

The Shift to Head-Initial VP in Germanic

Paul Kiparsky
Stanford University

1 Introduction

An interesting asymmetry in syntactic change is that OV base order is commonly replaced by VO, whereas the reverse development is quite rare in languages.¹ A shift to VO has taken place in several branches of the Indo-European family, as well as in Finno-Ugric. The Germanic languages conform to this trend in that the original OV order seen in its older representatives, and (in more rigid form) in modern German, Dutch, and Frisian, has given way to a consistently head-initial syntax in English, Scandinavian, and Yiddish:

- (1) a. German: *das Buch lesen*
Dutch: *het boek lezen*
b. English: *to read the book*
Swedish: *att läsa boken*
Yiddish: *leyenen dos bukh*

The Germanic family offers an excellent laboratory for investigating the causes and conditions of this drift. By approaching the languages in comparative perspective, we can hope to discover what triggered this convergence towards VO word order, why it took place precisely in these three branches of Germanic, and what relation (if any) it has to other changes that they underwent, such as the rise of V2 syntax, the impoverishment of case and agreement morphology, and the loss of “scrambling”.

Modulo various differences in theoretical assumptions, the literature offers three main types of answers to these questions:

(1) *Language contact*: VO replaced OV under the influence of neighboring VO languages.

¹Thanks to Andrew Garrett, Hans-Martin Gärtner, Ans van Kemenade, Anthony Kroch, and three reviewers for their trenchant comments and suggestions on an earlier version of this paper. I am also grateful to Cleo Condoravdi, Peter Sells, and Elizabeth Traugott for helpful advice at various stages of the work.

(2) *Grammaticalization of word order*: The loss of inflectional morphology gave rise to VO, because it is the word order best suited for marking the distinction between subject and object.

(3) *Reanalysis of derived VO sequences*: When increasingly frequent application of optional movement rules made the main clause evidence for OV opaque, VO arose through reanalysis of synchronically derived Verb-Object sequences.

The solution to be defended here resembles the latter two in attributing the word order change to structural causes and linking it to other changes that took place in these languages, but otherwise differs from them. While grammaticalization theories posit an efficient cause (a preference for VO, whether functionally grounded in the unambiguous marking of grammatical relations, or formally in some intrinsic D-structure asymmetry), and reanalysis theories posit an enabling cause (opacity of OV), I contend that internal causation of particular events of linguistic change (as opposed to borrowing) normally involves an interaction of efficient and enabling causes, and propose an account of the word order shift with such a dynamics. And the causal factors themselves which I propose are also different:

(4) *Harmonization of the direction of complementation*: Where the subordinate clause evidence for the OV base became opaque through the rise of embedded finite verb fronting in embedded clauses (the enabling cause), VO arose as a generalization of rightward complementation, already in force in the other functional and lexical categories in Germanic (the efficient cause).

In what follows I first examine these theories more closely and lay out the theoretical and empirical arguments in favor of the fourth, discussing in particular its relation to Kayne's hypothesis that heads universally have their complement on the right. I then present an analysis of Old English word order variation which supports the proposed historical account.

2 Explaining the Word Order Shift

2.1 Language Contact

The word order shift has been seen as part of a "pan-European drift" (Hock 1986), extending beyond Germanic to Baltic, Slavic, and Greek, and even further to non-Indo-European languages such as Finnish and Estonian. Danchev 1991 goes so far as to speak of "creolization" as the factor responsible for the change in Old English.

A priori, language contact is a reasonable hypothesis here, for it is clearly a major factor in syntactic change. For example, the syntactic divergence of Balto-Finnic from the rest of Finno-Ugric was undoubtedly caused at least in

part by protracted contact with Indo-European languages over a long period of time, including over 2000 years of Germanic influence.

However, the distribution of VO vs. OV within the different branches of the Germanic family correlates poorly with the degree of contact which they have had with VO languages. Dutch and German have had more contact with other European language families (especially Romance) than the Scandinavian languages have. Icelandic, of course, has been especially isolated throughout its history, and has consequently borrowed very little in any part of its grammar or lexicon. If contact were the primary cause of the change, we might have expected Dutch and German to have adopted VO and Scandinavian to have retained OV, which is just the reverse of what actually happened. English, to be sure, was profoundly influenced by French, and before that by Scandinavian. But the VO option is established in Old English prior to the French contact that began with the Norman conquest (Pintzuk 1991).² During that early period, on the other hand, the Scandinavian languages were themselves still predominantly OV, as far as is known, though they were already Infl-medial.³ Therefore, it is unlikely that the word order shift in Old English is wholly due to the influence of other languages. The contact story looks best for Yiddish. The change to VO might be attributed to borrowing from the coterritorial Slavic languages,⁴ or, as a reviewer suggests, it may have arisen even earlier on the basis of the VO syntax of the Ladino spoken by the Jewish migrants from Spain, from whose German the Yiddish language presumably originated.⁵

The attribution of the OV to VO shift to borrowing would also not explain the slow and steady spread of the change in each of the languages (Santorini 1989, 1993b, Pintzuk 1991, Rögnvaldsson 1993). The change does not accelerate during periods of intensive language contact or slow down during periods of relative isolation, as far as we can tell. Instead, OV and VO coexist for a long time — as long as six centuries in Icelandic — during which VO steadily encroaches on OV, at a rate seemingly unaffected by changes in the sociolinguistic situation.

I conclude that, although borrowing might be an ancillary factor in the change from OV to VO in some of the languages, and quite possibly even its main cause in Yiddish, there probably were some internal factors involved as well, working in favor of the change in languages like Icelandic and against it in

²Stockwell and Minkova 1991 also reject French influence as a major factor in the word order change, on the grounds that the chronology is wrong.

³According to Falk (1993, 161), early Old Swedish was OV and Infl-medial; in a lecture at the Melbu Summer School in 1993 Rögnvaldsson estimated, on the basis of a statistical study of the OV to VO shift, that Old Icelandic in 1000 A.D. would have had roughly 90% basic OV.

⁴See Santorini 1989 for discussion of several possible scenarios for how Yiddish syntax might have developed under Slavic contact.

⁵Santorini (1989, 1993b) establishes that the earliest Yiddish was Infl-medial, but it is not clear whether it was already VO.

languages like Dutch. What these internal factors were is the question to which the rest of this paper is devoted.

2.2 Grammaticalization

The idea that the shift to VO was caused by the loss of inflections and is directly associated with the loss of free word order can be found in Sapir 1921 and Jespersen 1938; Vennemann 1975 is probably the most influential modern formulation of it. Roberts 1994 reinterprets this link within the minimalist program (Chomsky 1993), replacing the traditional functionalist understanding of the privileged role of VO with a formal account based on the hypothesis that the head-complement order is universal (Kayne 1992, Zwart 1993), to which we will return below.

The historical and comparative evidence does not support a direct causal connection between the loss of inflection and the shift to head-initial VP, at least if we take inflection to mean overt case morphology and/or rich agreement in the sense of Roberts 1993. Whether a language has inflection in that sense does not seem to correlate with the VP's orientation at all. All four possible combinations are well-attested cross-linguistically, and all are instantiated even in the Germanic family:

(2)

	OV	VO
inflection	German, Old English	Yiddish, Early Middle English
no inflection	Dutch, Frisian	Swedish, Modern English

So the the orientation of VP and inflection must be considered independent properties.

The specific course of the change in the individual Germanic languages and its chronology point in the same direction. The shift to VO began in Old English before the collapse of the case system and before the loss of subject-verb agreement. Late Old English is well on its way to VO order, yet retains the morphological distinction between nominative and accusative case in noun inflection. Icelandic has even maintained the full four-case system and morphological agreement intact, yet has adopted fixed VO order. And on the other hand, Dutch did lose morphological case, yet kept OV as the basic word order. The same thing is true of the diachronic relation between the change in headedness of the VP and the establishment of fixed word order. In English, the VO base was dominant by 1200, whereas the loss of scrambling can be dated to the last quarter of the 14th century (Allen 1992).

What the historical and comparative evidence does support, though, is the claim that the loss of inflectional morphology entails rigid order of direct nominal arguments. Every Germanic language which has lost case and agreement

morphology, whether VO (English, Swedish, Danish, Norwegian) or OV (Dutch, West Flemish, Frisian, Afrikaans), has imposed a strict mutual ordering requirement on its nominal arguments, without changing the orientation of its VP. The order is always that subjects precede objects, and indirect objects (NPs) precede direct objects. It is the same as the unmarked order of those languages which do scramble freely, again whether VO, as in Yiddish, or OV, as in German (Haider 1993).⁶

(3) a. German:

1. ... dass Jan seinem Vater das Buch gibt.
... that Jan his father the book gives
'that Jan gives his father the book.'
2. ... dass Jan das Buch seinem Vater gibt.
3. ... dass seinem Vater Jan das Buch gibt.
4. ... dass seinem Vater das Buch Jan gibt.
5. ... dass das Buch Jan seinem Vater gibt.
6. ... dass das Buch seinem Vater Jan gibt.

b. Dutch (Vikner 1991, Ch. 4, Haegeman 1993, Neeleman 1994, Zwart 1993, 303):

1. ... dat Jan zijn vader het boek geeft.
2. ... *dat Jan het boek zijn vader geeft.
3. ... *dat zijn vader Jan het boek geeft.
4. ... *dat zijn vader het boek Jan geeft.
5. ... *dat het boek Jan zijn vader geeft.
6. ... *dat het boek zijn vader Jan geeft.

c. Swedish:

1. ... att Jan ger sin far boken.
2. ... *att Jan ger boken sin far.
3. ... *att sin far ger Jan boken.
4. ... *att sin far ger boken Jan.

⁶ I set aside here so-called focus scrambling as being a separate phenomenon (Zwart 1993, p. 246, with references). Also, it is important that the implication only goes one way, from lack of morphology to fixed argument order. Many languages with rich inflection require fixed word order anyway (among the Germanic languages, Icelandic is the most notorious case, but there are others). A second key point is that languages with fixed order of nominal arguments differ systematically in one important additional respect, which exactly correlates with the VO/OV difference (Neeleman 1994). Rigid VO languages (English, Swedish, Danish, Norwegian, Icelandic) require adjacency of verb and object, and of the two objects in a double object construction, whereas rigid OV languages (Dutch, West Flemish) allow adverbs to intervene freely in both these positions (Holmberg 1985, Vikner 1991, 310, Zwart 1993, 302). Adjacency presumably confers a perceptual advantage, but it is not clear that processing considerations can explain why the adjacency requirement is imposed only in VO languages. See Neeleman 1994 for a proposal for explaining this asymmetry within grammatical theory.

5. ... *att boken ger Jan sin far.
6. ... *att boken ger sin far Jan.

These three generalizations — that presence or absence of inflection is independent of the directionality of the VP, that lack of inflection implies fixed order of nominal arguments, and that the fixed order is identical with the unmarked order of nominal arguments in “free word order” languages — hold widely outside of the Germanic family as well; in fact, there are very few known exceptions to them in the languages of the world.⁷ Why this should be so must be considered an open question. The T-model architecture of GB-theory, and those of most of its competitors and successors, separate syntax from overt morphology (“spellout”) in such a way as to preclude a principled theoretical connection between them. In Kiparsky MS I develop a licensing theory in which the three generalizations, as well as the observations in fn. 6, follow from grammatical principles. That theory exploits two main ideas: (1) that Th-role assignment must be licensed by case features which are borne by position as well as by inflectional affixes, and (2) that the order Subject – Indirect Object – Direct Object seen in (3) is a linearization of the Theta-hierarchy (Uszkoreit 1984, Wechsler 1991), which is itself a projection of semantic form (Bierwisch 1983).

2.3 Reanalysis

The majority of linguists who have recently discussed the change to head-initial VP have attributed it to reanalysis of derived surface VO sequences as basic VO sequences. Opinions differ on which kinds of surface VO sequences are assumed to provide the basis for the reanalysis. Most versions stress the role of various rightward movement rules, in particular, of NP shift and extraposition, and also of adjuncts positioned after the verb as “afterthoughts”, as being responsible for diluting the evidence for underlying OV order (Stockwell 1977, Canale 1978, Lightfoot 1979, and van Kemenade 1987 for English, Sigurðsson 1988 for Icelandic, and Santorini 1989, Ch. 4 for Yiddish). van Kemenade also points to verb raising as another source of surface VO syntax which could have contributed to the change by making the OV order opaque.

Another class of reanalysis scenarios invokes the surface VO sequences derived by V-to-C or V-to-I movement in main clauses as a source.⁸ Stockwell and Minkova 1991 propose that the English word order shift involved the following series of changes involving successive mutual adaptation of main clause

⁷Notably a group of South American languages in which objects normally precede subjects (Derbyshire and Pullum 1981). However, there are good reasons to treat the apparent nominal arguments in these languages as appositions, the real arguments being the pronominal elements associated (in the “right” order!) with the verb.

⁸Santorini (1989, 165 ff.) briefly considers this theory for Yiddish and rejects it on the grounds that it does not extend to languages which shifted from I-final to I-medial syntax *without* ever having observed a V2 constraint (citing Koopman 1984:152, fn. 5, on Vata).

and subordinate clause word order patterns: (1) subordinate clauses adopted V2 order by analogy with V2 in main clauses, (2) because subordinate clauses underwent Topicalization relatively rarely, the V2 of subordinate clauses was reinterpreted as SV, (3) the SVO order introduced into subordinate clauses by this reanalysis in turn triggered a reanalysis of main clause V2 word order as SVO order.⁹

Reanalysis based on main clause V-to-C movement is also advocated by Lightfoot 1991, in the context of a parameter-setting model of change which crucially assumes that acquisition must be based on main clauses only, so that the verb-final syntax overtly displayed by subordinate clauses is in principle not accessible to the learner. Here the triggering factor of the reanalysis is seen as random variation in the input, resulting in increasingly frequent exercise of the V2 option: “This no more reflects a difference in grammars than if some speaker were shown to use a greater number of passive or imperative sentences. Rather, it reflects the kind of accidental variation that is familiar from studies in population genetics. Nonetheless, changes in the primary linguistic data, if they show a slight cumulative effect, might have the consequence of setting a grammatical parameter differently” (p. 67-68).

Implicit in Lightfoot’s account is the claim that the change could as well have gone in the opposite direction and that there is no specific reason why it happened precisely when and where it did happen. One might question this on the grounds that random fluctuations in the deployment of linguistic options can plausibly trigger qualitative changes if they involve low-frequency data points, such as the use of particular lexical items, but surely not rules as central to the grammar as V-to-C movement, to which practically every sentence bears witness. Variation studies show that speakers replicate the major variation patterns of the community quite faithfully, and crucially, that even young speakers do so (Labov 1989). Moreover, if random variation in frequency of usage were the cause of change, as proposed by Lightfoot, we should expect oscillation in usage, rather than the steady increase in frequency of the innovation forms that the historical record actually shows.

Neither does the record show a sharp discontinuity such as would be expected at the putative point of reanalysis.¹⁰ Each language had OV and VO base orders

⁹The idea that the change originated in main clauses, and was generalized from there to all clauses is also explicit in Stockwell 1977, Lightfoot 1982, and Hock 1986. However, leveling of main and subordinate clause word order is problematic as an explanatory principle in view of the fact that there has been an equal amount of *differentiation* between main and subordinate clause word order within Germanic, viz. the rise of V2 itself (Kiparsky 1994), and the development of a distinction between main and subordinate clause order in all the mainland Scandinavian languages (analyzed as the loss of V-to-I movement, Platzack 1988).

¹⁰English is an apparent exception here. The statistics of Hiltunen 1983 and Bean 1983 seem to show that the shift from OV to VO was relatively rapid in embedded clauses (I am not aware of similar indications for the other languages). Lightfoot 1991, 65 ff. indeed takes this as evidence that the reanalysis in embedded clauses is “catastrophic”. A. Kroch, in discussion,

side by side for a long transitional period, during which VO steadily increased in frequency until it finally became categorical. This period began in Old Icelandic and was not definitively completed in the literary language until the first quarter of the 19th century (Rögnvaldsson 1993). In English, the the OV/VO phrase structure variation begins in Old English (it is not clear how early) and is not completed until late Middle English (Pintzuk 1991). And Santorini (1993a) shows that in Yiddish the change began at least 500 years ago and has still not gone to completion. This casts doubt on all “catastrophe” scenarios in which increasingly frequent use of VO structures (from whatever source) would at some point have critically altered the learner’s triggering experience, causing an abrupt reanalysis to VO, after which residual occurrences of OV orders would have been derived by a minor “backwards” movement rule.

Accordingly, Santorini, Pintzuk, and Rögnvaldsson argue instead that during this transitional period, the languages had two coexisting base structures, OV and VO; a similar position is taken in Weerman (1989).¹¹ A theoretical basis for this approach to syntactic variation has been laid by Kroch (1989); see also Santorini (1992, 1993b). An important additional advantage of this assumption, as Pintzuk points out, is that it makes unnecessary the complex reverse (“adaptive”) rules that would have to be posited for the period of variation if a single uniform base is assumed. The present paper will add support to this so-called double-base hypothesis.

At the same time, the discovery of the gradual actuation of the change poses a challenge to *any* historical account, irrespective of whether it is based on reanalysis or not. In the initial stages of the change, when the evidence available to the learner overwhelmingly favors OV, what allows the innovating VO base to gain a foothold? And what keeps it going forward at a steady rate? In the cases of reanalysis familiar from phonology, a plausible driving mechanism often exists, such as simplification (reanalysis of derived forms as basic allows elimination of the rule that derived them) or elimination of abstractness (reanalysis of derived forms as basic makes underlying forms easier to learn). Neither of these causal factors can be responsible for the word order shift under discussion, though. Rule loss (whatever that might mean in syntax) could not be at work because the relevant rightward movement processes, such as Extraposition and Heavy NP Shift, however they are to be analyzed, clearly remain in operation

suggests that this asymmetry is probably a statistical artifact of the sparse attestation and changed sociolinguistic status of Early Middle English during the twelfth century (see also Kroch and Taylor MS.). After the Norman Conquest, standardized literary West Saxon gave way to a variety of previously unwritten vernacular forms of English. This resulted in a sharp transition in the written literature, which cannot be taken as evidence of a catastrophic change in the language itself.

¹¹More precisely, Rögnvaldsson’s position is not that there were two competing systems but that there was one system in which the value of the VO/OV parameter is unspecified. The difference at first sight seems unimportant. But in section 3 below I show that they are empirically distinct, and that the evidence supports the Kroch/Santorini/Pintzuk competing grammars hypothesis.

even after the VO base gets established (indeed, they are still active in modern English). So do the verb-fronting processes (V-to-C and/or V-to-I), even though these are later lost in some of the languages. As for abstractness (whatever *that* might mean in syntax), it is not at stake either because the reanalysis from OV to VO does not make V-to-I or V-to-C movement, or any other process for that matter, any less “abstract” by *any* reasonable interpretation of the term. So we are left with the question: what is the “motor” of this change?

Besides explaining how the direct evidence for OV from overtly verb-final subordinate clauses became unavailable to learners or was overlooked by them, a reanalysis account of the OV to VO shift must explain how learners could have neglected the more indirect main clause OV evidence as well. And, like any account of the change, it has to say something about why the change took place only in English, Scandinavian, and Yiddish, and not in German, Dutch, and Frisian. The strongest possible confirmation of a specific reanalysis solution would be a demonstration that the initial conditions prior to reanalysis differed in the relevant way in the two sets of languages. In the case at hand, this would involve showing that learners of the continental Germanic languages at an earlier stage had access to evidence for OV which was absent in Old English and early Scandinavian, and that Yiddish at a later date diverged from German in such a way that this evidence became unavailable there too.

An important step in this direction was taken by Lightfoot (1991, 60) in the context of his theory of Degree-0 learnability, which restricts the evidence accessible to learners for setting parameters to main clauses. Lightfoot’s thesis is that OV was less learnable in Old English than in the continental Germanic languages, because the clues which reveal it in main clauses were weaker. He deals with four types of evidence: the position of particles relative to the verb, the position of negatives relative to the verb, the position of infinitives relative to the verb, and verb-final main clauses.

In fact, however, the evidence for the OV base in Old English was in no way weaker than in the Germanic languages which retained OV. The opposite is true. Consider first Lightfoot’s argument that negatives “move with the verb”, as in (4):

- (4) a. ne geseah ic næfre ða burh (*ÆCHom* I.527.3)
 not saw I never the city
 ‘I never saw the city’
- b. þeah se lærow ðis eall smealice & openlice gecyde, ne forstent it
 though the teacher this all carefully and openly tells, not avails it
 him noht (*CP* 163.18)
 him nought
 ‘though the teacher tells all this carefully and openly, it is of no avail
 to him’

One can counter with the point that the doubled negatives (such as *næfre* and *noht* in (4)) provide evidence for the underlying place of the verb. In any case, the same construction also exists in Old High German:

- (5) ni mac unser neheinir chomen in daz himelrih (MSD 280,34)
 not can of us no-one come into the kingdom of heaven
 ‘no-one of us can come into the kingdom of heaven’ (Behaghel 1924, 73)

Lightfoot’s argument from the order of infinitive and finite verbs is similar; again, verb raising is uncontroversially found in Old English as well as in Dutch/German (van Kemenade 1987, 55 ff., Pintzuk 1991, 99 ff.).

Heavy NP Shift and related rightward “leakages”, which most of the other authors see as the source of the reanalysis, are also found in all the old Germanic languages, and they are indeed quite common in earlier stages of German and Dutch. For Middle High German, see e.g. Paul-Schmitt 1950, 142 ff. In Middle Dutch, Burridge (1993, 80) finds that 14th century Brabantish and Hollandish texts have a frequency of “exbraciation” ranging from 24% to 58% in various types of subordinate clauses, and no less than 52% to 64% in main clauses, involving every one of the processes to which reanalysis theories have attributed the Old English OV to VO shift.

Lightfoot’s argument that Old English sometimes permitted particles to shift, so that they no longer provided reliable evidence for the underlying verb, cannot be maintained. Pintzuk (1991, 178) shows that Old English *had* no particle movement, on the grounds that particles never appear postverbally in Infl-final clauses. Therefore, the position of particles must have available as an OV clue prior to the word order change. On the other hand, rightward particle movement does occur in older German and Dutch. Neeleman and Weerman (1993, 190) note that, in addition to extraposed objects and postposed adverbials, particles and resultatives can also appear postverbally in Middle Dutch embedded clauses:

- (6) die dese niemare hoerden an (Middle Dutch, MNW 1, 110)
 who these messages heard at
 ‘who listened to these messages’ (from Neeleman & Weerman 1993, 190)

In this respect, then, the evidence for OV was *stronger* in pre-Old English than in continental Germanic.

As for verb-final main clauses, it is true that they are not very common to begin with (outside of coordination) and that they diminish in frequency during the Old English period, as Lightfoot points out. That said, the very existence of such clauses in Old English, such as (7a,b), also speaks against Lightfoot’s theory, since they are reportedly almost absent in Old High German.¹² Moreover conjoined verb-final main clauses are quite frequent in Old English, e.g. (7c).

¹²To be sure, even V2 clauses can end up as verb-final if everything after C° is moved or deleted. The occurring examples of non-trivially verb-final main clauses in Old High German are attributed by Behaghel 1932, §1432 to Latin influence.

- (7) a. He þa his here on tu todælde (*Orosius* 116.16)
 he then his army in two divided
 ‘he then divided his army into two’
- b. Hie þa hine on rode ahengan (*BH* 177.25)
 they then him on cross hung
 ‘they then hanged him on the cross’
- c. & him þa Ioseph, rihtwis man, mid godcunde fultume gehealp
 and him then Joseph, righteous man, with holy aid helped
 (*Orosius* 32.26)
- ‘and Joseph, the righteous man, then helped him with holy aid’

Lightfoot’s learnability thesis does not exclude conjoined main clauses from the evidence available to the language learner. Therefore, the relatively frequent verb-final conjoined main clauses in Old English constitute a type of main clause OV evidence that was not available in continental Germanic.

I conclude that Old English, prior to the word order change, offered the learner the same OV clues in main clauses as the continental Germanic languages did; in addition, its V-final main clauses and lack of particle movement provided two additional signposts that they lacked. Moreover, every process that compromised the OV clues in Old English was also applicable in the other languages. Qualitatively, at least, the main clause evidence for the OV base was actually stronger in Old English than in OHG.

For these reasons, a Lightfootian main-clause based reanalysis story would have to resort to a frequentistic argument. It would have to postulate that the main clause diagnostics were *quantitatively* weaker in those languages that that later underwent OV to VO shift. In Old English, they would indeed have had to be so feeble that learners could fail to acquire OV order even given the direct evidence of overtly verb-final main clauses and of particles (which as noted don’t shift and therefore give reliable evidence for the underlying verb position). The claim would presumably have to be that the cumulative force of the OV diagnostics fell below some critical threshold (the size of which would of course have to be specified). The relevant processes would include at least extraposition, heavy NP shift, verb raising and verb projection raising, particle movement (in those languages that have it), as well as the placement of adverbs, resultatives, and negation. There is little statistical data to go on, but from Pintzuk (1991, 113 ff.) we do know that Old English has *less* verb projection raising than continental Germanic does, again the opposite of what a main-clause based reanalysis account would require. In any case, how likely is it that usage with respect to these numerous independent processes would shift enough to jointly cause a significant decrement in the evidence for the OV base, and that this would happen independently in several branches of the Germanic family? If such a conspiracy of shifts in usage were discovered, surely the plausible

interpretation of them would be as so many symptoms of an emerging VO base option.

A frequentistic alternative which is not open to these objections is the conjecture that “complex verbs” (presumably infinitives + auxiliaries or finite verbs) might have been more frequent in Dutch than in Old English, in virtue of which Dutch would have provided proportionately more evidence for verb-final base order (Gerritsen 1984, Burridge 1993, 114). If this asymmetry holds up,¹³ and turns out to divide also the other Germanic languages in the right way,¹⁴ it would constitute a single factor, rather than an unexplained conspiracy, and one not open to the suspicion of being an effect of the word order shift rather than a cause of it.

However, the notion that there is some critical threshold of frequency with which a parameter setting must be attested in the data in order to be acquired is itself questionable for several reasons. It would seem incompatible with historical studies which show that fairly marginal options can be passed on from one generation to the next for several centuries. In Yiddish, for example, Santorini (1989, 134) finds that Topicalization in embedded clauses has remained a stable low-frequency option (exercised in 2%-4% of the possible cases) from the time it is first attested in the seventeenth century down to the present. Prior to that time, it was not a possible option at all. Opinions diverge on what the difference between the grammars of these two stages is (Santorini 1989, van Kemenade MS), but it has a typically parametric flavor. Another example is the parameter setting allowing (optional) long-distance anaphora out of infinitive complements in Swedish. This is acquired robustly even though the proportion of reflexives encountered in the data that actually require long-distance antecedents is small.¹⁵ Such cases show that learners can acquire and transmit parameter settings on the basis of evidence which is vastly less fre-

¹³Dennison 1986, 1993:287 expresses doubts about the alleged rarity of compound verbs in Old English. Burridge’s comparison is based solely on the Parker MS, which because of its chronicle style is perhaps not representative in this respect. A count of random passages of Bede’s History and Alfred’s Boethius turned up 20-30% of complex verbs, which agrees with Dennison’s figure of 27%. This is less than what Burridge reports for Middle Dutch, but not dramatically so (e.g. 31% - 52% for the 14th century).

¹⁴Which seems even more doubtful. Yiddish would seem to use *more* compound tense forms than German or Dutch, for example.

¹⁵As far as I am aware, proponents of frequency thresholds in acquisition have not specified what the threshold is, nor even what exactly it should be a statistical measure of. In Swedish, most reflexives have local antecedents (over 99% in my sample), simply because the infinitive constructions in which long-distance binding is allowed are not all that common. Perhaps, though, it is only the “crucial cases” that count for the learnability threshold, i.e. the proportion of reflexives with actual long-distance antecedents to reflexives with *potential* long-distance antecedents. In my sample, the relevant constructions had 4 examples of reflexives with an intended long-distance antecedent against 8 examples of intended local binding, and 2 where the long-distance antecedent is referred to by a pronoun rather than by a reflexive (“counterevidence” which the learner must overcome in order to learn that reflexives *can* have long-distance antecedents).

quent in speech than the main clause evidence for the OV base in Old English. Therefore, short of assuming that the critical threshold is parameter-specific (or, worse still, construction-specific), *the loss of OV cannot be explained solely from a insufficiency of data supporting it.*

The claim that learners ignore statistically marginal evidence has sometimes been justified on the grounds of their ability to detect production errors in the corpus. The fact that production errors are, in most cases, successfully weeded out from the primary linguistic data can however be explained on the non-statistical grounds that they normally cannot be accommodated in the grammar without violating principles of UG. In fact, it has been argued that those production errors which *are* consistent with grammars permitted by UG *can* get acquired, and constitute one of the sources of analogical change, e.g. certain kinds of analogical regularizations in morphology (Haber 1975).

The assumption that learners at a certain stage simply miss the evidence for OV is also inconsistent with the unanimous verdict of the acquisition studies summarized in Wexler 1994 (see in addition Boser et al. 1992, Poeppel & Wexler 1993). These show that children master the differing distribution of finite and nonfinite verbs as early as their second birthday. Learners of German and Dutch at that age use finite forms consistently in V2 position (and in V1 position in questions). That is, they know from a very early age that nonfinite verbs do not move. Crucially, the evidence includes studies from VO languages such as Swedish. Wexler, examining Platzack's 1990 Swedish data from three children aged 20-26 months, finds that V2 is found only with finite verbs (30 occurrences in the corpus); nonfinite verbs never appear in V2 position. If two-year olds have already got V2 down perfectly in VO languages, how could VO itself be the result of learners' misanalysis of V2?

In summary, pure reanalysis scenarios of the word order shift are inconsistent with the gradual spread of the change, underestimate the ability of language learners, and do not explain why the Germanic languages diverged as they did. There is no evidence for the assumption that parameter settings must be supported by some critical proportion of the relevant input data: in the present state of the art, the opposite view that a parameter setting can in principle be acquired from a single critical example has an equal claim to plausibility.

2.4 Endogenous optimization

Grammaticalization and reanalysis present essentially complementary perspectives on the causes of change. Grammaticalization theories of the word order shift specify the motivation of the change (optimal encoding of grammatical relations, or the privileged status of rightward complementation), but they have little to say about the specific circumstances which permit the transition to a new system. Reanalysis theories focus precisely on that question, by locating

points of structure which are ambiguous or features which are weakly entrenched in the primary linguistic data available to the learner at the stage just prior to the change, and by specifying constraints on the learning mechanism that allow such ambiguities and data gaps to trigger the observed reanalyses.

It would be a mistake to think of these two kinds of explanations as mutually exclusive competing alternatives. In fact, in so far as explanations of particular linguistic changes (*qua* internally caused historical events) are possible at all, they must arguably involve the interaction of both factors. Grammaticalization (and more generally any kind of structural optimization) does not by itself solve the actuation problem, or explain how a given change can take place in one language but not in another closely related one. Theories relying on such mechanisms typically fail to address the fact that suboptimal structural features do arise in language change, and can be maintained stably for a long time. Reanalysis theories on the other hand attempt to characterize the preconditions for change, but not its motivation, and typically fail to address the problem of drift and directionality in linguistic evolution.

A genuine explanation of a specific instance of change, then, must spell out two things: the factors that induced learners to prefer the innovating structure, and the factors that made that innovating structure attainable at a certain point on the basis of the data to which learners were exposed. In the case at hand, this would mean explaining why VO was preferable to OV in the Germanic system, and what weakened the evidence for OV enough in English, Scandinavian, and Yiddish to enable that preference to be implemented in those languages.

The two components of this push-pull mechanism — the effective cause and the enabling cause — are rather different in character. The effective cause is a preference which is general, but which only manifests itself when conditions allow — for example, when the evidence for the old structure (right-headed VP, in this case) became attenuated. This preference, structural optimization, in this case, could be either absolute and language-independent, or relativized to the specific structure of the language (an “implicational universal”). The enabling cause is by its very nature language-specific; this is where we would look for the point of difference between those Germanic languages which shifted to VO order and those which did not. On a reanalysis account of the change it would have to do with the degree of learnability of the old structure and the degree of saliency of the evidence for it. If we adopt the optimization account proposed here, these considerations remain relevant, of course. However, the pull factor might then also involve other conditions, such as a requirement that the successive stages of a language should differ minimally in output terms, whether for sociolinguistic or intelligibility-related reasons.¹⁶ It might well depend also on extragrammatical considerations, such as the sociolinguistic context, the radius

¹⁶The notion of continuity is fundamental in the theory of linguistic change developed by Tabor 1994.

of communication, or the existence of diglossia or multiple norms. And it would be applicable to change from any source, including borrowing.

A change could then be triggered either by an increased push towards the new (by another change which made the innovating structure more highly valued) or by a decrease in the pull of the old (by something which caused the old structure to be less entrenched or less attainable).

My proposal has the latter character. As the fixed push factor which causes learners to opt for the VO analysis I propose the preference for uniform directionality of head-complement relations. The Germanic system had a complex mixture of rightward and leftward Th-role assignment. Functional categories took their complements on the right. This is clearly true for Comp, and Infl, in so far as it existed (see below). Prepositions also Th-marked rightward. For other lexical categories the picture is mixed. Verbs took their complements on the left, except for sentential complements, which were obligatorily on the right. Adjectives (and perhaps nouns) seem to have taken nominal complements on the left and PP complements are on the right (as is still the case in modern German: *ihm nützlich* / *nützlich für ihn* “useful for him”). In VO languages, all these complications have been eliminated. The mixed system of Germanic has been reduced to a fully harmonic system of consistent rightward complementation.¹⁷ Thus, the shift to VO is a simplification of the grammatical system: in the new grammar, *all complements of all heads, lexical and functional, are licensed uniformly on the right*.

A deeper explanation of the shift to VO shift is suggested by Kayne’s (1993) idea that the Spec-Head-Complement order is universal, so that *all* languages are basically VO, and apparent OV languages simply have extensive leftward movement processes. The change from (apparent) OV to VO could then be simply a loss of those leftward movement processes. It would then be motivated not only internally to the Germanic system, as discussed, but on a language-independent basis as a grammatical simplification. The theory would then predict not only that the mixed system of head-complement relations of Germanic would become uniform, but also the direction in which the system was harmonized. In addition, it would also account for the cross-linguistic asymmetry we noted earlier, that OV commonly changes to VO but the reverse does not happen.

But this more ambitious version of the story raises the stakes. If VO is the universally preferred system, why don’t all languages have it? Or putting the question in historical terms, how do OV type systems, *ex hypothesi* suboptimal, come about in the first place? This problem is sometimes considered a decisive flaw of optimization theories. As Lightfoot (1987, 155) puts it, in a discussion of a study by Keyser and O’Neil: “But for optional rules to be lost, they must first

¹⁷See Webelhuth 1992 for a comparative analysis of directionality the modern Germanic systems. On cross-category harmony of head-dependent ordering as a typological principle see Hawkins (1983); Dryer (1990) argues that the real parameter is direction of branching.

be introduced; if we are predisposed not to attain optional rules, one wonders how they would be triggered in the first place. ... historical linguists often succumb to the temptation to see a general directionality to change, and to explain this either by invoking laws of history (see Lightfoot 1981) or by attributing historical effect to genetic predispositions.” This objection misses the mark. The fact is that there is not a single dimension of optimization but numerous partially conflicting ones. As historical linguists have always recognized, there are many undeniably valid general tendencies in language change, such as assimilation and other natural sound changes, paradigmatic leveling and other forms of analogical simplification, and grammaticalization. There is nothing objectionable or conceptually mysterious about such tendencies, and indeed no intrinsic problem with viewing them as optimizations, as long as we recognize that the putatively more complex structures which they function to eliminate can themselves be the results of other, cross-motivated processes and tendencies, and that they in turn can give rise to complexity at some other point in the system. To take a typical phonological scenario, the natural weakening process of deleting unstressed vowels can create consonant clusters, which are then simplified by assimilation into geminate consonants, which are in turn degeminated, then weakened, and eventually deleted, causing vowel hiatus, which is then eliminated, and so on. Thus complexity in language can arise as a by-product of simplificatory processes.

We can therefore legitimately posit a universal preference which is not universally instantiated provided that we spell out the other factors that allow (or force) it to be subverted. In principle, they might be either intersecting structural or functional constraints,¹⁸ or historical processes. Motivating the latter would, in the present instance, amount to demonstrating a natural origin for SOV syntax.

A historical explanation along these lines might be that SOV systems develop from systems in which the syntactic arguments are pronouns cliticized to the verb and the nominals are optional appositions to them, normally occurring in left-dislocation position. In time, the clitic pronominal arguments become morphologically incorporated into the verb. An SOV system then develops by the usual grammaticalization process in which the incorporated pronominal arguments on the verb become agreement affixes, and their function as arguments is taken over by the former nominal adjuncts, which initially retain their linear position.¹⁹ The life-cycle would thus be: clitic pronouns → incorporated pronouns → agreement (with SOV order), followed by loss of agreement and/or by SOV → SVO.

¹⁸For example, marked feature specifications incur an extra cost, but every language must have some of them in order to have enough phonemes in its system.

¹⁹One episode of this development can actually be seen taking place within late Indo-European, namely the shift of sentential complements from their Indo-European status as adjuncts of pronominal arguments, to arguments in their own right (Kiparsky 1995).

But why should the grammaticalization of nominal appositions to subject and object arguments yield an SOV system, rather than forcing a direct transition to the putatively optimal SOV structure? Indeed, how can we account for the fact that SOV is maintained stably for millennia in numerous languages of the world (including of course continental Germanic)? Zwart 1993 suggests that OV order can in fact be motivated by specific structural properties of the languages that have it. On his analysis, it arises by syntactic movement of objects from their underlying postverbal position to Spec-Agr-O. The idea is that movement to Spec-Agr-O takes place overtly in the syntax if necessary, and covertly at LF if possible. The distinction is a matter of whether the relevant agreement feature is “strong” or “weak”. Pending some independent motivation of the postulated strong/weak distinction in the status of Spec-Agr-O, this relocates the problem but does not solve it.²⁰

However, this may be the wrong way of putting the question. Historical linguistics rarely allows us to specify a sufficient trigger of change. Even though *ki* is more marked, or more complex, than *či*, a language may retain its *k*'s unchanged for millennia, while its close relative undergoes the natural sound change $k \rightarrow \check{c}$ before front vowels. It is not hard to see why this might be the case. For a property to be marked might mean no more than having a very small margin of extra complexity, insignificant next to massive positive evidence and easily swamped out by it in the acquisition process. Such marked linguistic features would be expected to be stable for long periods if they are well entrenched in the system. Optimization in such cases would act a weak force, of the type familiar from evolution, which manifests itself only over long periods of time, yet in a consistent direction (the reverse change $\check{c} \rightarrow k$ would never take place). Therefore, a marked system might well be quite stable until something happens to compromise its learnability.

To summarize, I have argued that the shift to VO is an endogenous structural optimization, and have discussed the causal factor behind it, in two versions. The first, weaker version is that it harmonizes the mixed Germanic system in which heads took their complements sometimes to the left and sometimes to the right, depending on a complex set of conditions. This analogical generalization (simplification) of the Germanic-internal structural constraints provides a motive for the change independently of Kayne's theory. The stronger version relies on Kayne's theory, which allows the change to be seen, in addition, as a reflex of a language-independent preference for VO.

In any case, from this perspective the shift to VO is not an isolated change affecting just the structure of the verb phrase, but part of a process which involves the other phrasal categories as well, viz. adjectives and nouns, and, at an earlier stage, lies behind the establishment of prepositions.

²⁰For other problems with the Kayne/Zwart account, see Gärtner and Steinbach 1994, Rohrbacher 1994.

Turning now to the second part of our proposal, the enabling cause, the claim is that the entrenchment of the old system was reduced in a subset of the Germanic languages by the introduction of *verb-fronting in embedded clauses*. This rendered the base position of the verb substantially more opaque in subordinate clauses (as it already was in main clauses because of V-to-C movement), and this made the OV base decisively harder to learn, in spite of the indirect evidence discussed above.

All Germanic languages with the exception of English front the finite verb to the vacant head of some functional projection; let us refer to this as *verb-fronting*.²¹ These languages fall into two types (Vikner 1991): *general verb-fronting languages*, where verb-fronting applies in all types of subordinate clauses, either obligatorily or optionally, and *restricted verb-fronting languages*, where it applies only in main clauses and complementizerless subordinate clauses (and, to a varying extent, also in “asserted” subordinate clauses even if they have complementizers). Yiddish and German respectively exemplify these two types:

- (8) a. A jidisch mejdl hot sech barimt, as efscher hundert mol **hot** men
 a Jewish girl has self bragged, that perhaps hundred times has one
 si schojn gebetn, as si **sol** chassene hobn. (Landmann 302)
 her already asked, that she should marry
 ‘A Jewish girl bragged that she had already been asked perhaps a
 hundred times to marry.’
- b. Ein jüdisches Mädle hat sich gerühmt, dass man sie vielleicht hundert
 Mal schon gebeten **hat**, dass sie heiraten **soll**.

In German, neither subordinate clause of (8b) allows fronting; thus German is a strict verb-fronting language which does not allow verb-fronting even in “asserted” subordinate clauses with a complementizer, such as the first subordinate clause in (8b) (whose equivalent would allow fronting in Swedish). It is evidence such as the second subordinate clause in (8a) which shows that Yiddish is a general verb-fronting language.

With respect to this property, the Germanic languages divide as follows:

- (9) a. Restricted verb-fronting: German, Dutch, Frisian, modern mainland
 Scandinavian²²

²¹I choose this non-committal term over “V2” because the fronted verb need not have exactly one constituent before it; the existence and derivation of V1 and V3 clauses will be discussed below.

²²Frisian and the Scandinavian languages allow verb-fronting rather freely in “asserted” embedded clauses (see Wechsler 1990 for an analysis of this phenomenon). Though not considered acceptable in standard German, this is frequently heard in colloquial speech, especially in causal and concessive clauses (*weil* “because”, *obwohl* “even though”), but not only in them. The following is a real-life apparent relative clause heard recently: ... *aus Gründen, die sind wirklich grösstenteils unabhängig von der Sprache*. “...for reasons which really are largely language-independent.” But see now Gärtner (MS) for evidence that relative clauses of this type should actually be analyzed as conjoined main clauses.

b. General verb-fronting:

1. Obligatory: Icelandic, older mainland Scandinavian, modern Yiddish
2. Optional: Faroese, Old English, old Yiddish

The generalization is that *the shift from a head-final base to a head-initial base took place in exactly those languages which developed general verb-fronting in embedded clauses*. Moreover, the languages that developed obligatory verb-fronting in embedded clauses also developed obligatory VO, and when languages developed optional embedded verb-fronting they also developed optional VO base order. And the rise of embedded verb-fronting in each language antedates the rise of VO.

That the Scandinavian languages had general verb-fronting at the time prior to the shift in VP headedness is uncontroversial. Although the modern continental Scandinavian literary languages have lost general verb-fronting, modern Icelandic retains it (Rögnavaldsson and Thráinsson 1990).

As for Yiddish, Santorini 1989 shows that it possessed verb-fronting as an option in embedded clauses from the earliest period of its attestation. The constituent that precedes it is in the early stages nearly always the subject.

It has been claimed by some authors that Yiddish is underlyingly not VO but OV, like German (Hall 1979, Geilfuss 1990, Vikner 1991). However, I am persuaded by Santorini's 1993 elegant statistical argument that modern Yiddish is predominantly VO, though it retains a marginal OV option as well: where the rightward movement transformations such as extraposition and scrambling can be seen at work independently, they operate at a low frequency, which remains fairly constant throughout the history of Yiddish; VO order on the other hand is very common and increases in frequency over time. Therefore rightward movement can only be a small factor contributing to VO order; the major source of it (96%, by Santorini's estimate) must be the base order.²³

The upshot is that there is indeed a causal connection between verb-fronting and the shift to VO, but it is not the one suggested by Stockwell and Lightfoot. It involves verb-fronting not in main clauses but in subordinate clauses, and not as a model for the new VO structure — this was provided by the head-initial structures in the rest of the system — but rather as a decrement of the evidence for OV: the fronting of finite verbs in general verb-fronting languages merely rendered the VO base sufficiently opaque to allow the optimal VO syntax to break through.

This shows that the position of the finite verb in subordinate clauses is the primary evidence for the OV base in the acquisition of languages like German.

²³Thus Yiddish is another instance of the “double-based” VP structure to be discussed for Old English in the next section.

We have seen that whenever general verb-fronting makes that evidence opaque, and only then, the languages give in to the analogical pressure to reorganize the VP as head-initial. Such a pattern of change would be incomprehensible unless data from subordinate clauses were accessible in the acquisition of syntax, contrary to the hypothesis of Degree-0 learnability.

To summarize: in this section I have argued that the shift from OV to VO is not, or at least not wholly, due to language contact, nor to the loss of inflectional morphology and of scrambling, nor to reanalysis of synchronically derived VO sequences. Instead, I proposed that it is a generalization of rightward complementation, that is, in essence an optimization, or “analogical change”. The change is motivated within the system, and arguably also grounded in a language-independent preference for head-complement order. It was activated in those languages in which generalized verb-fronting caused attrition of the embedded OV structures that provided the principal supporting data for the old system. The following section spells out this scenario and its synchronic underpinnings in more detail for Old English.

3 Old English

3.1 van Kemenade 1987

van Kemenade argued that Old English fits the analysis of German and Dutch proposed by den Besten 1983, in which verb-second and verb-first order are derived by fronting of the finite verb from the end of the sentence to C if that position is vacant, and the preceding Spec-CP position is filled by a Wh-element or topicalized phrase. She showed that Old English displays many of the word order patterns predicted by this analysis, and that important classes of apparent deviations from it can be accounted for by processes of extraposition and verb (projection) raising which can be independently demonstrated for the language.

However, van Kemenade also called attention to a number of syntactic generalizations which do not readily square with this “Dutch/German” model of Old English syntax. Subsequent efforts to explain these generalizations have modified one or both of the model’s two above-mentioned basic assumptions, by having the finite verb move to I rather than to C, and/or by letting topicalized constituents occupy a left-peripheral operator position distinct from that occupied by Wh-constituents (together with negation and certain adverbs such as *þa*, *þonne* ‘then’ and *ne* ‘not’, which pattern in the same way).

Perhaps the clearest piece of evidence that distinguishes Topicalization and Wh-movement is the positioning of pronouns. The regularities governing them (Allen 1977, 49, Mitchell *ó*3907) were recognized by van Kemenade as reflecting

their clitic status.²⁴ The generalization is that if there is a fronted Topic, clitic subject pronouns come directly after it, followed, either directly or at a distance, by the finite verb:

- (10) a. Hefonas **he** þurhfor mid his modes sceawinge (*CP* 99.23)
 heaven he traversed with his mind's vision
- b. & ealle þa oðre þeoda þe on Crecum wæron **he** to gafolgieldum
 and all the other peoples that with Greeks were he to tributaries
 gedyde (*Orosius* 124.6)
 made
 ‘and he made all the other peoples that were among the Greeks into tributaries’

But if the fronted constituent is a Wh-phrase (or *ne*, “not”, *þa*, *þonne* “then”), clitic subject pronouns are placed in third position, *after* the verb:

- (11) a. Hwæt sceal **ic** ðonne má secgean fram Sancte Iohanne (*BH*
 What shall I then more say from Saint John
 169.23)
 ‘What more then shall I say about St. John?’
- b. Ne mæge **we** awritan ne mid wordum asecgan ealle þa wundra
 not can we write nor with words say all those wonders
 (*ÆLS* 21.242)
 ‘We can neither write nor express with words all those wonders’
- c. þa ondwyrdon **hie** him tweolice (*Orosius* 156.2-3)
 then answered they him doubtfully
 ‘Then they answered him doubtfully’

Clitic object pronouns and adverbials may remain in the VP, but if fronted, they obey the same regularities as were just summarized for subject pronouns.

- (12) a. & se cyng **him** eac wel feoh sealde (*ASC*, A.D. 894)
 and the king him also much property gave
 ‘and in addition the king gave him much property’
- b. God **him** worhte þa reaf of fellum (*AHT^h* I.18)
 God them made then garments of skins
 ‘then God made garments of skin for them’
- c. ða nietenu **ðonne** beoð hwæthuguningas from eorðan áhæfen
 the beasts then are somewhat from earth raised
 (*CP* 155.15-16)

‘for beasts are to a certain extent raised from the earth’

²⁴Independent evidence for cliticization rather than syntactic movement is that only pronouns in Old English allow preposition stranding (Allen 1977, 53). See van Kemenade 1987, Ch. 4-5 for other arguments that Old English weak pronouns are “syntactic clitics” and for extensive discussion of preposition stranding.

Example (12a) also illustrates the availability of *verb-final main clauses* in Old English, which are scarcely found in German or any other old or modern Germanic language. The fact that these verb-final main clauses can have topics, as in (13b,c), but not Wh-phrases (Allen 1977, 48 ff.), is another diagnostic of the structural difference between topicalization and Wh-movement.

- (13) a. He þa his here on tu **todælde** (*Orosius* 116.16)
 he then his army in two divided
 ‘He then divided his army into two.’
- b. Her hæþne men ærest on Sceapige ofer winter **sætun** (*ASC*, A.D.
 here heathen men first in S. over winter sat
 855)
 ‘In this year heathen men first encamped in S. over the winter’
- c. & beforan Moyses & his folce he þone Raedan Sae on twelf
 and before Moses and his people he the Red Sea in twelve
 wegas adrigde (*Or.* 38.28)
 directions drained
 ‘and before Moses and his people he drained the Red Sea in twelve
 directions.’

Thus verb-fronting in Old English main clauses is sensitive to the same difference as cliticization is.

The most straightforward evidence that there are two positions is that they can co-occur:²⁵

- (14) a. witodlice þisum lēofan leorningnihte befæste sē hælend his
 truly to this beloved disciple entrusted the savior his
 modor
 mother
 (Aelfric,
 ‘Truly, it was this beloved disciple to whom the savior entrusted his mother.’
St. John 17)
- b. Soðlice swa micle lufe hæfde eall se ceasterwaru to him
 Indeed so much love had all the citizens for him
 (*Apollonius* 8)

Still another surprise for the “Dutch/German model” of Old English syntax is the apparently unrestricted availability of verb-fronting in subordinate clauses. As discussed above, German resists verb-fronting in embedded clauses, even in those which convey “assertions”. Old English not only allows verb-fronting freely after topicalization in in embedded assertions, as in (15),

²⁵Swan 1988 extensively documents the occurrence of multiple adverbials preceding the finite verb in Old English; I am grateful to Elizabeth Traugott for calling my attention to this study. See now also Swan 1994. The examples in (14) are cited from these studies.

- (15) a. Ic secge þæt behefe ic **eom** ge cingce & ealdormannum (*ÆColl.* 150)
 ‘I say that useful I am to the king and the chiefs.’
 b. Be ðam is awriten ðæt betera **beo** se geðyldlega wer ðonne se gielpna
 (*CP* 217.10)
 ‘Therefore it is written that better is the patient man than the boastful (one)’

but in principle allows verb-fronting in any type of subordinate clause (Pintzuk 1991, 1992):²⁶

- (16) a. þa wæs he sona mid godcundne onbryrdnysse innan monad,
 then was he at once with divine inspiration inwardly warned,
 þæt he **wearp** þæt sword **onweg** (*Bede* 38.19)
 so that he threw the sword away
 b. þæt wære swiðe gilplic dæd gif Crist **scute** ða **adún**
 that would have been a very proud deed if Christ had plunged then down
 (*ÆCHom* 170.2)
 c. ... þæt hie **noldan** leng heora hlaforda ne heora weras
 that they not wanted any more their husbands’ nor their men’s
 ræstgamanan secean (*BH* 173.16)
 bed-pleasure seek
 ‘(they turned so steadfastly to Christ’s love and faith) that they
 would no longer seek intercourse with their husbands or their men’

That these are indeed clauses with a fronted finite verb is guaranteed by the particles in (16a,b) and the infinitive in (16c), which reveal the finite verb’s underlying location.

3.2 Pintzuk 1991

Pintzuk proposes to account for Old English embedded verb-fronting by assuming that the finite verb in Old English normally does not move to C, but stays in I. This would make Old English a language of the Icelandic and Yiddish type, rather than of the German and Dutch type. In agreement with the analyses offered for Icelandic and Yiddish by Rögnvaldsson and Thráinsson 1990, Santorini 1989, and Diesing 1990, Pintzuk situates Old English subjects in Spec-VP,

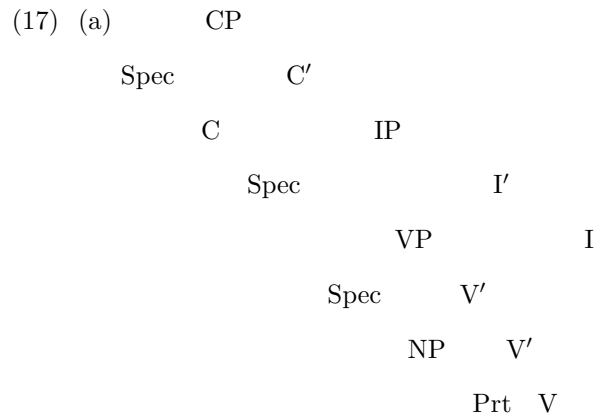
²⁶Pintzuk shows (1991, 346) that verb-fronting is twice as common in sentential complements as in relative clauses (62.6% vs. 29.1%). It is not surprising that a substantial proportion of verb-fronting subordinate clauses are of the “embedded assertion type”. See also Bacquet 1962 p. 403, 419, who notes the association of the relevant structures with the verb *secgan* ‘say’, and remarks that “l’utilisation de cette tournure donne au discours indirect un curieux air de style direct,” and compare Mitchell 1985, ¶ 1939-1943 on the “fluid boundary between true indirect speech and direct quotation” in Old English. The point however is that verb-fronting is amply attested outside of these contexts as well.

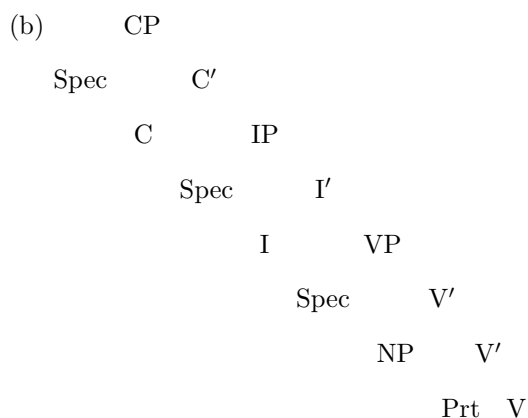
where they stay unless topicalized, in which case they move to Spec-IP, assumed to be an \bar{A} -position in these languages, rather than an A-position as elsewhere in Germanic. While topics move to Spec-IP, Wh-elements move to Spec-CP, from where they attract the verb to C. The abovementioned positional differences between topics and Wh-elements then make sense on the assumption that pronouns cliticize to the left edge of IP. In that position, they would follow verbs raised to C, but precede verbs that remain in I, as in case of ordinary topicalization.

Pintzuk actually assumes that other elements than Wh can also be placed in Spec-CP, although for some reason they don't attract the verb there. This possibility gives the "V3" construction already exemplified in (14).

The most far-reaching of Pintzuk's proposals, though, is that Old English represents a transitional system, whose diversity of syntactic patterns is derived from three distinct base-generated phrase structures in synchronic competition with each other.

Pintzuk's three base structures are derived by varying the direction of headedness at two levels of phrase structure. The IP is taken to be either head-final or head-initial:



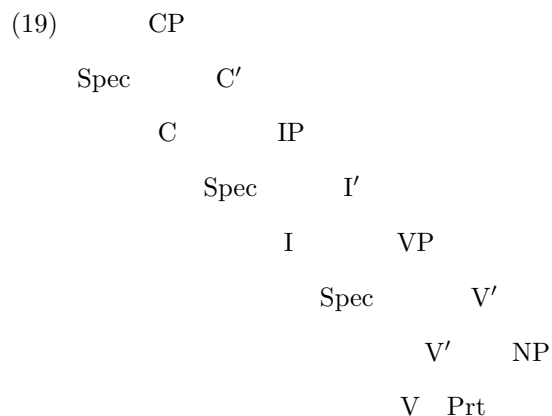


Pintzuk gives evidence that Old English had in addition an underlying VO phrase structure, from which sentences of the following type are derived:

- (18) he wolde adraefan ut anne aþeling
 he would drive out a prince

Since there is no particle movement in Old English, as Pintzuk shows on other grounds (p. 87 ff., p. 178 ff.) the relative order of the infinitive and the particle in such sentences must reflect the base order.

Pintzuk accounts for such sentences by allowing the head-initial version of the IP to dominate, in addition to a head-final VP, as in (17b), also a head-initial VP:



These three phrase structures are illustrated in (20a-c) by three grammatical renderings of “(that) the bishop wanted to lift up the child”:

- (20) a. (þæt) se biscop wolde þæt cild up aheafan. ... I_o [[... V]_{VP} t]_{VP}
 b. (þæt) se biscop þæt cild up aheafan wolde. ... [[... V]_{VP} t]_{VP} I_o
 c. (þæt) se biscop wolde aheafan up þæt cild. ... I_o [t [V ...]_{VP}]_{VP}

d. *(þæt) se biscop aheafan up þæt cild wolde. ... [t [V ...]_{VP}]_{VP}
 I_o

What about the fourth logically possible combination, of a head-final IP with head-initial VP? This phrase structure must be excluded somehow because sentences like (20d) are unattested in Old English. A consistently head-final phrase structure such as van Kemenade’s would explain their absence, but in Pintzuk’s multiple phrase structure system they constitute an unaccountable gap.

A second problem is that Pintzuk’s analysis, in consequence of situating the finite verb at I^o and treating the Spec-IP position that precedes it as the site of topicalization, is committed to the prediction that embedded clauses should allow topicalization, which is incorrect (van Kemenade MS).²⁷ Kroch & Taylor (MS) suggest that the analysis of Old English as an Infl-type V2 language could be made to work by sending the topic all the way to Spec-CP. The tensed V would still go to I^o (in Wh-contexts to C^o), so that the clitics can move to the left edge of IP as before.²⁸ The new difficulty now is how to ensure that C^o stays empty, and how to prevent Wh-expressions (and *þa*, *þonne*, *ne*) from preceding topics in sentences where both occur.²⁹

Both the problems raised by van Kemenade’s analysis and those raised by Pintzuk’s analysis can, I think, be solved by the following proposal, adapted with modifications from Kiparsky 1994.³⁰

3.3 A quasi-minimalist analysis

In Old English (as in the other early Germanic languages and in modern English) the specifier of CP is a focus position. Wh-phrases, certain demonstratives such as *þa*, and negation are intrinsic focus elements and obligatorily move to Spec-CP. This is also the site of the functional counterpart of “focus scrambling” in these languages. C^o is also obligatory in subordinate clauses, because sentences

²⁷Embedded assertions, such as those in (15), do allow Topicalization, but as with verb-fronting in such clauses, this is a separate phenomenon that does not vitiate the generalization at hand. The apparent cases of topicalization in subordinate clauses that were cited by Allen 1977, 52 can now be seen to involve cliticization.

²⁸According to Kroch & Taylor, another virtue of this analysis is that it eliminates the need for a rule which shifts clitics from the left edge of IP to second position if there is nothing before it. But on Halpern’s (1992) theory of cliticization (adopted by Taylor herself in Taylor 1994) this is no advantage because the movement in question (*prosodic inversion*) is part of UG.

²⁹As a reviewer notes, the latter is a problem for *any* analysis which assumes that Wh-expressions land before topics (Heycock and Santorini 1994).

³⁰That paper itself was inspired by Allen’s and van Kemenade’s pathbreaking work; the published version also benefited from Pintzuk’s (1990) criticisms of an earlier draft of it, as well as from her own analysis in that work.

must be turned into CPs in order to function as arguments and modifiers. Sentences with a fronted constituent of this type are therefore CPs, headed by a C^o position obligatorily filled by a complementizer or finite verb. Otherwise, where C^o is not required for the satisfaction of some principle of grammar, its presence is language-specifically determined. In strict V2 languages such as German, ordinary declarative main clauses are always CPs.³¹ In Old English, as in modern English, ordinary declarative main clauses are simply IPs. Topicalization and left dislocation is adjunction to the highest projection. Subjects originate in Spec-VP and (subject to conditions involving specificity) move to Spec-IP.

In subordinate clauses, the obligatory C^o position is usually filled by a lexical complementizer, which blocks V-to-C^o movement.³² Certain subordinate clauses may however lack a lexical complementizer; the empty C^o is then obligatorily filled by the finite verb, yielding the obligatory V1 pattern seen in bare concessive and conditional clauses, such as (21):

- (21) **Wære** se mon on swelcum lande swelce he wære þe hi ahte,
 were the man in whatever land whoever he were that them possessed,
 ðonne wære his wela & his weorðscipe mid him (*ConsPhil* 63.21)
 then were his wealth and his dignity with him
 ‘Whatever land the man who possessed them might be in, his wealth and his dignity would be with him.’

Conversely, Old English allows main clauses to have lexical complementizers. These main clause CPs are instantiated by yes/no questions introduced by *hwæðer*, which have subordinate clause word order, hence are verb-final (with a verb-fronting option as in all subordinate clauses):

- (22) a. Hwæðer ðu nu swelces auht wyrcan mæge? (Boethius,
 whether thou now such-Gen anything do can
ConsPhil 29.22)

‘How can you do anything like that?’

- b. Hwæþer he wolde þam forcuðestum monnum folgian? (*ibid.* 37.15)
 whether he would the wickedest men follow
 ‘Would he follow the most wicked men?’

Ordinary yes-no questions presumably have an empty operator in Spec-CP, as standardly assumed.

The reason why main clause questions, with an overt or covert Wh-operator, as well as sentences with the fronted adverbs *þa*, *þonne*, and *ne*, always have verb-second order is then that those elements are fronted to Spec-CP, hence

³¹In addition to the V2 constraint and the well-known complementarity of overt C and verb-fronting, this explains its main clauses with complementizers, such as questions with *ob*, and warnings with *dass*.

³²In relative and adverbial clauses with an overt Wh-element in Spec-CP, the complementizer may be phonologically empty, though syntactically present.

sentences containing them must be CPs, therefore contain a C^o , to which the finite Verb must move.³³ I assume that the reason pronominal clitics follow the finite verb in just those cases is that the focus element (*Wh*, *ba*, etc.) forms a phonological constituent with the verb or complementizer that follows it, and that the clitic comes in second position after this constituent.³⁴ Topics, being adjoined to the clause, do not form a prosodic constituent with the following verb.

When an NP or adverbial/PP is focused, for contrast or emphasis, it can be placed in Spec-CP, in which case the finite verb follows it in C. I propose this structure for the sentences in (10), as well as for cases like:

- (23) a. & wif **ic** lærde þæt hie heora weras lufedan (*BH* 185.22)
 and wives I taught that they their husbands should love
 ‘(landowners I taught to pay their taxes), and wives I taught to love their husbands.’
- b. Sopllice **unc** gecyþeþ ure Drihten Hæland Crist his mægen
 Surely us two will manifest our Lord Savior Christ his power
 (*BH* 189.3)
 ‘Surely our Lord Jesus Christ will manifest to us his power.’
- c. Forðon **we** sceolan mid ealle mod & mægene to Gode gecyrran
 therefore we must with all mind and power to God turn
 (*BH* 97.32)
 ‘That is why we must turn to God with all our mind and power.’

General verb-fronting in embedded clauses is V-to-I raising, as in Pintzuk’s analysis. Unrestricted topicalization in embedded clauses is however excluded because Spec-IP is normally reserved for subjects, and adjunction is limited to main clauses (which has to be said anyway on any account in order to block topicalization to CP in embedded clauses).

Adjunction of non-focused adverbs and PPs to IP yields V3 in main clauses:³⁵

³³A reviewer suggests that *ne*, being a proclitic, may not itself be in Spec-CP but rather agreeing with an empty operator in that position.

³⁴This is a variant of van Kemenade’s idea that operators have a special semantic relationship to the element that follows them, which blocks clitics from intervening between them, except that I assume the special relationship to be a matter of phonological phrasing, perhaps connected to the stress which falls on the focused constituent. This is more in line with the theory of Halpern (1992), in which clitics are positioned by the interaction of independent syntactic and phonological constraints. They are syntactically associated with edges of maximal projections, and phonologically required to lean to their left or right on a prosodic constituent, whose size is parametrically specifiable as a phonological word or a phonological phrase. The phonological constraint is dominant, so when necessary for the satisfaction of the phonological requirement, the clitic is moved by “prosodic inversion” at PF to a position after the first prosodic constituent of its syntactic host (or, if it is a proclitic, before the last).

³⁵Pintzuk (1991, 101 ff.) argues that adjunction is limited to temporal adverb phrases, as

- (24) a. & þy ilcan geara Tatwine **was** gehealgod to arcebisċ (*ASC*,
and the same year Tatwin was consecrated as archbishop
731.6)
- b. & for Godes naman ándetnesse Petrus **wæs** on rode gefæstnod
and for God's name's confession Peter was on cross fastened
(*BH* 173.3)

‘And for confessing God’s name Peter was fastened on the cross.’

In addition, V1 clauses such as (25) occur:

- (25) a. **Wæs** on ða tid heora heretoga and latteow Ambrosius haten
was in that time their duke and general A. called
(*Bede* 1.16)
- ‘Their duke and general at that time was called Ambrosius.’
- b. **Wæs** eac munuc se ylca bysceop Aidan; **wæs** he sended of þam
was also monk the same bishop Aidan; was he sent from the
ealande and of þam mynstre þe Hii is nemned (*Bede* 3.3)
island and from the monastery that Hii is named
‘The same bishop Aidan was also a monk; he was sent from the island
and the monastery called Hii.’
- c. *Comon* he of þrim folcum ðam strangestan Germanie (*Bede*
came they from three peoples the mightiest of Germany
1.14)
- ‘They came from three of the mightiest nations of Germany.’
- d. **Hæfde** se cyning his fierd on tu tonumen (*ASC*, A.D. 893)
had the king his army in two divided
‘The king had divided his army in two’

These are usually taken to be parallel to V2 clauses, but without an overt preposed constituent, and I will follow this analysis.³⁶ In so far as they involve unaccusative and impersonal (subjectless) verbs (as in (25a,b)), they can be considered IPs in which Spec-IP remains vacant, and the subject (when there is one) remains in its original Spec-VP position. It is suggestive that V1 in declarative main clauses occurs in exactly the subset of Germanic languages that have general V2 (Vikner 1991, 98), viz. Yiddish (Diesing 1990, 56), Icelandic

in (24a). Unquestionably that is the most common case, but examples like (24b) suggest that it is more generally available for adverbs and PPs. As in modern English, NPs can only be left dislocated in such cases, though the left dislocated element arguably lands in the same adjunction position (Kiparsky 1995, pp. 143-144).

³⁶It has been suggested that the verb is preceded by an empty element of some sort, perhaps an adverbial (Diesing 1990).

(Sigurðsson 1990), and Old English.³⁷ Nevertheless, an analysis as CPs with an empty operator in Spec-CP is plausible.

Just in case Spec-IP remains vacant because there is no external argument, it can be filled by another constituent; this is shown by the fact that topicalization in embedded clauses occurs with unaccusatives and impersonal verbs (van Kemenade MS). This type of V2 clause has a non-focused preposed constituent, such as a “locative subject”:

- (26) a. In ðeosse abbudissan mynstre wæs sum broðor syndriglice mid
in this abbess’ monastery was a certain brother especially with
godcundre gife gemæred (*Bede* 4.24)
divine grace celebrated
‘In this abbess’ monastery there was a brother especially noted for
his divine grace’
- b. æfter his gebede **he** ahof þæt cild up
after his prayer he lifted the child up
(*AHTh* II.28, cited from van Kemenade 1987, 110)

If the main clause is a CP, adjunction is also possible. This is the analysis of sentences where *þa*, *þonne*, or *ne*, followed as usual by the finite verb in C (and then by the clitic pronoun if there is one) is preceded by an adjoined adverbial, as in (27a-c):

- (27) a. þeah se lærow ðis eall smealice & openlice gecyde, ne **forstent**
though the teacher this all carefully and openly tells, not avails
it him noht (*CP* 163.18)
it him naught
‘Though the teacher tells all this carefully and openly, it avails him
nought.’
- b. þeah þe we þas þing cweþe, ne **tellað** we synne weosan
though that we these things say, not count we sin to be
gesinscipe (*Bede* 1.18, p. 82)
wedlock
‘Though we say these things, we do not count wedlock as sin.’
- c. Gif þonne ge lifes bæð forhycgað, ne **magon** ge ænige þinga lifes
if then you life’s bath despise, not may you in any thing life’s
hlaf onfon. (*Bede* 2.5)
bread receive
‘(But) if you despise the bath of life, you may not in any wise receive
the bread of life.’

³⁷The colloquial German verb-initial construction cited by Vikner has a rather different character; it may be phonological rather than syntactic in nature, involving the elision of an unstressed pronoun before the verb and not syntactic V1: Diesing’s Yiddish example *Bin ikh geblibn aleyh*. “So I was left alone.” does not correspond to German (**Bin ich allein geblieben*.); German only has *Bin allein geblieben*.

The same class of adverbials can precede Wh-elements, which pattern positionally with *ne* in other respects as well, as we have seen.

- (28) a. Gif micel feornis siiðfætēs betweohn ligeð, þætte biscoopes æþelice
 if much distance bishops between lies so that bishops easily
 cuman ne magon, **hwæðer** mot biscop halgad beon buton
 come not may, whether may bishop consecrated be without
 oðera biscoopa ondweardnesse? (*Bede* 1.27)
 other bishop's presence
 'If the distance between bishops is great, so that bishops may not
 easily travel, may a bishop be consecrated without the presence of
 other bishops?'
- b. Gif hwa nu bið mid hwelcum welum geweorþod & mid hwelcum
 if who now be with whatever riches endowed and with whatever
 deorwyrþum æhtum gegyrewod, **hu** ne belimpþ se
 valuable possessions adorned, whether not belongs the
 weorþscipe þonne to þam þe hine geweorþað? (*ConsPhil* 96.12)
 glory then to him that him adorned?
 'Now if anyone is endowed with all riches and adorned with all valu-
 able possessions, does not the glory then belong to him who adorned
 him?'
- c. Gif þonne for micle arwyrðnesse **hwylc** mon ne geþyrstgað
 if then for much veneration which man not ventures
 onfón, se is to herienne. (*Bede* 1.27)
 to receive, he is to praise
 'Whoever, out of great veneration, does not venture to receive, then,
 is to be praised.'³⁸

The upshot is that Old English is not really a strict “V2 language” in the sense that German and Old Icelandic are.³⁹ It permits both V1 and V3 declarative main clauses, and its V2 clauses are of several types. There are at least two configurations in which the finite verb sits in second position: (1) a focused element in Spec-CP, with the verb in C position after it (e.g. (10), (11), (23)), (2) a subject (or, in sentences without external arguments, some other constituent, as in (26)) in Spec-IP, with the verb in I position after it, e.g. (12). Adjoining an adverbial or PP to these two structures in turn yields two very different types of V3 order, respectively illustrated by (27)-(28) and (24).

The puzzling asymmetry revealed by Pintzuk's theory, where one of the four predicted phrase structure combinations (left-headed VP with right-headed IP)

³⁸This sentence looks like a case of topicalization in an embedded clause, which is generally excluded in Old English, as I assume on Kemenade's evidence. However, it seems to be an adjoined subordinate clause of the archaic Indo-European type, and topicalization is always permitted in them (Kiparsky 1995).

³⁹See Swan 1994 for a similar assessment. However, genuine V2 dialects are reported to occur in certain Middle English writers, notably Chaucer (Kemenade MS.), and in northern dialects (Kroch & Taylor MS).

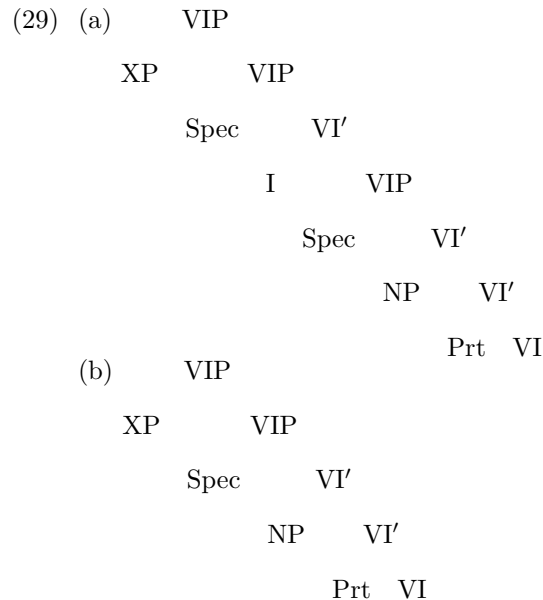
is systematically excluded, can now be resolved. The solution has to do with a more general point to which students of modern Verb-final Germanic languages have repeatedly drawn attention, namely that *there is no rightward V-to-I movement* (Reuland 1990, Haider 1993, Zwart 1993, 68, van Gelderen 1993, Ackema, Neeleman, and Weerman 1993). The argument is basically that the putative V and I positions are never distinct in OV languages. Zwart observes that, if complement clauses are generated in the same position as nominal complements (as would be required if we take seriously the fact that they bear the same kinds of thematic roles), V-to-I raising would require them to be extraposed to the right periphery of the IP. One problem is that extraposition in this case would have to be obligatory, whereas it is otherwise optional. Moreover, the extraposed clauses might be expected to be islands. A compelling argument due to Haider is that obligatory extraposition is incompatible with coordination data, which require complements to be inside VP. Both Zwart and Haider also point out that the inseparability of the verb cluster and the particle+verb combination is hard to explain on the assumption that the finite verb is raised rightwards to I. van Gelderen 1993 makes a similar point, and argues that it holds for Old English as well. Another strong argument is Reuland's demonstration that V-to-I raising predicts non-existing readings for adverbs on the assumption that their scope is determined by constituent structure.

In those cases where there is no V-to-I movement in the syntax, it must either take place somewhere else — in the lexicon or at LF — or be eliminated altogether. The solution suggested by Reuland is that V and I merge in the lexicon and constitute a complex category V/I in the syntax. Zwart takes IP to be left-headed in Dutch (a corollary of Kayne's hypothesis that heads universally precede their complements in the base) and posits covert movement of V to I at LF, in line with the minimalist program of Chomsky 1992. Ackema, Neeleman, and Weerman contend, still more radically, that the syntactic category I does not exist, and suggest the general constraint that each lexical category has only one functional projection in the syntax (see also Weerman 1989, p. 81). van Gelderen, discussing the question in comparative and historical perspective, reaches the conclusion that the presence of I (her T) as a syntactic category is a site of parametric variation and change; she dates the rise of this category in English to the end of the Middle English period (ca. 1380).

If Kayne's hypothesis is valid for functional categories, then both the lack of rightward V-to-I movement and the mysterious directional asymmetry revealed by Pintzuk's analysis can be explained. There is then no such thing as an I-final language, and the unfronted finite verb sits *in situ* at the end of its VP. The various ways of realizing this idea differ in important ways, as we have just seen, but each of them excludes the missing fourth pattern (20d) in principle, and translates Pintzuk's I-final/I-medial distinction into a structural difference which does *not* cross-classify with the VO/OV distinction with any unwanted results. Depending on which implementation we adopt, the change

in English would then be either from covert (LF) V-to-I movement to overt V-to-I movement, or from lexical V/I merger to syntactic V-to-I movement, or simply the introduction of I as a syntactic category. Though the choice between them would of course have to be decided on more general grounds, the history of English syntax might contribute evidence towards its resolution. Following van Gelderen 1993, I will assume that the word order shift is connected with the rise of the category I (Aux, her T), which she argues became obligatory around 1380, on the evidence of the emergence of *to* as an I, *do*-support, and other related syntactic innovations at this time.

I shall here adopt the position that the finite verb emerges fully inflected from the lexicon, as assumed in Lexical Phonology and in the Minimalist Program, and that it constitutes a complex category V/I à la Reuland, that is, simultaneously an I and a V. A bare declarative clause will then be simultaneously an IP and a VP, and so on. I will assume that Old English at the stage under discussion is a transitional system in which I functions either as a syntactic category (as in modern English) or a purely morphological category of the verb (as in modern German). The finite verb is inflected in the lexicon and as such intrinsically both an I and a V (denoted by VI in the trees below).



In IPs, the finite verb obligatorily moves to I; in VIPs it remains in situ: this is the analysis of verb-final sentences such as (13). Since CP is also optional (unless required for subordination or focusing as discussed above), both IP and VIP can be main clauses by themselves, or as complements to C. I return to this point shortly.

The relevant options here are thus: (1) right-headed vs. left-headed VP, and

(2) presence or absence of a separate syntactic I projection (which by assumption is necessarily left-headed like all functional projections). The four structures in (20) accordingly receive the following analysis:

- (30) a. (þæt) se biscop wolde þæt cild up aheafan. ... I_o [[... V]_{VP} t]_{VP}
 b. (þæt) se biscop þæt cild up aheafan wolde. ... [[... V]_{VP} V]_{VP}
 c. (þæt) se biscop wolde aheafan up þæt cild. ... I_o [t [V ...]_{VP}]_{VP}
 ... [V [V ...]_{VP}]_{VP}
 d. *(þæt) se biscop aheafan up þæt cild wolde. [[V ...]_{VP} V]_{VP}

The fourth, nonexistent word order type is excluded because it has an inconsistent syntax, namely a right-branching VP nested in a left-branching VP.⁴⁰ Given our assumption that the left-branching and right-branching structures reflect competing grammars, such nesting would require the two grammars to be in force simultaneously, which should be impossible. In effect, this approach entails that variation between the two word orders is a matter of *code-switching*, only in this case between minimally distinct grammars. The expectation is that competing grammatical systems could be able to alternate under the same constraints that govern code-switching.

Although Pintzuk adopts Kroch’s approach to syntactic variation as grammar competition, her analysis of Old English syntax is inconsistent with the explanation just proposed. For her, the alternative word orders in (30) reflect variation in the headedness of both VP and IP, and these would have to be independent in order for the three occurring patterns to be derived. The missing fourth pattern (30d) (then analyzed as (20d)) would have to be excluded by an arbitrary constraint.

We now have an empirical argument favoring Kroch’s competing grammars approach over two alternative accounts of syntactic variation, respectively invoking syntactic movement rules and underspecified phrase structures.

Movement rules formally rich enough to derive the variant word orders from a single base structure are no longer available in current syntactic theory, but in traditional transformational grammar VO orders could be derived from an OV base (or vice versa) by “scrambling” of NPs and PPs/adverbs, plus some form of particle movement.⁴¹ Against this approach, we have already cited Pintzuk’s and Santorini’s arguments that the “catastrophic” OV to VO reanalyses which it entails are not supported by the historical evidence, and that the particle movement rules which it requires lack independent motivation in the synchronic grammars of Old English and Yiddish, and in fact would generate ungrammatical sentences in both languages. The results of this section provide additional

⁴⁰Note that there are still four structures but that two of them coincide in (20c).

⁴¹Verb movement alone was also proposed in some analyses, but it would clearly be both insufficient and undesirable on theoretical grounds.

empirical arguments against it. First, it does not generalize to the IP/VIP variation in the status of main clauses during the transitional periods after the introduction of the new functional category I. Specifically, it offers no account for the ungrammaticality of the pattern (30d). If the descriptive generalization is that Infl occurs only with left-headed VP, as I have argued, this distribution cannot be characterized by any movement rules.

Granting that the VO/OV variation is a matter of alternative base structures, we could still attempt to avoid the assumption of competing grammars in a different way, by assuming that OV and VO are generated within a single grammar in which the orientation of the VP is *unspecified*. In perhaps the earliest study of free word order in generative grammar, Staal (1967) suggested that the phrase structures of languages like Sanskrit are unspecified for constituent order, like rotating mobiles. The move from phrase structure rules to licensing principles in modern syntax can provide a formal basis for Staal's suggestion; the same is true of GPSG/HPSG's separation of constituency and linear ordering into separate constraints. For Germanic, the idea of unspecified base order has been adopted by Rögnavaldsson (1992), who proposes that the VO/OV parameter is unspecified in Icelandic. In terms of a parametric approach, the idea could be articulated in several ways, with different theoretical and empirical consequences. If we assume that each parameter is supplied with a default value which will be acquired in the absence of positive evidence, then learners would have to leave parameter values unspecified if there is conflicting positive evidence for both settings. On the other hand, if we assume that there is no default value, then a parameter could be left unspecified in the absence of positive evidence for either setting.⁴² In that case, we would predict that the grammar resulting from an unspecified VO/OV parameter is the simplest to acquire; on the view that a default value is assigned, it would be on the contrary the most complex alternative, requiring the most data to acquire.

Both these variants of the underspecification solution to the VO/OV alternation are contradicted by our findings.⁴³ Consider again Old English clause structure as schematized in (30). The reason why pattern (30d) is ungrammatical is that nested VPs must have the same orientation. As noted above, this follows (on certain plausible assumptions) if left and right-headed VPs reflect distinct grammatical systems. But an analysis that posits a single grammar with unspecified headedness would again require an otherwise unmotivated constraint to exclude a left-headed VP from being embedded in a right-headed VP.

In conclusion, by assuming the possibility of bare VIP without a syntactically separate I projection, and that clauses are not CPs unless they must be, we

⁴²For example, in the case at hand, if the accessible data contained only main clauses with simple finite verbs in Infl or Comp, and indirect clues from particles, negation and other phenomena were for some reason unavailable.

⁴³Whether the theory might be appropriate for other cases of "free word order" is a separate question which I leave open here.

predict the possibility of verb-final main clauses such as those cited earlier in (7) and (13), and their properties. As the analysis also predicts, such verb-final main clauses cannot be introduced by focus elements such as *Wh*, *þa*, and *ne*, which force a CP. However, they should be capable of adjoining topics, a possibility confirmed by examples like (13b,c).

Let us now take another look at the OV to VO change in the light of the analysis of Old English clause structure that we have arrived at in this section.

4 The changes behind the variation

The competing grammars that we found in Old English are the result of two major syntactic changes: the introduction of I, and the fixing of left-headed VP. Both innovations fall in with other changes as part of two long-term tendencies: the development of functional categories, including complementizers (Kiparsky 1995), articles, and prepositions, and the imposition of uniform rightward complementation. Starting from an OV structure without I, the first properly Old English major syntactic innovation was the introduction of I, which then made possible the analogical generalization of left-headed VP, as discussed in section 1.4, yielding VO with I, which is the modern English structure. The large repertoire of verb positions in Old English, compared to that of most Germanic languages, results in part from the coexistence of each stage in this development as a synchronic grammar of the language.

The same relationship between the rise of I and the change of OV to VO is found in several other Germanic languages. Yiddish seems to have gone through the same development as English, plus some additional changes of its own. According to Santorini (1989), the modern Yiddish system evolved in three stages: (1) general embedded verb-fronting became possible, and VO is introduced as an option, (2) topicalization became possible in subordinate clauses (this is where it differs from English), (3) embedded verb-fronting became obligatory, at about the same time as OV phrase structure was lost. The Scandinavian languages seem to have gone through the same stages as well,⁴⁴ though the mainland Scandinavian languages subsequently lost V-to-I movement as well as topicalization in subordinate clauses.

Looking beyond Germanic, a parallel development is plausible for a number of other Indo-European and Finno-Ugric languages which shifted from OV to VO syntax. The Western Finno-Ugric languages, including Finnish, have both a syntactic I projection and VO order; the Eastern Finno-Ugric languages, though poorly studied, are uncontroversially OV, and in this respect can be assumed to retain the original syntax of the family.

⁴⁴According to Falk (1993, 162) Swedish underwent “a development from OV to VO within a basically I^o-medial system”.

Interestingly, a recent study of classical Greek suggests that the rise of the IP category is implicated in the OV to VO word order shift, and that functional categories (and IP in particular) are left-headed. Although classical Greek word order is “free”, Taylor (1994) shows that is predominantly OV in the Homeric dialect, and that it changes to VO by the time of the Koine dialect of the New Testament. She concludes that the Homeric dialect had an OV base, which had changed to VO by the time of the Koine, via intervening stages which on her analysis involve OV and VO in competition, such as that represented by Herodotus. The occurring surface word orders, under these assumptions, result from rightward and leftward movement processes which apply at a constant rate throughout the period under consideration. Concurrently with the shift from underlying OV to VO, Taylor finds a change in the positioning of clitics. In Homeric Greek, clitics appear at the left edge of IP; in New Testament Greek they appear at the left edge of VP; Herodotus again shows an intermediate stage which she argues is a mix of both systems. The question is why the final versus medial position of the verb should correlate with whether the domain of cliticization is IP or VP. If we assume as before that IP is always left-headed, this mysterious co-variation is at once explained. For then Homeric Greek, like any SOV language, must lack a separate category I in the syntax, so that it has no IP distinct from VP. The inflected verb in Homeric Greek thus belongs, as in early Germanic, to the composite category VI, the specifier of whose clausal projection is occupied by the subject. In post-Homeric Greek, I becomes a separate category; in fact, the change to VO order studied by Taylor could be seen precisely as the result of the rise of a functional I node which hosts the finite verb.⁴⁵ The apparent change in cliticization is then a direct consequence of the emergence of IP. We can maintain exactly the same cliticization rule for all stages of ancient Greek, namely that clitics go to the beginning of the lowest containing maximal projection, i.e. VIP in Homer, and VP in New Testament Greek:

- (31) a. CP[VIP[*clitic* ...]] (Homeric)
 b. CP[IP[VP[*clitic* ...]]] (New Testament)

Thus there is convergent evidence from outside Germanic that the OV to VO shift is causally connected to the rise of I as a syntactic head. The latter change itself may be a compensatory enrichment of the inventory of functional categories, replacing the receding aspectual and modal inflections of the Indo-European verb. Within Germanic, it is suggestive that the languages which did not develop IPs, i.e. continental Germanic, also do not have a syntactic class of intrinsically finite verbs, such as the English modals, which are presumably members of the functional category I.

⁴⁵Quite likely a left-headed verb phrase emerged concurrently with the I-projection, just as it did in Germanic. In that case, we would have *three* structures for some period of post-Homeric Greek, until both IP and the VO base became categorical, as in English.

In conclusion, I have argued that the OV to VO shift in Germanic is a generalization of rightward complementation, resulting in a fully harmonic head-complement structure in all categories. This change was able to penetrate in those Germanic languages in which the subordinate clause evidence for OV was weak due to application of general embedded verb fronting, itself due to the rise of the category I. The analysis is in line with a view of syntactic change as endogenous optimization, a fundamentally analogical process targeting certain preferred structural properties and enabled by opacity. It supports the idea that the site of synchronic variation is in coexisting grammars and is inconsistent with Lightfoot's degree zero learnability hypothesis, and with pure reanalysis theories in general. Support was also found for Kayne's directionality hypothesis, at least in the domain of functional categories.

Bibliography

- ACKEMA, P., A. NEELEMAN, AND FRED WEERMAN. 1993. "Deriving functional projections." *NELS* 23, Vol. 1.
- ALLEN, CYNTHIA. 1977. *Topics in diachronic syntax*. Ph.D. dissertation, University of Massachusetts, Amherst, Mass.
- ALLEN, CYNTHIA. 1986. "Reconsidering the history of *like*." *Journal of Linguistics* 22.375-409.
- ALLEN, CYNTHIA. 1992. "Variation and syntactic change: the English 'impersonal' constructions." *Second diachronic generative syntax workshop, Preprints*. Institute for Research in Cognitive Science, University of Pennsylvania.
- BACQUET, PAUL. 1962. *La structure de la phrase verbale à l'époque Alfrédienne*. Université de Strasbourg: Faculté des lettres.
- BEAN, M.C. 1983. *The development of word order patterns in Old English*. London: Croom Helm.
- BEHAGHEL, PAUL. 1924. *Deutsche Syntax: eine geschichtliche Darstellung*. Vol. 2. Heidelberg: Winter.
- BEHAGHEL, PAUL. 1932. *Deutsche Syntax: eine geschichtliche Darstellung*. Vol. 4. Heidelberg: Winter.
- BOSER, K., B. LUST, L. SANTELMANN, J. WHITMAN. 1992. "The syntax of CP and V2 in early child Germanic." *NELS*. 22.
- BIERWISCH, MANFRED. 1983. "Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In W. Motsch and R. Ruzicka (eds.) *Untersuchungen zur Semantik*. Berlin: Akademie-Verlag.
- BURRIDGE, KATE. 1993. *Syntactic change in Germanic: aspects of language change in Germanic, with particular reference to Middle Dutch*. Amsterdam: Benjamins.
- CANALE, WILLIAM MICHAEL. 1978. *Word order change in Old English: base reanalysis in generative grammar*. Ph.D. dissertation, University of Toronto.
- CHOMSKY, NOAM. 1993. "A minimalist program for linguistic theory." In K. Hale and S.J. Keyser (eds.) *The view from Building 20: Essays in honor of Sylvain Bromberger*. Cambridge, Mass: MIT Press.
- DANCHEV, A. 1991. "Language change typology and some aspects of the SVO development in English." In Dieter Kastovsky (ed.), *Historical English syntax*. Berlin-New York: Mouton de Gruyter.
- DENNISON, DAVID. 1986. "On word order in Old English." *Dutch Quarterly Review* 16.277-95.
- DENNISON, DAVID. 1993. *English historical syntax*. London: Longmans.
- DERBYSHIRE, D., AND G.K. PULLUM 1981. "Object initial languages." *IJAL* 47.

- DIESING, MOLLY. 1990. "Verb movement and the subject position in Yiddish." *Natural Language and Linguistic Theory* 8.41-81.
- DRYER, M.S.. 1990. "Object-verb and adjective-noun order." *Lingua* 74.185-217.
- FALK, CECILIA. 1993. *Non-referential subjects in the history of Swedish*. Ph.D. Dissertation, Department of Scandinavian Languages, University of Lund.
- GÄRTNER, HANS-MARTIN. MS. "Are there V2 relative clauses in German?" Max-Planck-Gesellschaft, Arbeitsstelle Strukturelle Grammatik, Berlin.
- GÄRTNER, HANS-MARTIN, AND MARKUS STEINBACH. 1994. "Economy, verb-second, and the *SVO-SOV* distinction." *WPSS* 53.1-59.
- GEILFUSS, JOCHEN. 1990. "Jiddisch als SOV-Sprache." *Zeitschrift für Sprachwissenschaft* 9.
- GERRITSEN, MARINEL. 1984. "Divergent word-order developments in Germanic languages: a description and tentative explanation" In Jacek Fisiak, *Historical Syntax*. Berlin: de Gruyter.
- HABER, LYNN. 1975. "The muzzy theory." *CLS* 11.
- HAIDER, HUBERT. 1993. *Deutsche Syntax — generativ*. Tübingen: Narr.
- HALL, BEATRICE. 1979. "Accounting for Yiddish word order, or what's a nice NP like you doing in a place like this?" In Jürgen Meisel and Martin D. Pam, *Linear order and generative theory*. Amsterdam: Benjamins.
- HALPERN, AARON. 1992. *Topics in the placement and morphology of clitics*. Ph.D. dissertation, Stanford University, Stanford, California.
- HAWKINS, JOHN. 1983. *Word order universals*. New York: Academic Press.
- HILTUNEN, RISTO. 1983. *The decline of the prefixes and the beginnings of the English phrasal verb*. Turku: Turun Yliopisto.
- HOCK, HANS HEINRICH. 1985. "Pronoun fronting and the notion 'verb-second' position in Beowulf." In J. T. Faarlund (ed.) *Germanic Linguistics: Paper from a symposium at the University of Chicago*. Bloomington, Ind.: Indiana University Linguistics Club.
- HOCK, HANS HEINRICH. 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.
- HOLMBERG, ANDERS. 1985. "Icelandic word order and binary branching." *Nordic Journal of Linguistics* 8.161-195.
- JESPERSEN, OTTO. 1938. *The growth and structure of the English language*⁹. Garden City: Doubleday.
- KAYNE, RICHARD. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- VAN KEMENADE, ANS. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.
- VAN KEMENADE, ANS. MS. "V2 and embedded topicalization in Old and Middle English." University of Amsterdam.

- KIPARSKY, PAUL. 1994. "Indo-European origins of Germanic syntax." In Adrian Battye and Ian Roberts (ed.).
- KIPARSKY, PAUL. MS. "The rise of positional licensing." To appear in Ans van Kemenade and Nigel Vincent (eds.). *Proceedings of the Third Diachronic Generative Syntax Conference*, University of Amsterdam.
- KOHONEN, VILJO. 1978. *On the development of English word order in religious prose around 1000 and 1200 A.D.*. Åbo: Åbo Akademi.
- KOOPMAN, HILDA. 1984. *The syntax of verbs*. Dordrecht: Foris.
- KROCH, ANTHONY, AND ANN TAYLOR. MS. "The syntax of verb movement in Middle English: dialect variation and language contact." University of Pennsylvania.
- KROCH, ANTHONY. 1989. "Reflexes of grammar in patterns of language change." *Language variation and change* 1.199-244.
- LABOV, WILLIAM. 1989. "The child as linguistic historian." *Language variation and change* 1.85-94.
- LANDMANN, SALCIA. 1988. *Jiddisch. Das Abenteuer einer Sprache*. Frankfurt: Ullstein.
- LIGHTFOOT, DAVID. 1979. *Principles of diachronic syntax*. Cambridge: University Press.
- LIGHTFOOT, DAVID. 1982. *The language lottery*. Cambridge, Mass: MIT Press.
- LIGHTFOOT, DAVID. 1987. Review of Keyser and O'Neil, *Lg.* 63.151-155.
- LIGHTFOOT, DAVID. 1991. *How to set parameters*. Cambridge, Mass.: MIT Press.
- LIGHTFOOT, DAVID. 1993. "Why UG needs a learning theory: triggering verb movement." in C. Jones (ed.) *Historical linguistics: problems and perspectives*. London: Longman.
- MALING, JOAN. 1990. "Inversion in embedded clauses in Modern Icelandic." In Joan Maling and Annie Zaenen (eds.) *Syntax and Semantics, Vol. 24: Modern Icelandic Syntax*. San Diego: Academic Press.
- MITCHELL, BRUCE. 1985. *Old English Syntax*. Oxford: Clarendon Press.
- NEELEMAN, AD. 1994. *Complex predicates*. Utrecht: Onderzoeksinstituut voor Taal en Spraak.
- NEELEMAN, AD AND FRED WEERMAN. 1993. "Case theory and the diachrony of complex predicates in Dutch." *Folia Linguistica Historica* 13:189-217.
- PAUL, HERMANN AND E. SCHMITT. 1950. *Mittelhochdeutsche Grammatik*. Tübingen: Niemeyer.
- PINTZUK, SUSAN. 1991. *Phrase structures in competition: variation and change in Old English word order*. Ph.D. dissertation, University of Pennsylvania.
- PINTZUK, SUSAN. 1992. "Phrase structure variation in Old English." *Second diachronic generative syntax workshop, Preprints*. Institute for Research in Cognitive Science, University of Pennsylvania.

- PLATZACK, CHRISTER. 1990. "A grammar without functional categories: a syntactic study of early Swedish child language." *Nordic Journal of Linguistics* 13:107-126.
- POEPEL, DAVID. 1993. "The full competence hypothesis of clause structure in early German." *Language* 69:1-33.
- ROBERTS, IAN. 1993. *Verbs and diachronic syntax*. Dordrecht: Kluwer.
- ROBERTS, IAN. 1994. "Cliticisation and word order change in the history of English." Talk Presented at the the Third Diachronic Generative Syntax Conference, University of Amsterdam.
- ROHRBACHER, BERNHARD. 1994. "Notes on the antisymmetry of syntax." *University of Pennsylvania Working Papers in Linguistics*, 1.
- RÖGNVALDSSON, EIRÍKUR AND HÖSKULDUR THRÁINSSON. 1990. "On Icelandic word order once more." In Joan Maling and Annie Zaenen (eds.) *Syntax and Semantics, Vol. 24: Modern Icelandic Syntax*. San Diego: Academic Press.
- RÖGNVALDSSON, EIRÍKUR . 1992. "Word order changes in the VP in Old Icelandic." *Second diachronic generative syntax workshop, Preprints*. Institute for Research in Cognitive Science, University of Pennsylvania. Revised as "Word order variation in the VP in Old Icelandic". MS. Reykjavik 1992.
- SANTORINI, BEATRICE . 1989. *The generalization of the verb second constraint in the history of Yiddish*. Ph.D. dissertation, University of Pennsylvania.
- SANTORINI, BEATRICE. 1992. "Variation and change in Yiddish subordinate clauses." *NLLT* 10.595-640.
- SANTORINI, BEATRICE. 1993a. "Jiddisch als SVO-Sprache." In Werner Abraham and Josef Bayer (eds.), *Dialektsyntax*. Special issue of *Linguistische Berichte*.
- SANTORINI, BEATRICE. 1993b. "The rate of phrase structure change in the history of Yiddish." *Language variation and change* 5.257-283.
- SAPIR, EDWARD. 1921. *Language*. Chicago: University of Chicago Press.
- SIGURDSSON, HALLDÓR ÁRMAN. 1988. "From OV to VO: evidence from Old Icelandic." *Working Papers in Scandinavian Syntax* 34.
- SIGURDSSON, HALLDÓR ÁRMAN. 1990. "V1 declaratives and verb raising in Icelandic." In Joan Maling and Annie Zaenen (eds.) *Syntax and Semantics, Vol. 24: Modern Icelandic Syntax*. San Diego: Academic Press.
- STAAL, J.F. 1967. *Word order in Sanskrit and universal grammar*. Dordrecht: Reidel.
- STOCKWELL, ROBERT P. 1977. "Motivations for exbraciation in Old English." In Charles Li (ed.), *Mechanisms of syntactic change*. Austin: University of Texas Press.

- STOCKWELL, ROBERT P. AND DONKA MINKOVA. 1991. "Subordination and word order change in the history of English." In Dieter Kastovsky (ed.), *Historical English syntax*. Berlin-New York: Mouton de Gruyter.
- SWAN, TORIL. 1988. *Sentence adverbials in English: a synchronic and diachronic investigation*. Oslo: Novus.
- SWAN, TORIL. 1994. "Old English and Old Norse initial adverbials and word order." In Toril Swan, Endre Mørck, and Olaf Janse Westvik (edd.), *Language change and language structure: Older Germanic languages in comparative perspective*. Berlin-New York: Mouton de Gruyter.
- TABOR, WHITNEY. 1985. *Syntactic innovation: a connectionist model*. Ph.D. Dissertation, Stanford.
- TAYLOR, ANN. 1994. "The change from SOV to SVO in ancient Greek." *Language Variation and Change* 6.1-37.
- THRÁINSSON, HÖSKULDUR. 1994. "Comments on the paper by Vikner." In David Lightfoot and Norbert Hornstein (eds.), *Verb movement*. Cambridge: University Press.
- USZKOREIT, HANS. 1985. *Constraints on order*. Stanford: SRI and CSLI.
- VENNEMANN, THEO. 1975. "An explanation of drift." In Charles Li (ed.), *Word order and word order change*. Austin: University of Texas Press.
- VIKNER, STEN. 1991. *Verb movement and the licensing of NP-positions in the Germanic languages*. Ph.D. dissertation, University of Geneva.
- VIKNER, STEN. 1994. "Finite verb movement in Scandinavian embedded clauses". In David Lightfoot and Norbert Hornstein (eds.), *Verb movement*. Cambridge: University Press.
- WEBELHUTH, GERT. 1991. *Syntactic saturation and the modern Germanic languages*. Oxford: Oxford University Press.
- WECHSLER, STEVEN. 1990. "Verb second and illocutionary force in Swedish." *Edinburgh Working Papers in Cognitive Science* 6:229-24.
- WECHSLER, STEVEN. 1991. *Argument structure and linking*. Ph.D. dissertation, Stanford University, Stanford, California.
- WEERMAN, FRED. 1989. *The V2 conspiracy*. Dordrecht: Foris.
- WEXLER, KEN. 1994. "Finiteness and head movement in early child grammars." In David Lightfoot and Norbert Hornstein (eds.), *Verb movement*. Cambridge: University Press.
- ZWART, C. JAN WOUTER. 1993. *Dutch syntax: a minimalist approach*. Groningen: Groningen Dissertations in Linguistics.