1$*
 …that we him this problem solve let have-to want should
 'that we should$_1$ want$_2$ to have to$_3$ let$_4$ him solve$_5$ this problem'
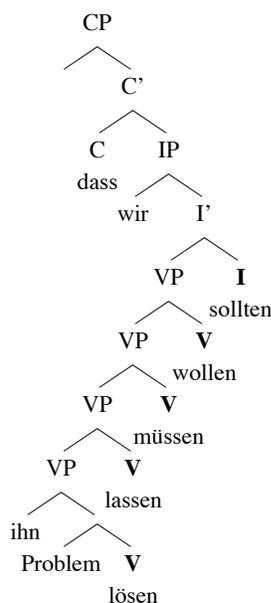
(van Riemsdijk 1998b:7)

Presented with these facts, the most straightforward account would model the German/English difference by positing a single, semantically-determined hierarchical phrase structure, common to the two languages, and supplemented with language-particular precedence rules which reverse the order of heads and complements reversed within the VP (and possibly IP). The German embedded clause in (1) could thus have exactly the same syntactic structure as its English translation, and parametric variation is kept to the minimal assumption of a headedness rule or parameter. This is illustrated in (2).

(2) a. English   b. German

```
         CP                                    CP
        /  \                                  /  \
       C'                                    C'
      /  \                                  /  \
     C    IP                               C    IP
    that                                  dass
        we    I'                              wir    I'
             /  \                                   /  \
            I    VP                               VP    I
          should /  \                            /  \  sollten
               V    VP                          VP    V
             want  /  \                        /  \  wollen
                  V    VP                      VP    V
               to-have /  \                   /  \  müssen
                      V    VP                VP    V
                    let   /  \              /  \  lassen
                        him   /  \        ihn   /  \
                             V  problem  Problem  V
                           solve                lösen
```

This pretty picture is challenged by the properties of Germanic verb clusters discussed throughout the present volume. Dutch, for example, so like German in other relevant respects, has the expected gross word order—verbs canonically follow their non-verbal complements—but unexpectedly reverses the order of the verbs relative to one another.[2] The Dutch equivalent of (1) is given in (3).

(3) *...dat wij hem dit problem zouden₁ willen₂ moeten₃ laten₄ oplossen₅*
     …that we   him this problem should   want   have-to   let   solve
     'that we should want to have to let him solve this problem' (van Riemsdijk 1998b:7)

Unlike the German example in (1) or the English paraphrases, the Dutch example in (3) displays a discontinuous dependency—the object *dit probleem* 'this problem' and its selecting verb do not form a constituent on the surface—all of the higher verbs intervene.

Dutch and German are no strangers of course to discontinuous dependencies on the surface. Scrambling (4a) and topicalization (4b), for example, readily create word order patterns in which the object and its selecting verb do not constitute a surface constituent, and such surface discontinuity forms the basis of the theory of movement and its analogues in other frameworks.

(4) a. *Ohne Hilfe kann dieses Problem nur der Hans lösen.*
         without help   can   this   problem only the Hans solve
         'Only Hans can solve this problem without help.'

---

[2] This way of stating the headedness of Dutch, German, Afrikaans etc. captures the fact that clusters are always right-peripheral (up to extraposition), along with the distribution of cluster-internal non-verbal constituents in Verb Projection Raising constructions (see section 4). Also, if as Wurmbrand (2001b) argues, restructuring complements are VPs, while other infinitive complements are clausal, then this same generalization may also form the basis for an account of why various aspects of cluster formation are limited to restructuring verbs.

b. *Dieses Problem   kann   nur   der   Hans   lösen.*
   this   problem   can   only   the   Hans   solve
   'Only Hans can solve this problem.'

However, the word order variation associated with well-motivated instances of syntactic (phrasal) movement typically has interpretive correlates. For example, to a first approximation, a scrambled object is interpreted as old information whereas an unscrambled object, adjacent to its selecting verb, is interpreted as new information. Once scrambling, V2 and the like are controlled for, the literature offers no systematic interpretive difference that can be attributed to the German versus Dutch difference in (1) and (3), or other examples internal to a single speech variety. In the limiting case, true optionality is reported, as in the case of clusters involving a finite verb and its non-finite complement in Dutch, as in (5).

(5) a. *(Ik denk)*   *dat*   *Jan*   *het*   *boek*   *gelezen*   *heeft*
       (I think)   that   Jan   the   book   read   has

    b. *(Ik denk)*   *dat*   *Jan*   *het*   *boek*   *heeft*   *gelezen*
       (I think)   that   Jan   the   book   has   read

    both:'(I think) that Jan has read the book.   (Zwart 1996: 232)

Even where broadly accepted movement dependencies lack a clear interpretive correlate, as in the movement of the finite main verb to second position in a V2 clause, such instances of movement are typically robust and invariant within a speech community.

The discontinuous verb-object dependencies evidenced in examples such as (3) and (5) thus differ from canonical movement dependencies like those in (4) in that the former lack an effect on interpretation and show pervasive inter- and even intra-speaker variation (see Wurmbrand, this volume). The use of the term *verb cluster* in these cases is thus an extension of other uses of the term *cluster* in the generative literature (see especially Bonet 1991 on Romance *clitic clusters*), referring basically to discontinuous dependencies in which the familiar motivations or diagnostics for movement are not obviously attested, i.e., strings of similar elements in which the surface order departs from what simple rules of precedence will derive from the semantically motivated structure.

How are the verb cluster phenomena to be accommodated within a constrained theory? There are broadly speaking three possibilities instantiated among the contributions to this volume.

First, there are what we might call *inheritance* approaches. These approaches accept that the surface constituency is that indicated by the word order, i.e., that the object *het boek* 'the book' does not form a constituent with the verb *gelezen* 'read' in (5b), and they thus allow for a mediated dependency between the object and the verb. In particular, some mechanism is posited in order to allow a complex constituent [*heeft gelezen*] (importantly, with no traces of the object) to inherit the subcategorization of the non-head, i.e., the participle. In Williams (this volume), this inheritance is referred to as *Geach's Rule* and was first proposed for Germanic verb clusters in Categorial Grammar by Steedman (1985). Approaches within HPSG (Hinrichs & Nakezawa

1994, Kathol 2000) use such a mechanism as well.[3] In such theories, there are no truly discontinuous constituents as the object is not selected by *gelezen* alone but rather by the complex predicate [*heeft gelezen*] with which it is indeed adjacent.

The second family of theories represented in this volume originates with Haegeman and van Riemsdijk (1986). These approaches posit traceless surface constituents isomorphic to those posited by the inheritance theories. For proponents of these theories, though, the constituents are derived. Subcategorization requirements are taken to be met at some early stage of the derivation (such as D-structure) and a subsequent operation alters the constituency. Typical of this family is the positing of an operation of *reanalysis* or *rebracketing*, often restricted to adjacent input constituents (see especially section 2.3, below). In the instantiations of this idea that are most clearly delineated from the other frameworks, rebracketing is held to operate in the mapping from a purely hierarchical representation (the "narrow syntax" of Chomsky & Lasnik 1995) to the phonological string. Within models where semantics interprets narrow syntax only, these theories thus differ from both the inheritance and the antisymmetry models in that cluster formation cannot feed semantic interpretation and thus word order differences in such clusters could not correlate with interpretive.

The final family of accounts of verb cluster phenomena are *movement* approaches originating at least in part with Evers (1975), and now encompassing studies within the Antisymmetry framework of Kayne (1994) (see especially den Dikken 1996 and Zwart 1996). Such approaches share with the reanalysis theories the view that subcategorization requirements are met at D-structure (or something like it) and that the surface constituencies are derived. They part company, though, in (presumably) denying that the mechanisms which derive the discontinuous dependencies on the surface are anything other than the familiar kinds of movement illustrated in (4).

To the extent that this grouping of the various proposals is reasonable, the major issues on which opinions are divided seem to me to be:

the treatment of variations in headedness (primitive or derived)

the place of arbitrary variation in the grammar (represented, but ignored, in the structures that are the input to semantics, or derived in the *Spell-Out* module)

the role of linear relations, including adjacency, in (narrow) syntax

With the landscape mapped out in this way, the following sections present some of my own views on the strengths and challenges for the various families of approaches.

2. Headedness and Constituency

2.1 Extended headedness

The inheritance and the reanalysis approaches taken together share an acceptance of primitive headedness / directionality parameters, though individual accounts vary in the manner in which

---

[3]Williams's theory (this volume) is clearly related to the CG and HPSG approaches, but differs in being embedded in a framework that does accept movement for the canonical cases such as (4), see below.

these are represented. In particular, these approaches are distinguished from the movement approaches in allowing for the possibility that two languages or constructions might (at least in principle) vary *only* in precedence relations, with no concomitant difference in constituency relations. The clear case of this is headedness, as in (2) up to Saussurean arbitrariness in the lexical items.

Cluster-internal variation can be assimilated to headedness variation on the assumption that the verb cluster, i.e., the sequence of verbs at the right edge in (3), forms a constituent to the exclusion of the nominal arguments. We can call this constituent V* to avoid any commitments about projection level (X° vs XP). Once such a constituent is derived at some level of representation, it may serve as the input to language / construction specific headedness rules. For example, the German versus Dutch contrast could then be modeled as in (6). In these trees, V* is final within VP in both languages, but internal to V* German is head-final and Dutch is head-initial.[4]
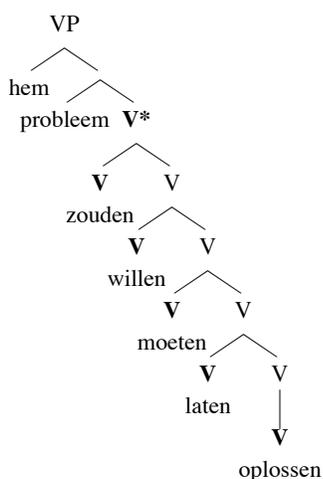
(6) a. Dutch   b. German
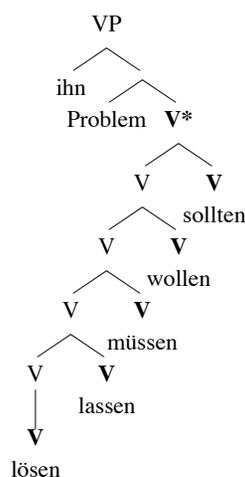
Many other attested orders (though not all—see below) can be cast as involving mixed headedness within V* and the Verb Projection Raising constructions in Flemish and Swiss are readily accommodated by allowing V* to include phrasal projections and thus internal complements. There is an empirical question as to the limit of variation that is allowed by UG. Thus Wurmbrand (this volume) suggests that different constructions or classes of verbs (modals, auxiliaries, participles) may vary in headedness, while the approach in Kathol 2000 allows for individual verbs to bear lexical specifications for the headedness of the constituent in which they occur. Another difference between approaches lies in which verb is responsible for the headedness of the construction in which it occurs; for Kathol, for example, it is the non-head, i.e., selected, verb which is specified as obligatorily preceding or following its selecting head, whereas for Williams, it is the head that is specified as preceding or following its complement.

---

[4] It is not necessary to posit a cluster or V* constituent for the German order here, as that is derived with a simple head-final VP as shown in (2). The cluster is posited here for expository convenience as it makes for the most minimal variation with the Dutch example selected. The choice is not innocuous in the bigger picture and I return to this below.

One might (and I will) call these families of approaches the *extended headedness* approaches since they preserve an important role for headedness as a point of minimal variation between languages or constructions.

2.2 Inheritance versus reanalysis

The extended headedness approaches differ amongst themselves in the claims that are made about how the special V* constituents in examples like (6) are established. Under the inheritance approaches, (originating, I believe, with Steedman 1985), the complex constituents like V* are base-generated as such and thus the difference between, English and Dutch is not only one of precedence but also one of syntactic constituency. Under the inheritance approaches, there is no level of representation at which the Dutch object *het boek* in (5b) ever forms an English-like VP constituent with the verb *gelezen* excluding the higher auxiliary. On the reanalysis approaches (originating with Haegeman and van Riemsdijk 1986) the complex predicate constituents are derived from more English/German-like VPs. the derivation of (5b) within the reanalysis framework is sketched in (7) (see Wurmbrand, forthcoming, for a far more detailed presentation of the various theories that have been proposed).

(7) a. $[_{VP1} [_{VP2}$ *het boek gelezen$_2$* $]$ *heeft$_1$* $]$ D-structure constituency

  b. $[_{VP1}$ *het boek* **[*gelezen$_2$ heeft$_1$*]** $]$ Reanalysis / Rebracketing

  c. $[_{VP1}$ *het boek* **[*heeft$_1$ gelezen$_2$*]** $]$ Flip / Inversion = Headedness

The derivation begins with a base structure (7a) in which the subcategorization restrictions are satisfied in the semantically motivated constituents for example, in which the object is the sister to its selecting verb. A transformation of reanalysis or rebracketing (7b) alters the constituency, though not the precedence relations among the terminal elements. Finally, the effect of surface discontinuous constituency is achieved by applying inversion rules within the derived constituents, as in (7c).

Although the two approaches converge on the surface constituency, it would seem that the presence versus absence of the structure in (7a) should have the potential to distinguish among the theories. On might, for example, expect to find interpretive correlates when scope bearing elements are involved. Consider, for example the scope interactions between a quantified object and an inherently negative verb as in (8).

(8) a. … (that) Ivan [forgot [to bring all the books]].
  b. … (that) Ivan [forgot all the books].

The example in (8a) admits of a reading in which the inherent negation in the verb *forget* scopes over the universally quantified object (i.e., Ivan remembered to bring some books, but not all), consistent with the surface c-command relations indicated. (The example is also consistent with a reading in which the universal scopes over negation though it is not trivial to show that this is a true ambiguity as opposed to vagueness.) The example in (8b) lacks such a reading; this sentence is infelicitous in a situation where Ivan remembered some of the books, but not all. Imagine, for the sake of argument, that the distinction has to do with constituency, say, that the "not all"

reading requires the verb *forget* to asymmetrically c-command the universally quantified DP in object position, which it does in (8a) but not in (8b). Making the appropriate lexical substitutions into a cluster example like Dutch (5b), the inheritance theory differs from the reanalysis theory precisely in whether or not there is a constituency meeting this requirement at any level of representation. In the reanalysis theory, the surface constituency in (9b) is derived from (9a), while in the inheritance theories (9b) is the only structure posited.

(9) a. … (that) Ivan [[all the books to bring] forgot] before reanalysis & inversion
   b. … (that) Ivan [all the books [forgot to bring]] inheritance

The presentation just above is intended to be illustrative only, the assumptions about scope are probably too naïve to be useful as stated. However, it is clear that scope-bearing elements do enter into clusters, and the theories do make different assumptions about basic constituent structures, from which it should in principle be possible to derive predictions about scope. As far as I know, the facts have not been tested.[5]

  The differences between the two classes of theories are particularly clear—at least in principle—if they are embedded in the Y or T model of grammatical architecture proposed in Chomsky & Lasnik (1977). The salient property of this model is the assumption that the syntax mediates the relation between sound and meaning at the phrasal level by deriving hierarchical structures that are interpreted by both the semantic and phonological components. A derivation is conceived of as containing a point (recently called "Spell-Out") at which the information is split and subsequent operations selectively manipulate either the semantic (e.g., QR) or phonological structure. Because of this architecture, differences in syntax before Spell Out (among languages or constructions) constitute differences in the structure that serves as the input to semantics and thus at least in principle could be expected to correlate with differences in semantic interpretation. For example, scrambling, considered above, is a syntactic operation happening before Spell-Out and therefore it feeds both word order (phonological form) and interpretation.

  One particularly strict version of this general framework holds that syntactic structure consists solely of hierarchical information (dominance relations, constituency) with precedence relations established in the mapping of the syntactic structure to the phonological string (see Marantz 1984: 7-8). An intuition shared by a number of proponents of these frameworks is that truly arbitrary, language-particular variation in precedence relations, such as differences in head-complement order (i.e., headedness) which are not known to systematically correlate with semantic interpretation, are therefore best taken to be the domain of what we might call *Spell-Out* rules, i.e., part of this mapping from syntactic structure to linear string. From this perspective, one might ask whether the reanalysis / rebracketing process is properly viewed as a syntactic or a post-syntactic operation. Such a question makes sense only when embedded in a theory that makes the distinction independently, and could be explored by asking whether the triggers and restrictions on reanlaysis (as diagnosed by cluster formation) are syntactic (referring

---

[5] Den Dikken (1996) has argued that the positions of arguments in Verb Projection Raising constructions in West Flemish have an influence on their scope-taking possibilities. What is not shown, but is directly relevant to the question here, is whether the linear order of the verbs alone influences the constituency of the clause, in a manner that may have implications for scope possibilities. Note also that in German, there is an effect of forced wide scope in unambiguously lexical restructuring contexts such as long passive (see Wurmbrand 2001a on which the above examples are modeled) but this effect appears to be tied to case assignment / agreement properties and not to surface constituency / linear order as the inheritance theories would predict (see also Bobaljik & Wurmbrand 2002 for discussion).

to meaning and hierarchy) or morphological/phonological (referring to notions of precedence alone). Though the full question is perhaps premature, Susi Wurmbrand presented some initial arguments in the course of the Working Group meetings that the cluster formation is dependent upon the morphological and not the syntactic category of the elements involved (see also Wurmbrand 2003). The IPP construction (*Ersatzinfinitiv*) is characterized by the use of a (morphological) infinitive form of the verb where a participle would be expected on general grounds (e.g., complement to auxiliary *have*). In terms of construction-specific ordering, this infinitive patterns with other (true) infinitives and not with its syntactic kin, the other participles. This is the kind of argument that would weigh in favour of a particular interpretation of the reanalysis view, especially as against views that take cluster formation to be (narrowly) syntactic.

2.3 Why OV?

Another point regarding headedness which comes up in the discussion of verb clusters, and which remains open so far as I can see, lies in the distribution of the phenomenon within Germanic. In particular, leaving aside Hungarian (and a quibble about Yiddish) to which we return, verb cluster formation is restricted to the OV Germanic languages. In the VO languages (English and the Scandinavian languages at least), departures from the order which directly reflects the syntactic hierarchy (e.g., (2a)) are limited to instances of movement to the clause initial position (VP-fronting, Stylistic Fronting in Icelandic and Old Swedish) or extraposition of a VP or larger constituent.

Inheritance-based theories such as Williams (this volume) and its antecedents in CG and HPSG seem ill-equipped to capture this restriction in a principled manner, where the reanalysis based theories may fare better.

To see the difference, consider first the analysis of Dutch (10a) offered in Steedman (1985) (example attributed to Huybregts 1976).

(10) a. *... omdat   ik Cecilia   de nijlpaarden   [zag   voeren]*
          because   I Cecilia   the hippos   saw   feed
          '…because I saw Cecilia feed the hippos.'

      b. …[because [I [saw [Cecilia [feed the hippos]]]]]

The non-finite verb *voeren* 'feed' is a normal transitive verb, and as such it subcategorizes for two NPs (first the object, then the subject). The result is what we might think of as a non-finite small clause: [SC NP [NP V]]. The verb *zag* 'saw' subcategorizes for such a small clause, and then for an external NP argument (the seer). These combinatorial properties are transparent in the English right-branching structure in (10b). In Dutch, however, Steedman invokes *Forward Partial Combination*, effectively the same as Williams's *Reassociate* to allow two things to happen. First, *zag* and *voeren* combine to yield the complex constituent as shown in (10a), and second, by Geach's Rule, this complex constituent inherits the internal subcategorization of its non-head. In effect, this yields a (complex) verb with the subcategorization of a ditransitive verb (like *give*): it combines with two internal arguments to yield a finite VP that subsequently combines with the subject. Note that the complex verb inherits the precedence requirements of the non-head as well and thus the complex [*zag voeren*] 'saw feed' follows all of its arguments.

Nothing in these accounts prohibits the same operation from applying in a VO language like English or the Scandinavian languages, however. The mechanisms invoked to account for Dutch (10a) also permit either (11a) or (11b) in English. The combination *saw-feed* should, in effect, have the same distribution as a normal ditransitive verb up to issues of precedence.[6]

(11) a. *… because I [saw-feed] Cecilia the hippos.
    b. *… because I Cecilia [saw-feed] the hippos.

An appeal to a lexical accident is implausible since constructions like (10a) appear to be systematically absent from the Germanic VO languages.

Note that what is in essence the same point can be made with respect to Williams's *Flip* in the special constituents. None of the inheritance (or movement-based, see below) theories currently able to capture the range of facts in West Germanic can exclude (12a) (or its word-for-word translation in any of the Scandinavian languages) for any principled reason. The derivation to be excluded is sketched in (12b-c), compare the derivations of Dutch on the extended headedness theories in (7). (Innumerable derivations are possible under Antisymmetry theory, and must be excluded by language-particular stipulations.)

(12) a. * He must bought have the book.

    b. $[_{IP}$ He must$_1$ $[_{VP}$ have$_2$ $[_{VP}$ bought$_3$ the book $]]]$
              ⇓

    c. $[_{IP}$ He must$_1$ $[_{VP}$ $[_{V°}$ bought$_3$ have$_2$] the book $]]$ by Reassociate / Head-Movement

An answer to the question of why this is restricted in this manner is one of the prize puzzles of studies of Germanic word order. The current best guess would appear to try to relate this to the intuition, shared among all the frameworks discussed above, that the special mechanisms that account for the ordering in clusters are operative in a "small" domain, a complex X° in the trees in (6), something slightly larger in Swiss German and West Flemish varieties that allow for Verb Projection Raising. The trick, then, will lie in restricting the formation of the appropriate constituents to the OV languages.

Haegeman and van Riemsdijk (1986) and van Riemsdijk (1998a) have suggested that the key factor lies in string adjacency among the affected elements. As van Riemsdijk (1998a) notes, string adjacency alone is insufficient; the verbs in English (12) are adjacent, yet cluster formation is impossible. Note, though, that the VO languages and the OV languages (within Germanic at least) are independently known to differ as to their tolerance of right-adjoined adverbs. The right edge of a VP or clausal projection appears to be an impossible site for base-generated adjunction in the OV languages, where such adjunction (along with left-edge adjunction) is freely available in the VO languages. It seems that there is a hope then of distinguishing between the OV languages, in which the consecutive verbs are *necessarily* adjacent (up to extraposition), and the

---

[6] The order in (11b) is expected if, as in Steedman's framework, precedence relations are part of subcategorization. In inheriting the subcategorization of *feed*, the complex *saw-feed* should combine with an object to its right and a subject to its left. If, as in Williams's theory, subcategorization is for type, but not for precedence, then (11a) might be expected as it displays the canonical word order for a predicate that combines with three NP arguments (cf. *because I showed Cecilia the hippos*).

VO languages, in which the verbs may be accidentally adjacent in any given sentence (e.g., (12)), but may in principle be separated by base-generated adjuncts (or perhaps specifier positions). Only the former permit cluster formation.[7]

The appeal to necessary adjacency idea strikes me as a promising direction, though at the moment it is in part a speculation regarding the reduction of one unsolved problem to another, namely why right-adjunction should correlate with headedness, where left-adjunction does not (though see Saito & Fukui 1998 for related thoughts). Nevertheless, if the appeal to adjacency will ultimately prove to be the correct account of the VO / OV asymmetry, it is hard to see this as anything other than an argument for an account that invokes a level of syntactic representation distinct from the surface constituent structure in verb clusters. The reason is that the adjacency relation must be computed over a structure that is distinct from the surface structure. The verbs in (12c) are adjacent to one another and the structure is consistent with all properties of Williams's CAT or any other inheritance theory, so far as I can see. In other words, adjacency must be computed before reassociation applies, but on the inheritance theories, the reassociated structures are the base-generated structures—there is nothing before them over which the adjacency condition can be stated.

2.4 Limits of extended headedness.

Under the extended headedness approaches, the verb cluster ordering phenomena are the analogue of general headedness variation, but within the special constituents noted as V* in (6). The restrictions on the order become, in essence, restrictions on what the constituents are that may feed the headedness rules. UG sets the possible bounds of variation; language and construction specific variation is then essentially arbitrary within these bounds. For example, Wurmbrand (forthcoming) carefully outlines the range of possible word orders that can be characterized by applying headedness rules (= Williams's *Flip*) to a three-member cluster. Since headedness maintains the integrity of constituents, it is impossible using only this mechanism to

---

[7] Yiddish is conspicuously exceptional in this regard, in that it is quite generally a VO language but allows a cluster-like complement-verb order with passive participles, in which the participle may precede the passive auxiliary, as illustrated in (i) from den Besten & Moed van Walraven (1986: 117) (with minor modifications following Diesing 1997: 386).

(i)	a.	*Matones*	*zaynen*	*gevorn*	*gebrakht*
		gifts	are	been	brought
	b.	*Matones*	*zaynen*	*gebrakht*	*gevorn*
		gifts	are	brought	been

both: 'Gifts have been brought.'

Even controlling for leftward movement of the object (scrambling, object shift), other examples of OV-like orders in the verbal domain in Yiddish include certain adverb-copula sequences and verb-particle combinations. These are discussed by Diesing (1997), who argues for movement analyses in all cases, in particular, arguing that the example in (ib) involves incorporation of the participle into the auxiliary. An approach along these lines is necessary if it is indeed true that true verb cluster effects are limited to head-final languages. Verb cluster orders with passive participles are special in Afrikaans as well, which allows the order 3-1-2 only when the "3" element is a passive participle (see Wurmbrand, this volume, note 1 to Table 2). As noted in (13b) below, this order is potentially problematic within some approaches.

derive orders in which underlying constituents are discontinuous on the surface, i.e., without reanalysis or reassoication. Thus, if reanalysis may feed cluster formation, then cluster-internal headedness rules expand the range of possible orders, but not without limit. From a three-member cluster [1 [2 3]], only four orders are possible:

(13) a.   possible: 1 2 3; 1 3 2; 3 2 1; 2 3 1
      b.   impossible: *3 1 2; *2 1 3

On the surface, the excluded orders do occur (see (17), below), and these must then receive a special treatment involving additional operations. In keeping with the leading intuition of the framework, it is expected that such operations should be detectable on independent grounds. For example, the surface order 2-1-3 might be expected to be the result of extraposition of the lowest VP, and thus would be restricted to contexts independently known to allow extraposition (this much appears to be correct; Wurmbrand, this volume). Williams's proposals in this volume are also clearly within the extended headedness family, though in addition to permitting *flip* to derive the orders in (13a), Williams admits of the additional mechanism of *reassociation* which allows the orders in (13b) as well. As Williams discusses, while under his theory UG sets no bounds on the variation in order in small clusters (less than three elements), it nevertheless imposes progressively more severe restrictions on larger clusters.

The extended headedness theories, at their core, share the intuition that dominance relations do not universally uniquely determine precedence relations, and thus, that two languages may (in the limiting case) differ only in precedence relations, with no corresponding syntactic differences.

3. Movement and antisymmetry

The extended headedness family of approaches, considered above, may be contrasted at least in part with analyses that involve syntactic movement to account for the surface discontinuous constituents. The most important contrast is with the accounts that subscribe to the framework of Kayne's (1994) *Antisymmetry* theory. These are important in that they deny the central tenet of the extended headedness approaches, namely, that headedness alone can be a point of possible cross-linguistic variation (at least in syntax). In Kayne's theory, all structures are linearized as Spec > Head > Complement, and thus all variation in linear order must involve variation in syntactic structure. Verb clusters, one might think, would provide a fertile testing ground for this divide between theories. As Wurmbrand (forthcoming) has shown, though, verb cluster phenomena turn out to provide no argument for the antisymmetry theory. Wurmbrand provides an extensive discussion of the various types of derivations that are proposed within the antisymmetry framework and a direct comparison with what needs to be said on a framework admitting of headedness parameters. I see no reason to repeat her discussion here, but I will reiterate and expand on some of the salient points.

Before proceeding to the discussion of the antisymmetry account, it is worth noting that movement accounts are not all committed to the antisymmetry approach, as the organization of this chapter might inadvertently suggest. Indeed, the first generative account of verb clusters in Dutch is a movement account, that of Evers 1975. Evers proposed a rule of Verb Raising that transformed an underlying left-branching constituent (a) into one with a complex verbal

constituent as in (b), via raising of the infinitive verb and (right) adjunction to its governor. (I have presented the trees with more current labels.)
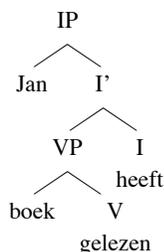
(14) a. Underlying   b.   Surface
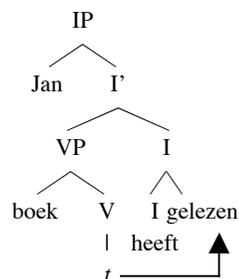
In terms of the constituent structures and the derivation relating them, Evers's account is the clear precursor to the reanalysis account in (7). It shares with that account those properties that are related to the base structure (such as possible interpretive correlations). If adjacency is relevant, and if reanalysis can be more plausibly linked to adjacency constraints than verb raising, then the inversion account may be an improvement on that point. In any event, the differences between a head-movement account (recognizing the possibility of head-final structures and of right-adjunction) is quite similar to the headedness accounts above, in particular, since it admits of directionality parameters as points of minimal variation among language.

An alternative school of thought is centered around Kayne's (1994) *Antisymmetry* theory. The leading intuition of this perspective is the hypothesis that all word order variation is syntactic in the narrow sense of implicating movement. The linearization of any syntactic structure is uniquely determined by UG; thus, two phonological strings *AB* and *BA* cannot correspond to identical syntactic structures. As a consequence, any difference in precedence relations must correspond to a difference in syntactic structure.[8] For the purposes to be discussed below, Bródy's *Mirror Theory* (this volume) is also a theory of this type, once we grant a distinction between morphological and syntactic objects, where the latter universally instantiate head-complement order up to movement and position of pronunciation, and the former universally instantiate head-final order.[9]

Verb clusters in Germanic would seem to provide a fertile ground for comparing these theoretical approaches. Here, as noted above, one finds exceptionally fine-grained differences among dialects, constructions and even among speakers of the same dialect. In particular, when

---

[8] There is at least one exception, namely, the possibility that the trace of a moved element is a copy of that element, and variation in linear order may arise from differences in the choice of copy pronounced. Kayne (1994: 96) endorses this as consistent with the framework, see also Bobaljik (2002) and references therein for arguments in favour of this approach in a model not reliant on the LCA. In order to set out the more interesting claim, I will disregard this possibility in the discussion here.

[9] Bródy introduces the term "morphological specifier" for the non-head member of a binary-branching morphological structure, for example, in an incorporation structure [N V] where V is the head, Bródy refers to the N as the "specifier".

the entire range of data is considered, the situation appears to be strikingly consistent with the extended headedness position.

One property of extended headedness theories that is occasionally commented upon is that the language (or construction) specific headedness rules are arbitrary. There are two related points that strike one in connection with this comment.

The first point is that, so far as the data have been presented to date, this seems to be an accurate characterization of the facts, especially within Germanic. That is, the orders evidenced in any particular speech variety are not predictable on the basis of any other known property of those varieties. There are a handful of implicational relations, but the fact that, for example, some Swiss German speakers allow (15) while others do not has not (to my knowledge) been shown to correlate with any other property.

(15) % *wil   er   si   gsee₃ ha₂   mues₁*
      since   he   her   seen   have   must
      '…since he must have seen her.'   (Wurmbrand, forthcoming: 2)

If this is the correct characterization of the facts, then the distribution of the ordering restrictions among languages is in point of fact arbitrary. It must therefore be learned, directly, and any theory which allows for the description of all the patterns must incorporate arbitrary statements somewhere in the description of any particular speech variety.

To criticize the extended headedness theories on the grounds of arbitrariness would require showing that these theories miss empirical generalizations. Where the headedness theories hold that there may generally be a syntactic symmetry between head-initial and head-final orders, some of the proposals incorporating Antisymmetry invoke a series of movements for one order that are not invoked for the other. For example, in the Koopman & Szabolcsi (2000) system, the head-complement (1>2) order involves pied-piping of CP, which is not involved in the inverted (or 'roll up') order (2>1). A theory like that of Koopman & Szabolcsi's would be non-arbitrary if independent evidence in support of this syntactic difference among the languages were presented. For this particular case, though, as Williams (this volume) discusses, Koopman & Szabolcsi resort to a number of language-particular stipulations which, importantly, are specific to the constructions at hand. Arbitrariness is thus not at issue; Koopman & Szabolcsi (2000) have merely traded one set of language-particular stipulations (including headedness) for another (e.g., the *Hungarian Complexity Filter*, 61).

If it is the case that the language variation is itself arbitrary, we may still ask the question of how (or how well) the various theoretical approaches succeed in encoding the arbitrariness. What I have called the extended headedness theories above simply encode the observation directly with learned rules of precedence (= learned values for headedness parameters). Surely there can be no objection to these theories on learnability grounds; compare left- versus right-prominence in metrical systems, generally taken to be an irreducible point of language variation as noted above. All current alternative theories, based on the Antisymmetry proposals, retain headedness parameters in some guise, as discussed in the preceding paragraph. Whatever mechanism is invoked to derive a V-final structure from a V-medial/initial structure (or vice versa), the mechanism is ultimately attributed to some set of language particular properties that must be learned. It is in my opinion a weakness of these approaches that, in the absence of surprising correlations or new generalizations, the essentially straightforward point of language variation evidenced by headedness tends to be obscured by a series of curiously unrelated

stipulations. Unless it can be shown that the verb cluster patterns correlate with other properties in interesting ways—that is, in ways that the Antisymmetry-based proposals can shed light on—the very arbitrariness of the data implies that verb cluster phenomena in Germanic cannot provide an argument for Antisymmetry theory (see also the remarks on Zwart 1996 in Wurmbrand, forthcoming).

Note also that the Antisymmetry-based theories share with the inheritance based theories the questions concerning the cross-linguistic distribution of verb clusters. Deriving the correlation between verb clusters and (surface) verb-final order would be feather in any theory's cap. If the connection to (surface) right-edge adverbs, and the related adjacency condition, is on the right track, once again the problem is at least far easier to state in theories admitting of headedness than in those adopting Antisymmetry (see in this connection Williams's comments on the problems in domain restriction faced by Koopman & Szabolcsi 2000).

Approaches to verb clusters within the Antisymmetry frameworks assume that the arguments for the framework are established on the basis of independent considerations. On such theories, the straightforward analysis of the data available on the extended headedness approaches is excluded a fortiori. The puzzle, then, lies in working through what assumptions about variation in movement parameters must be made in order to meet the burden of descriptive adequacy set by the data. In the best case scenario, the hope would be that the formalization of the observations will itself yield new and un-thought-of predictions that may turn out to be correct.

4. The Hungarian Connection

Verbal complementation yields variation in surface word order in Hungarian as well as in the West Germanic languages. One question addressed in some of the contributions to this volume is whether this should be considered the same phenomenon as the West Germanic clustering or not, and if so, to what degree. When we regard the verbs alone, there are some striking differences among these languages. On the working definition of the term "cluster" suggested in section 1, the orders of Hungarian verbs are not true clusters. Although there are various possibilities for the surface ordering of a series of infinitive-embedding verbs, there is essentially no arbitrary dialectal variation (see Szendrői and Tóth, this volume), and the possible orders conform quite closely to what is expected on the basis of a right-branching syntax, supplemented with successive-cyclic head movement (assuming, with Kayne 1994, that this always yields left adjunction). In particular, the derivation in (12b) which is impossible in all the VO Germanic languages looks suspiciously like what best characterizes the Hungarian Roll-Up construction. This is illustrated in (16b) straightforwardly derivable from (16a) via successive cyclic head movement (or its analogue in Mirror Theory, Bródy, this volume).

(16) a. *Utálok₁*   *kezdeni₂*   *járni₃*   *úszni₄*
　　　Hate.1SG   begin.INF   go.INF   swim.INF
　　　'I hate to begin to go swimming (regularly)'

　　b. *Utálok₁*   *úszni₄*   *járni₃*   *kezdeni₂*
　　　Hate.1SG   swim.INF   go.INF   begin.INF
　　　--same   (Bródy, this volume)

14

Note in this connection that the right-branching sequences such as (16a) appear to be truly phrasal, and may have phrasal elements (such as a subject DP) interspersed amongst them, while this is impossible amongst the verbs in the roll-up sequence in (16b) (see Ackema, this volume: 16, 20-21). This is as expected on the head movement and Mirror Theory approaches (but see below).

The various conditions on the roll-up and partial roll-up orders described in the papers here (see especially Szendrői and Tóth, this volume) are easily stated as conditions on this head-movement: it must be successive cyclic, it must start with the lowermost verb, the finite verb cannot be included in the roll-up etc. This view would suggest that Hungarian verbal complexes are formed by successive-cyclic head-movement from a head-initial base. Approaches differ in the degree to which they are able to make such conditions follow from deeper principles, or whether they need simply be stipulated add-ons to the theory. The difference between Hungarian and the VO Germanic languages would amount to one of head movement, a point on which the VO languages differ amongst themselves as well (for example English lacks V2, and there are differences among the Scandinavian languages regarding verb movement to Infl in non-V2 environments, see Vikner 1995).

The view that 'roll-up' in Hungarian is not the same process as cluster-formation in the OV Germanic languages allows for the exploration of the generalization suggested in section 2.3, namely, the restriction of the special cluster-forming mechanisms to *base-adjacent* contexts, and thus accounting for their distribution within Germanic. Alternatively—unless Ackema (this volume) is correct in positing an underlying OV structure for Hungarian—assimilating Hungarian to West Germanic jeopardizes this possible account from the outset.

Williams's contribution to this volume provides an interesting means of making this distinction between the languages sharper. Though Williams applies the *CAT* language to both Germanic and Hungarian, there is a striking difference between the two sets of accounts. Specifically, none of the Hungarian orders require the use of *Reassociate*, while all of the Germanic orders that depart from the strictly ascending order of German in (1) require this. For Williams's theory, as a matter of principle, this difference is simply the result of an arbitrary choice of which operations are obligatory or blocked in which contexts, no different from the arbitrary choices among various West Germanic varieties as to where *Flip* is and is not implicated in the definition. It strikes me that treating *Reassociate* and *Flip* in the same manner this way likewise misses the generalization about adjacency along the lines suggested in the preceding section. It is precisely where arbitrary variation is permitted by the theory (headedness, and reassociation under base-adjacency) that seemingly arbitrary variation abounds in the data (OV Germanic); Hungarian, like the VO Germanic languages, is characterized by a striking lack of such variation (see Szendrői and Tóth, this volume, 1), surely a fact we should like the theories to accommodate.

4.1 Preverbs and Particles

Notwithstanding the differences just discussed between Hungarian and West Germanic, there are various points of similarity between the Hungarian constructions and the Germanic counterparts which have suggested to various authors that a unified account is in order (see Ackema, this volume, and Csirmaz, this volume, for discussion of the similarities and differences between Hungarian and Dutch). These emerge particularly strikingly when one investigates particle/pre-

verb constructions and other instances where more than just verbs are included in the apparent clusters (e.g., VPR).

One observation is that the pre-verb / particle appears to participate in the cluster formation in both Hungarian and Dutch, however, it does so in rather different ways. A typical Dutch example of a particle verb *opbellen* 'up-call' embedded in a verb cluster is given in (17). (The verb here is in its participial form.)

(17)   *dat*   *hij*   *haar*   *{op}*   *kan$_1$*   *{op}*   *hebben$_2$* *{op}* *gebeld$_3$*
      that   he   her   (up)   can   (up)   have   (up)   called
      'that he may have called her'   (Ackema, this volume)

In Hungarian, though there is some disagreement, Szendrői and Tóth, this volume report that a significant majority of speakers allow the pre-verb only immediately preceding its selecting verb or in the clause-initial focus position (for semantically transparent pre-verbs, at least), but not freely interspersed within the cluster. This is partially illustrated in the contrast between (18a), the base sentence, acceptable or marginal for a majority of consultants, versus b-c, unacceptable to most consultants.

(18)   a.   *Már*   *egy*   *órája*   *ide*   *értek*   *a*   *t zoltók*   *és...*
           already   an   hour   here-to   arrived   the   firemen   and...

          $^\%$*...csak*   *most fognak*   *tudni*   *próbálni*   **be**   *mászni*   *a*   *barlangba.*
          ...only   now   will.they   can.to   try.to   into.PV   climb.to   the   cave.into

          'It's been an hour since the firemen arrived and it is only NOW that they will be able to try to climb into the cave.'   I16

     b.   *Már*   *egy*   *órája*   *ide*   *értek*   *a*   *t zoltók*   *és...*
           already   an   hour   here-to   arrived   the   firemen   and...

          *...csak   *most fognak*   **be**   *tudni*   *próbálni*   *mászni*   *a*   *barlangba.*
          ...only   now   will.they   into.PV   can.to   try.to   climb.to   the   cave.into

          'It's been an hour since the firemen arrived and it is only NOW that they will be able to try to climb into the cave.'   I10

     c.   *Már*   *egy*   *órája*   *ide*   *értek*   *a*   *t zoltók*   *és...*
           already   an   hour   here-to   arrived   the   firemen   and...

          *...csak   *most fognak*   *tudni*   **be**   *próbálni*   *mászni*   *a*   *barlangba.*
          ...only   now   will.they   can.to   into.PV   try.to   climb.to   the   cave.into

          'It's been an hour since the firemen arrived and it is only NOW that they will be able to try to climb into the cave.'   I1

Indeed, the impression which I take away from the articles collected here is that there is a shared optimism that the positioning of the particle in the Hungarian verb cluster is indeed derivable from deeper considerations, either of prosody (Csirmaz, Szendrői, this volume) or aspectual notions (Alberti, this volume) or perhaps both. Indeed in this sense the particle (or Verbal Modifier) placement in Hungarian seems to lack the arbitrariness (i.e., irreducibility to other linguistic factors) which characterizes many of the verb cluster effects in the West Germanic languages.

There is one shared aspect of the distribution of non-verbal material in the clusters, though, and this may be of interest in choosing among competing theories. Thus, while it is true that the particle *op* may occur interspersed throughout the construction in (17) (as may certain objects in West Flemish and Swiss German VPR constructions), there is nevertheless an inviolable condition that the particle (or light object) must precede its selecting verb. Thus, while the participle may precede the other two verbs (the 3-1-2 order), if it does so, the positions available to the particle are suddenly restricted.

(19) *dat hij haar {op} gebeld₃ {*op} kan₁ {*op} hebben₂ {*op}*
that he her (up) called (up) can (up) have (up)
'that he may have called her' (Ackema, this volume)

Parallel restrictions arise with objects in those varieties that allow the Verb Projection Raising construction. All of these patterns are consistent with the descriptive observation in section 1; the particle (or object) being a non-verbal element, the acceptable and unacceptable patterns in (17)-(19) (and their analogues in VPR) are distinguished by whether or not they follow their non-verbal complements and other associated expressions. This is illustrated with an object and an adverb in Swiss and German respectively in (20) and (21) (Wurmbrand, forthcoming: 23 and references therein).

(20) a. *...ob si hett₁ **d Prüeffig** besto₃ chöne₂*
… whether she had the exam pass can
'[who knows] whether she could have passed the exam.'

b. *...*ob si hett₁ besto₃ **d Prüeffig** chöne₂*
… whether she had pass the exam can
'[who knows] whether she could have passed the exam.'

(21) a. *...daß er das Buch hätte₁ **genau** durchsehen₃ sollen₂*
… that he the book had carefully through-look shall
'…that he should have looked through the book carefully.'

b. *...*daß er das Buch hätte₁ durchsehen₃ **genau** sollen₂*
… that he the book had through-look carefully shall
'…that he should have looked through the book carefully.'

The generalization that a non-verbal element in Germanic must precede (though not necessarily immediately) its associated (e.g., selecting) verb amounts to something very close to the generalization stated regarding the Hungarian constructions in (16). In both Hungarian and West

Germanic, the left-branching order (3>2>1) the sequence cannot be interlaced with XPs, while in the right-branching order (1>2>3) the sequence of verbs may be interrupted by XPs. In a mixed sequence, such as 1>3>2 illustrated above, the left-branching (i.e., right-headed) portion is inviolable (20b), (21b) while the right-branching (i.e., left-headed) part of the cluster (1>[3>2]) allows for intervening XPs, as expected.

Note that this broader generalization emerges from the separate accounts given for each of the constructions, though not for any particularly principled reason. For Hungarian, this situation emerges for the straightforward reasons described under (16)—the left-branching order (16b) might be taken to be the result of head-movement, or the morphological mirror of the syntactic complementation structure in (16a). Either way, only the left-branching structure is expected to be inviolable, as the right-branching structure has an XP node between each verb, a suitable target for XP-adjunction.

Why this situation emerges in cluster accounts of the West Germanic data lies in the generalization that reordering only applies to verbal elements. To accommodate the Dutch particle facts, and VPR generally, the extended headedness approaches all admit verbal projections in the construction of the constituents (like those in (6)) that are input to the headedness rules. Thus, Haegeman and van Riemsdijk (1986, 426) and Williams (this volume) allow the X'-level of verbal constituents subject to reanalysis/reassociation to be subject to parametric variation, but the headedness rules themselves (*flip, inversion*) are stated only over those verbs selecting verbal complements. For the Dutch particle cases, assuming the particle to originate as the complement of the lowest verb, [V'-V] reassociation preceding *flip* (in Williams's terminology) will yield the order in which the particle is intermediate in the cluster (following the higher verb), where [V-V] reassociation will "strand" the particle at the left edge of the cluster.[10] If *Flip* is restricted to the relative orders among verbs and their verbal complements, the Germanic generalization follows.

In principle, this should yields a less absolute impenetrability of the left-branching orders in Germanic than in Hungarian: the order 3>X(P)>2>1 or 3>X(P)>1>2 should be possible so long as X(P) is a dependent of one of the higher verbs. This might not be correct. Henk van Riemsdijk (personal communication, 7/2002) provides the following examples from Swiss German with an adverb and an object plausibly associated with the intermediate verbs (*wele* 'want' and *hälfe* 'help', respectively). These cannot immediately precede their associated verb when that verb is preceded by the lowest verb of the cluster.

(22) a.   *...wil*   *mer*   *s obst*   *hetted₁*   ***am liebste*** *wele₂* *verschnapse₃*
        ...because   we   the fruit   would've   most dearly want   distill
        'because we would have preferred most to distill the fruit into eau-de vie'

    b. [%]*...wil*   *mer*   *s obst*   *verschnapse₃* *hetted₁*   ***(*am liebste)*** *wele₂*
        ...because   we   the fruit   distill   would've   most dearly   want
        'because we would have preferred most to distill the fruit into eau-de vie'

---

[10] See Ackema (this volume) and Wurmbrand (forthcoming) and references therein for specific derivations; the latter also includes a survey of the various elements that may be interspersed in clusters in the various Germanic languages. Both Ackema and Wurmbrand also point out difficulties with deriving the precedence generalizations in antisymmetry-based approaches.

(23) a. *...wil   mer   wänd₁   **em Peter**   *hälfe₂ choche₃*
       ...because   we   want   DAT Peter   help   cook
       'because we want to help Peter cook.'

   b. $^{\%}$*...wil   mer   **em Peter**   choche₃   wänd₁ hälfe₂*
       ...because   we   DAT Peter   cook   want   help

   c. **...wil   mer   choche₃   {**em Peter**}   wänd₁ {**em Peter**}   hälfe₂*

In effect, current accounts from the extended headedness perspective imply that the correlation between order and impenetrability, if true, is at best a conspiracy of other factors (including, for example, the absence of right-edge of VP adverbs in German), at worst spurious. To the extent that the currently available extended headedness accounts allow us to understand (most of) the conspiracy of factors involved, the fact that the apparent correlation is not stated directly in the theory cannot constitute a criticism of these approaches. Nevertheless, these accounts might miss some important facts about VPR structures and internally leave a good deal to the imagination, (why should inversion be restricted to verbs selecting a verbal complement?) and it is not inconceivable that the development of an account that unifies the Hungarian facts more closely to the West Germanic facts might shed light on the various issues left as open questions in the present volume.

5. Concluding Remarks

Verb cluster constructions and the related topics of verb projection raising and particle climbing, seem to me to bear on some deep questions facing syntactic theory, specifically, those identified at the end of section 1: *headedness, arbitrariness*, and *linearity*. I have offered here my own reflections on the manner in which some of the material discussed in this book bears on these questions, or in some cases appears not to bear on the question where one might have originally hoped that it would. As I hope to have indicated above, though, it is clear that despite the significant progress made in the 25 years since Evers introduced the issue of the discontinuous dependencies in the Dutch clusters, there is still work to be done.

References

Ackema, P. This volume. Do preverbs climb?
Alberti, G. This volume. Climbing for aspect - with no rucksack.
Besten, H. den and Moed van Walraven, C. 1986. "The syntax of verbs in Yiddish". In *Verb Second Phenomena in Germanic Languages*, H. Haider and M. Prinzhorn (eds), 115-135. Dordrecht: Foris.
Bobaljik, J. D. and Wurmbrand, S. 2002. Long distance object agreement. To appear in *Proceedings of NELS* 33.
Bonet, E. 1991. Morphology after syntax. Unpublished Ph.D. Dissertation. Cambridge, MA: MIT.
Bródy, M. This volume. Word order, restructuring and mirror theory.
Chomsky, N. and Lasnik, H. 1977. "Filters and control". *Linguistic Inquiry* 8: 425-504.

Chomsky, N. and Lasnik, H. 1995. "The theory of principles and parameters". In *The Minimalist Program*, N. Chomsky (ed.), 13-127. Cambridge, MA: MIT Press.

Csirmaz, A. This volume. Particles and phonologically defective predicates.

Diesing, M. 1997. "Yiddish VP order and the typology of object shift in Germanic". *Natural Language and Linguistic Theory* 15(2): 369-427.

Dikken, M. den. 1996. "The minimal links of verb (projection) raising". In *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, W. Abraham, S. D. Epstein, H. Thráinsson and J. C. Zwart (eds), 67-96. Amsterdam: John Benjamins.

Evers, A. 1975. The transformational cycle of Dutch and German. Unpublished Ph.D. Dissertation. Utrecht: University of Utrecht.

Haegeman, L. and Riemsdijk, H. van. 1986. "Verb projection raising, scope, and the typology of rules affecting verbs". *Linguistic Inquiry* 17: 417-466.

Hinrichs, E. and Nakezawa, T. 1994. "Linearizing finite AUX in German verbal complexes". In *German in Head-Driven Phrase Structure Grammar*, J. Nerbonne, K. Netter and C. Pollard (eds), 11-38. Stanford: CSLI.

Huybregts, R. 1976. "Overlapping dependencies in Dutch". *Utrecht Working Papers in Linguistics* 1: 24-65.

Kathol, A. 2000. *Linear Syntax.* Oxford: Oxford University Press.

Kayne, R. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Koopman, H. and Szabolcsi, A. 2000. *Verbal Complexes.* Cambridge, MA: MIT Press.

Marantz, A. 1984. *On the Nature of Grammatical Relations.* Cambridge, MA: MIT Press.

Riemsdijk, H. van. 1998a. "Head movement and adjacency". *Natural Language and Linguistic Theory* 16(3): 633-678.

Riemsdijk, H. van. 1998b. Some remarks on verb clusters: the view from Germanic. Unpublished manuscript. Tilburg University.

Saito, M. and Fukui, N. 1998. "Order in phrase structure and movement". *Linguistic Inquiry* 29(3): 439-474.

Steedman, M. 1985. "Dependency and coördination in the grammar of Dutch and English". *Language* 61(3): 523-568.

Szendrői, K. This volume. A stress-based approach to climbing.

Szendrői, K. and Tóth, I. This volume. Hungarian verbal clusters - results of a questionnaire survey.

Twain, M. 1880. *A tramp abroad.* Toronto: Musson.

Williams, E. This volume. The structure of clusters.

Wurmbrand, S. 2001a. "How far can Agree see?". Paper presented at GLOW 24, Universidade do Minho, Braga, Protugal (April 2001).

Wurmbrand, S. 2001b. *Infinitives: Restructuring and Clause Structure.* Berlin: Mouton de Gruyter.

Wurmbrand, S. 2003. "Verb clusters: an overview". Talk given at the Joint conference of the Forum of Germanic Language studies and the Society for Germanic Linguistics. London (January 2003).

Wurmbrand, S. This volume. West Germanic verb clusters: the empirical domain.

Wurmbrand, S. Forthcoming. Verb clusters, verb raising and restructuring. Case study forthcoming. In The Syntax Companion. Oxford: Blackwell.

Zwart, C. J. 1996. "Verb clusters in Continental West Germanic dialects". In *Microparametric Syntax and Dialect Variation*, J. Black and V. Montapanyane (eds), 229-258. Amsterdam: John Benjamins.