Structure and order: asymmetric merge

Jan-Wouter Zwart

2009

To appear in: Cedric Boeckx, ed., *Oxford Handbook of Linguistic Minimalism*.

Abstract.

The paper discusses the conversion from hierarchical structure to linear order, and argues that the conversion can be automatic if merge is asymmetric (considered in both bottom-up and top-down derivations). Assuming a layered derivation architecture, which requires elements constructed in separate derivations to pass through the interfaces before being included in the numeration of the next derivation, the paper argues that linear orders deviating from the automatic structure-to-order conversion may originate at the sound interface separating two derivation layers, and suggests that head-finality in a head-initial language may be a linguistic sign, signaling derivation layering.

Keywords.

merge; linear order; asymmetric merge; linear correspondence axiom; ov/vo-order; layered derivations; lexicalism; recursion; bare phrase structure

Structure and order: asymmetric merge

Jan-Wouter Zwart

1. Introduction

In this chapter I understand '(linear) order' (said of linguistic elements α,β) in terms of temporal organization, such that α precedes β (is ordered before β) if and only if the time at which α is realized (in sound or gesture) precedes the time of at which β is realized. Order in this sense is traditionally considered to be the domain of syntax (cf. Ries 1927:141), but in linguistic minimalism, order is not established in narrow syntax but at the interface component dealing with sound (cf. Chomsky 1995:334-335).

This assumes a model of grammar where syntax in the narrow sense ('narrow syntax') is a computational system that takes elements from a lexicon (hereafter also called 'resource' or 'numeration') and merges them, creating a structure to be delivered for interpretation at interface components dealing with sound ('PF') and meaning ('LF'). In this model, linear order comes in only 'after syntax', i.e. as a modality-specific realization of a structure that is ordered hierarchically, but not linearly.

Syntax in the Minimalist Program, then, retreats to its core business of defining the way elements combine to create larger units (cf. Ries 1927:142), referring many traditional aspects of the theory of syntax (including also inflectional morphology) to more peripheral components. Nevertheless, the question of how order relates to structure has been a formative element of minimalist syntactic theory since Kayne (1992, 1994), in that it prompted the articulation of the 'bare phrase structure' theory (cf. Chomsky 1995:249) replacing traditional X-bar theory.

2. The Linear Correspondence Axiom

Traditional X-bar theory (Chomsky 1970, Jackendoff 1977) specifies a universal format for the structure of phrases, distinguishing heads, complements, specifiers and adjuncts as occupying well-defined structural positions in the phrase. The order of elements in the phrase is a function, partly of structure (in that complements appear closer to the head than specifiers and adjuncts) and partly of language-specific properties of the head, taking a complement to its right or left (the 'head parameter' or 'directionality parameter', yielding head-initial and head-final syntax, respectively).

In the bare phrase structure theory, structure is a function of the merger operation combining elements from the lexicon (hence 'merge'), which creates sets. Merge is recursive, in that the set created by merge is extended by each next operation merge, yielding the familiar hierarchical phrase structure organization. We return to the details of merge below, but the point to be made here is that the operation is autonomous, i.e. not bounded by requirements posed by a theory of phrase structure like X-bar theory.

Since the notion 'head' is not given up in the bare phrase structure theory (cf. Chomsky 1995:245), it remains possible to describe linear order in terms of the setting of a head parameter.

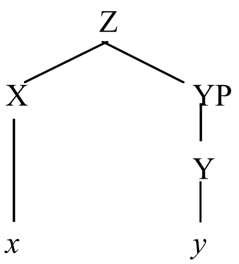
In bare phrase structure theory, the head determines the category of the output of merge, and it may be stipulated that it precedes or follows the nonhead (cf. Saito and Fukui 1998:452). But Kayne (1994) proposes that linear order is an automatic reflection of structural organization, through his Linear Correspondence Axiom, leaving no room for a directionality parameter (Kayne 1994:47).

The Linear Correspondence Axiom (LCA) was formulated with the traditional X-bar theory in mind, and proposes that the linear order of the terminals of a phrase structure Π reflects the asymmetric command relations among nonterminals of Π (the terminals are the actual linguistic items, and the nonterminals nodes in the phrase structure dominating the terminals). Command is the familiar notion of c-command, where α c-commands β iff β is (dominated by) the sister of α (and α is the sister of β iff α and β are merged together, and α dominates β iff α is the outcome of an operation merge involving [γ dominating] β). In order for the structure-to-order mapping envisioned in the LCA to be possible, nonterminals must be in asymmetric c-command relations, which in turn leads to proposals about structure to that effect (i.e. ensuring asymmetry of c-command relations among nonterminals) which are in part incompatible with the bare phrase structure approach.

The relevance of the LCA for the viability of a directionality parameter is the following. We observe within a language that linear order is antisymmetric, i.e. given two elements *x* and *y* we (generally) don't find both orders *xy* and *yx* (antisymmetric ordering of *x* and *y* means that *xy* and *yx* do not both occur). Kayne (1992) observes that this antisymmetry applies across languages, in that certain phenomena involving directionality (such as movement) do not co-vary with a supposed directionality parameter. For example, movements to the left in a head-initial language like English (wh-movement, subject-auxiliary inversion, etc.) are not mirrored as movements to the right in a head-final language like Japanese. Therefore, if the moved element is *x* and its trace is *y*, we do not find both *xy* and *yx* in this domain across languages. Antisymmetry, then, is a characteristic of 'universal grammar' (the language faculty as realized in all languages). If this can be generalized, i.e. if languages show no mirror effects at all, a directionality parameter cannot be part of universal grammar.

Kayne (1994) explicitly argues against a directionality parameter on theoretical and empirical grounds. Theoretically, he derives from the LCA that the specifier and the complement must be on opposite sides of a head (Kayne 1994:35). For linguistic items *x*, *y* functioning as a head and its complement to be ordered, the nonterminals *X* and *Y* associated with *x* and *y* must not c-command each other. Let *x* be the head; then *y* cannot be a head, since the complement of a head must be a phrase (Kayne 1994:8). If *y* is not complex, its nonterminals must include both a head-nonterminal and a phrase-nonterminal. The structure of the head-complement configuration, then, is as in (1):

(1) *head-complement configuration (Kayne 1994:10)*



nonterminals: Z, X, YP, Y

terminals: *x*, *y*

head: *x* complement: *y*

In (1), an asymmetric c-command relation exists between nonterminals *X* and *Y*, so that the terminals can be ordered as *xy*, but if *YP* were not present, *X* and *Y* would c-command each other, and no ordering of the terminals *x* and *y* would be possible. It is easy to see that a specifier *Q* (dominating terminal *q*) merged to *Z* in (1) would asymmetrically c-command both *X* and *Y*, so that the order of the terminals becomes *qxy*. Hence, the specifier and complement must be on opposite sides of the head, leaving no room for a directionality parameter regulating the head-complement order.

Empirically, the observations lead us to reject the only alternative ordering of the terminals of a specifier-head-complement structure, namely the inverse order *yxq*. This is because the specifier is predominantly on a left branch in both head-initial and head-final languages, witness the distribution of typical specifier occupants such as subjects (preceding predicates) and displaced wh-elements (fronted, i.e. moved to the left) across languages.

The combined theoretical and empirical observations lead Kayne to conclude that the linear order of phrases is universally specifier-head-complement. Deviations from the universal order must be the result of movement (where movement occurs when a term of Π is merged with Π), and languages differ not in a directionality parameter setting, but in the amount (and perhaps the type) of movement.

In assessing the LCA, it is important to separate the empirical observations from the theoretical proposal. The empirical observations (essentially of a typological nature) raise important questions relating to the significant *absence* of otherwise expected phenomena (e.g. Why is there no verb-second-to-last, Kayne 1992). These questions find a useful answer in a sweeping generalization like the LCA. But perhaps equally as important is the observation that few (if any) languages are altogether free from disharmonic word order phenomena. This requires that we define domains where deviations from the unmarked structure-to-order correspondence might originate (see section 6).

3. Bare phrase structure theory

In the bare phrase structure theory, structure is a function of merge, i.e. no nodes exist that are not the product of an operation merging two items. Projection levels (head, phrase) are contextually defined, and not given in advance by rewrite rules. It follows that a nonbranching complement (such as *y* in (1)) is simultaneously a head and a phrase: it is a head because it does not branch, and it is a phrase because it acts as the complement to another head. Furthermore, the bare phrase structure theory does not distinguish lexical items from nodes in the structure: the items *are* the elements merged, and so the items themselves constitute the structure.

These properties of the bare phrase structure theory are incompatible with Kayne's proposal to root the LCA (i.e. the structure-order correspondence) in asymmetric command relations. *YP* and *Y* in (1) are collapsed in a single node, yielding a symmetric sister pair of *X* and *YP/Y*. While Chomsky (1995:340) accepts the main empirical conclusions of Kayne (1994), including the universal head-complement order and the idea that deviations from the universal head-complement order are caused by movement, he proposes to rethink the role of the LCA in the theory of grammar.

In the bare phrase structure theory, the LCA no longer blocks the generation of symmetric structures—it's just not clear how to convert a symmetric structure to an ordered string (a sequence of sounds). If the LCA is a principle of the phonological component, the problem posed by symmetric structures disappears if the phonological component may ignore one of the elements of the symmetric structure. Chomsky therefore proposes that a nonbranching complement has to move (cliticize, incorporate) before the structure is turned over to the interfaces (Chomsky 1995:337). A trace of a moved category is ignored (or deleted) at the sound interface, obviating any ordering requirements (see Moro 2000 for extensive discussion of this proposal and its consequences).

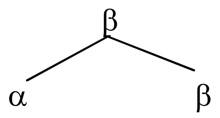
This leads to the strong prediction that any rightbranching structure ends in a trace. This raises at least two questions: 1. is it true that every rightbranching structure ends in a trace? and 2. if so, what triggers the movements creating these traces? Note that the movement trigger must be independent of the need to create a structure that is interpretable at the sound interface, as the movement takes place in narrow syntax and must be oblivious of interface requirements such as the LCA (*pace* Moro 2000:28-29). These questions have not been vigorously pursued in the literature, as far as I am aware.

If structure is a function of merge, as in the bare phrase structure theory, it becomes possible (and perhaps necessary) to think of structure not in terms of tree configurations, but in terms of sets. Chomsky (1995:243) therefore describes the output of merge of α and β as the set $K = \{ \alpha, \beta \}$. K may be merged again yielding another set containing K. Every phrase, then, is a recursively defined set of sets.

It is easy to see that when merge yields sets, asymmetric c-command relations can be expressed in terms of set membership properties. Thus, a specifier γ merged with $K = \{ \alpha, \beta \}$ yielding $L = \{ \gamma, \{ \alpha, \beta \} \}$ and α and β are not elements of the same set, as α and β are elements of K, a co-member of γ in L. Corresponding to this, γ c-commands α and β, but not vice versa. However, ordering by set membership yields no result among sisters, i.e. does not derive head-complement linear order (and there are other problems, having to do with the fact that set membership is not a transitive relation).

A separate question is posed by adjuncts. In traditional X'-theory, adjuncts are merged to maximal projections via Chomsky-adjunction (i.e., the node dominating the adjunct after adjunction is identical to the node to which the adjunct is merged; cf. Chomsky 1986:6). As a result, the sister of the adjunct involves two segments, with only the higher segment including the adjunct:

(2) *adjunction structure*



One interpretation of the configuration in (2) would be to state that α is neither included in nor excluded by β (cf. Chomsky 1986:7,9). If so, linear ordering of adjuncts cannot be a function of the set membership created by merge.

Replacing the familiar tree structure notation of phrase structure with set notation requires a rethinking of the notion projection (i.e. the determination of the features of a whole based on features of its parts). Chomsky (1995:244) views K, the set resulting from merger of α and β, as a slightly more complex object $\{ \alpha, \{ \alpha, \beta \} \}$, where α is the head of K and projects. α, then, is the *label* of K. For adjunction, Chomsky (1995:248) proposes that the label reflect the two-segment character of the object construed, and he suggests using to that end the ordered pair $\langle \alpha, \alpha \rangle$ as the label instead of just α. The set notation of (2) would then become $\{ \langle \alpha, \alpha \rangle, \{ \alpha, \beta \} \}$.

Notice that the label α of the output of merge $\{ \alpha, \beta \}$ is not itself ordered before or after the elements α and β at the sound interface PF (likewise with the adjunction label $\langle \alpha, \alpha \rangle$). This suggests that the label is a mere notational device, needed to express an inherent asymmetry among elements merged (cf. Collins 2002). This asymmetry (that one element is the head and the other is not) allows us to think of K as an ordered pair $\langle \alpha, \beta \rangle$, on the understanding that any dissimilarity among α and β in property P renders α and β ordered with respect to P. In this connection, Langendoen (2003:310) notes that $\{ \alpha, \{ \alpha, \beta \} \}$ is the set-theoretical definition of the ordered pair $\langle \alpha, \beta \rangle$ (more exactly, the set-theoretical definition of $\langle \alpha,\beta \rangle$ is $\{ \{\alpha\}, \{\alpha,\beta\} \}$, cf. Kuratowski 1921:171).

If the output of merge is an ordered pair by definition, all that is needed for the structure-to-order conversion is a correspondence rule that says (where material between slashes is ordered in time):

(3) *Structure-to-order conversion*
$\langle \alpha, \beta \rangle = / \alpha \; \beta /$

On this view, the head-complement distinction in itself is sufficient to bring on the asymmetry required for ordering at the interface. This raises the question (not further addressed here) whether the device of the label (essentially, the property of *projection*) is necessary and sufficient to turn the output of merge into an ordered pair.

I take (3) to be the 'silver bullet' of structure-to-order conversion. If it can be derived that merge itself yields an ordered pair rather than an unordered set, linear order follows almost trivially.

4. Timing and nesting

If the derivation of a syntactic structure D involves a sequence of steps, then stages of D $d_1, \ldots, d_n$ may be distinguished, and elements of D may be differentiated as to their existence in each stage. Assuming a bottom-up derivation, elements merged to D at stage d_i are not part of D at stage d_{i-1}. At each step of the derivation, then, an asymmetry exists between the two sisters being merged, in that one of the two sisters is already part of a derivation to which the other is newly merged (Jaspers 1998:109).

It follows that the output of merge is inherently asymmetric, except with first merge (assuming binary merge, i.e. involving exactly two elements). For most of the derivation, then, we may conclude that merge yields an ordered pair rather than an unordered set.

What about first merge? Let us assume that a derivation involves a single resource (lexicon, numeration) and a single target, the object under construction. If so, first merge is special in that it involves the selection (from a resource) of *two* elements. For each next step, it suffices to select a single element from the resource, the other element involved in the merger being the derivation under construction itself. Now if the target can be held constant (i.e. it is the unique object under construction), then merge can be simplified as in (4) (Zwart 2004):

(4) *Unary merge*
Merge selects a single element from a resource and includes it in the object under construction.

Before first merge, the object under construction is empty. First merge, then, simply includes an element from the resource in an empty workspace. At the next step, the workspace is no longer empty, and 'including an element in the object under construction' implies merger with that object.

Fortuny (2008:18f) demonstrates that this sequence of steps can be described in set-theoretical terms, where merge is an operation of set formation taking two sets A and B and producing the union of A and B. For merge to be successive, one of A and B (say, B) must be the output of the immediately preceding operation merge. At first merge, where there was no preceding operation, this output is zero, so that in that case B is the empty set $\varnothing$. Since the empty set is not phonetically realized, Fortuny derives the result of Chomsky (1995:337) and Moro (2000) that one of the elements involved in first merge must be empty at the sound interface PF.

It now follows that the asymmetry among the elements merged brought on by the derivational history (i.e. of two elements merged, one was already part of the derivation and the other is newly merged) applies to all stages of the derivation, including first merge. This allows us to think of merge as yielding an ordered pair, i.e. as being inherently asymmetric. If so, (3) may be taken to hold.

In formulating the structure-to-order rule (3), I assumed that merge yields an ordered pair automatically, i.e. as a function of the difference between the existing element (the derivation under construction) and the newly merged element. Fortuny (2008) shows that linear order can be derived from the history of the derivation, even if merge yields a set rather than an ordered pair.

Fortuny (2008:19) takes a derivation which successively merges the members of a resource S = { α, β, γ, δ, ε } to yield a derivational record K = { {α}, {α,β}, {α,β,γ}, {α,β,γ,δ}, {α,β,γ,δ,ε} }, which is the set of sets of elements merged at each stage of the derivation (for notational clarity, the empty set is left out in the member-sets of K). K is a *nest* (i.e. every set in K is a subset or superset of another set in K), which is shown by Kuratowski (1921:164) to provide a linear ordering of the members of S, i.e. the ordered n-tuple ⟨ α, β, γ, δ, ε ⟩. The order can be seen as a function of the number of sets in which elements are included, α being included in all sets of K, β in all but one, etc.

Note that the linear order of the example derivation would be / ε δ γ β α /, assuming bottom-up derivation and continuing to assume that specifiers are ordered before their associated heads and complements (i.e. if α and β form a head-complement combination, then γ is a specifier or a higher head, and must precede α and β, based on Kayne's observations). On Fortuny's derivation of the structure-to-order conversion, (3) reads as (3'):

(3') ⟨ α, β ⟩ = / β α /

Notice that this timing/nesting approach to structure-to-order conversion has nothing to say on the question of the order of the (most deeply embedded) head and complement: it is not immediately clear whether the first element merged should be a head or its complement. However, since complements are typically transparent (allowing subextraction), they cannot be 'lexical' in the sense of section 6.1, i.e. they cannot be construed in a separate derivation, and hence they cannot be merged as single items, as specifiers/adjuncts must be (see also Toyoshima 1997). The derivation, then, must start with the construction of the complement, as Fortuny (2008:20) also assumes.

The general approach discussed here makes no special provisions for adjuncts, and tacitly assumes that adjunction has no special status as far as syntactic structure is concerned (i.e. all merge = adjunction).

5. Order without merge

The approaches to structure-to-order conversion discussed above share the assumption that merge is a process transfering elements from a resource to a workspace (the structure under construction). Order can be derived from the circumstance that this transfer process involves a sequence of steps, yielding an ever increasing structure. This approach to structure building is questioned in Bobaljik (1995), who instead proposes that the derivation merely forges relations among the members of the resource. On this approach, the elements of 'transfer' and 'structure building' are just metaphors. (Consequently, Bobaljik makes no distinction between a 'resource' and a 'workspace', using the term 'workspace' for the numeration, which is now not depleted in the course of the derivation, but instead grows.)

Ignoring functional elements, the example *The man hit a ball* might be analysed in Bobaljik's system as involving an initial workspace (5), which is expanded to (7) after the steps in (6):

(5)	{the}	(6)	1. relate {the} and {man}
	{man}		2. relate {a} and {ball}
	{hit}		3. relate {hit} and {{a}, {ball}}
	{a}		4. relate {{the}, {man}} and {{hit}, {{a}, {ball}}}
	{ball}		

(7) {the}
{man}
{hit}
{a}
{ball}
{{the}, {man}}
{{a}, {ball}}
{{hit}, {{a}, {ball}}}
{{{the}, {man}}, {{hit}, {{a}, {ball}}}}

The family of sets in (7) is not a nest, suggesting that the timing/nesting approach to the structure-to-order conversion is lost in a system without a transfer process taking elements from the resource to the structure under construction. However, the timing/nesting approach can be restored if Bobaljik's system is sufficiently sharpened.

I believe that the system proposed in Bobaljik (1995) is not sufficiently restrictive, in that any (original or created) element may enter into a relation with any other element. Bobaljik (1995:56) capitalizes on this property of the system to derive what corresponds to interarboreal operations: this occurs when an element *x* previously merged with *y* (yielding *A*) then enters into a second relation with *z* (yielding *B*), after which *B* may enter a relation with *A* (or an element including *A*). Bobaljik (1995) argues that such operations are needed to derive a movement operation, such as head movement, where the moved element does not merge with the root node of the structure (i.e. it violates the Extension Condition of Chomsky 1993:22); in Bobaljik's analysis, a verb moving out of VP to T takes a sidestep to merge with T in a separate tree structure, yielding a complex head which is itself merged with the VP, yielding TP.

The system as a whole seems too unrestricted however. What is missing is a sense of direction in the derivation, and that is what precludes a straightforward structure-to-order conversion.

Two other remarks on Bobaljik's system are relevant here. First, the system does not involve as a first step 'merger with nothing/the empty set'. As a result, no ordering between *a* and *ball* results. Second, *the* and *man* in (7) occur in three subsets, as does the verb *hit*, suggesting that *the*, *man*, and *hit* have equal depth, which would not allow a straightforward transfer from structure to order. But that problem is remedied if *the man* can be construed in a separate derivation and is included in the workspace (5) as a single item (see section 6.1). In what follows, I assume that this is the correct approach to complex specifier elements: they are construed in separate derivations and included as a single item in the numeration that feeds the derivation in which they are to appear as specifiers.

What I would like to show here is that the intuitive appeal of Bobaljik's system, namely that the derivation merely establishes relations among elements in a resource, can be maintained without losing Fortuny's result that the structure-to-order conversion is a straightforward rule of the type in (3)/(3'), interpreting an ordered n-tuple as a sequence of elements ordered in time.

The proposal discussed here, developed in Zwart (2008), takes "merge" to be a process that splits the resource into a pair consisting of one item from the resource and the resource's residu ("split merge"; see also Fukui and Takano 1998, where a similar operation, "demerge", is part of the linearization process). The syntactic position and the grammatical function of the element split off from the resource are contextually defined, as a function of the relation with its sister, the residu of the resource. The residu itself becomes a dependent of the element split off. The derivation proceeds by split-merging the residu of the resource, splitting off one element with each step, until the resource is empty.

On this approach, starting from a resource S = { α, β, γ, δ, ε }, and splitting off α first, β next, etc., the derivation proceeds as in (8):

(8) *split merge*

STEP	SPLIT	RESOURCE
1.		{ α, β, γ, δ, ε }
2.	α	{ β, γ, δ, ε }
3.	β	{ γ, δ, ε }
4.	γ	{ δ, ε }
5.	δ	{ ε }
6.	ε	∅

The derivational record K can now be defined as a the set of sets of elements split off from the resource at each step, i.e. K = { {α}, {α,β}, {α,β,γ}, {α,β,γ,δ}, {α,β,γ,δ,ε} }, which is a nest yielding the ordered n-tuple ⟨ α, β, γ, δ, ε ⟩. The structure-to-order conversion then follows straightforwardly from rule (3).

The split-merge system shares with Bobaljik's system that it does not need to involve 'merger with nothing' as a first step, here a straightforward result of the top-down orientation of the derivation. Likewise, it involves no transfer from a resource to a workspace, reducing the importance of the concept of movement significantly (i.e., movement, 'internal merge' is now an additional mechanism, no longer modeled on the basic structure building operation '(external) merge'; cf. Zwart 2008 for discussion).

Note that in the split-merge system, the only constituents are, at each step, a) the elements split off from the resource, and b) the state of the resource (the combination of these two elements was defined as a constituent at the preceding step in the derivation). It follows that complex elements split off from the resource (such as specifiers, adjuncts) must be included in the resource as single items, i.e. must be the output of a separate derivation. Therefore, the derivation must be layered, as I believe is inevitable in a restrictive system.

In the minimalist literature, top-down derivations have been explored several times, most notably by Phillips (2003), and, building on Phillips' work, Richards (1999), and Chesi (2007). These proposals involve transfer from the resource to the structure, and the key concept is that merge expands structure to the right. A difference with the split-merge approach is that the right branch at each stage of the derivation is a linguistic item (rather than an unordered set), which is then replaced by a newly created branching structure at the next step. Space considerations prevent me from discussing this line of research in more detail.

6. Deviations

The pursuit of a regular and automatic structure-to-order conversion was motivated empirically by the word order asymmetries noted by Kayne (1992, 1994), and theoretically by the desire to keep narrow syntax free from ordering considerations. I have suggested that the output of merge may be an ordered pair rather than a set, and that the output of the derivation as a whole may be defined as an ordered n-tuple (Fortuny 2008, Zwart 2008). If so, a straightforward structure-to-order conversion of the type in (3) may be maintained.

However, if structure is created uniformly across languages, and something like (3) applies, how come the surface syntax of languages is riddled with deviations from the expected linear order pattern? It is to this question that we now turn, expecting to make some minor inroads at best.

6.1 Lexical and morphological

To begin with, a distinction must be made between head-final orders that may and may not be brought about by movement. The question of deviating word order is acute only for word orders that cannot be (or are unlikely to be) brought about by movement. This is because movement (a subcase of merge) of *x*, a complement of *y*, merges *x* with a phrase dominating *y*, establishing a new hierarchical relation between *x* and *y*. This is what allowed Kayne (1994) to maintain that word order variation need not be a function of a directionality parameter. So the question of deviating linear order has to abstract away from the effects of movement, and needs to address construction types where a movement analysis is impossible or unmotivated.

(In a split-merge approach, the status of movement is unclear, but we may assume as a starting point that the moved category is split off first, and that its 'base position' is filled by a contextually interpreted empty category; see Zwart 2008 for some discussion.)

More telling, then, is the observation that even strict head-initial languages like English show some amount of head-finality, for instance in the formation of compounds. English does not generally reorder heads and complements via movement, so a syntactic explanation of the complement-head order is not obvious (even if technically possible).

In the Minimalist Program, linear order is established at the sound interface, and it needs to be established to what extent deviations from the automatic structure-to-order conversion can be understood in terms of processes particular to the sound interface, i.e. processes that are not syntactic but rather 'morphological' or 'lexical'.

I take 'morphology' to be the inventory of forms expressing the properties of syntactic objects. Syntactic objects are created by merge, but at the sound interface they must be realized in forms which are stored and may have idiosyncratic properties (as is obvious from the example of inflectional paradigms). Syntactic features are instrumental in selecting the most suitable candidate from the set of stored forms (cf. Halle and Marantz 1993:121-122). On this view, complex forms that we mostly consider to be morphological, such as compounds, are construed in syntax via merge, but realized at the sound interface after an exchange with the morphological component (see Ackema and Neeleman 2004 for a general discussion of the syntactic nature of derivational morphology).

I take a form to be ‘lexical’ if it is enlisted in a numeration as a single item. That is, an item is lexical only in the context of a single numeration-derivation pair. Lexical integrity is the property of items of a numeration N that their parts are not manipulated in the course of the derivation built on N. Crucially, ‘lexical’ is opposed to ‘syntactic’ only in the context of a single (sub)derivation. Thus, a compound C may be created in derivation D_1, realized at the sound interface at the conclusion of D_1, and listed as a single item in the numeration for the next derivation D_2; C then is syntactic in the context of D_1 and lexical in the context of D_2. Idiosyncratic properties, including potentially deviating linear order, arise at the sound interface between D_1 and D_2. This definition of ‘lexical’ crucially assumes that derivations are layered, for which see Zwart (2008).

The model of grammar assumed here identifies the sound interface as the precise point of contact between automatic creative processes (merge) and stored knowledge (morphology). It remains to be shown that deviating linear order, overriding the automatic structure-to-order conversion, has this lexical/morphological character. If so, we expect constructions with deviating linear order to show idiosyncratic sound/meaning properties and/or reduced productivity.

6.2 Typological generalizations

Caballero et al. (2008) studied ordering effects in productive and unproductive noun incorporation in head-final and head-initial languages. They found that NV order is dominant in noun incorporation generally (36 out of 49), but exclusive in unproductive noun incorporation (8 out of 8). Moreover, about half of the VO-languages (9 out of 10) has NV order with productive noun incorporation (with OV-languages this is 15 out of 17). The survey shows more deviation from the expected structure-to-order conversion (i.e. from head-initial order) in unproductive types and almost exclusive deviation in the direction of head-final order. I take this to support the idea that linear order deviation is a function of morphology, on the assumption that a noun incorporation complex (certainly of the less productive type) is construed in a separate derivation and passes through the sound interface before being enlisted in the numeration for the next derivation as a single lexical item.

Caballero et al. (2008:398f) point out that similar results are obtained from a typological survey of synthetic compounds (of the type *skyscraper*, *witch hunt*) where they find “morphologically driven departures from syntactic word order” exclusively in VO-languages with NV synthetic compounds (in 8 out of 23 languages, with 0 out of 26 OV-languages showing VN compounds; cf. Bauer 2001). In this context, Caballero et al. (2008:400) note that clearly lexical incorporations (Type I of Mithun 1984) are NV in 3 out of 6 VO- languages, but VN in 0 out of 9 OV-languages, and the results are even more striking with grammaticalized noun incorporations where instrumental nouns have become derivational affixes (Caballero et al. 2008:400-401). Again, it appears that head-finality, certainly in head-initial languages, is brought about by morphology:

(9) *Generalization I*
Head-finality in a head-initial language is established at the sound interface.

Conversely, it appears that head-initiality in a head-final language is syntactic, i.e. a function of regular structure to order conversion. This may be concluded from a study of noun phrase coordination in head-final languages (Zwart 2005, 2009). If we take coordination to be a prototypical syntactic operation (developed out of mere juxtaposition if Mithun 1988 is correct), we expect the pattern to be the inverse from the pattern observed in noun incorporation and the like. That is, we expect head-final languages to show head-initial noun phrase coordination, and we do not expect head-initial languages to show head-final noun phrase coordination.

The terms 'head-initial' and 'head-final' coordination are justified on the hypothesis of Kayne (1994:12) that coordination structures are regular binary branching structures, headed by the conjunction, where the second conjunct is the complement of the conjunction. Head-initial and head-final coordination, then, are of the types *A & B* and *A B &*, respectively. The survey reported in Zwart (2009, table 3) indicates that 47 out of 57 head-final languages display head-initial coordination, whereas 0 out of 85 head-initial languages show head-final coordination. When we ignore coordination strategies not using a pure coordinating conjunction, but an adposition or some other device, head-final languages show no head-final coordination at all (Zwart 2009:1598). These observations suggest that structure-to-order conversion is regular in the syntactic domain:

(10) *Generalization II*
Head-initiality in a head-final language is established in narrow syntax.

The generalizations in (9) and (10) predict that if a noun phrase coordination takes on lexical (idiomatic) properties, head-finality will again become possible. Precisely this seems to be the case in Waigali, which has regular head-initial noun phrase coordination (11a), but also the head-final fixed expression (11b) (in Degener's 1998:166 words, this requires that the two conjuncts form a natural group):

(11) *Indo-European, Iranian: Waigali* (Degener 1998:166)
a. e meši ye e muša
a womananda man
'a man and a woman'

b. meši-moša-y
woman-man-and
'men and women'

I have not found any examples of the reverse (regular head-final coordination and idiomatic head-initial coordination).

6.3 The Final-over-Final constraint

If head-final order is brought about at the sound interface, head-final constructions are lexical items in the sense understood here (single numeration items), and we expect head-final constructions to occur embedded in regular syntactic head-initial structures, but not the reverse (head-initial structures embedded in head-final constructions), or at least not generally. Precisely this generalization has been formulated in Holmberg (2000:124) as the Final-over-Final Constraint or FOFC (see also Biberauer, Holmberg and Roberts 2008):

(12) *Final-over-Final Constraint (FOFC)*
If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final.

The FOFC essentially states that head-finality must be lower in the structure than head-initiality. Consider how this might follow.

If a complement β of a head α is complex, i.e. β is not the output of a separate derivation, then the phrase [α β] is linearized as / αβ / even on the original formulation of the LCA, where merge is symmetric but the structure asymmetric (given that β is complex), and likewise if merge yields asymmetric structure regardless of the complexity of β (as suggested in this chapter). It follows that if β is head-initial, it will be dominated by a head-initial phrase, as stated in the first clause of the FOFC.

If a complement β of a head α is the output of a separate derivation (i.e. is not complex in the context of the derivation in which α and β are merged), then the LCA does not apply on its original formulation, given bare phrase structure theory (see section 3). On the bare phrase structure approach, β must move in order to create a pair of α and a trace (which is ignored at spell-out, ensuring vacuous linearization); moved β then is manipulated in a syntactic derivation (becoming a specifier or adjunct) where we find only head-initial structure (on both the original and more recent formulations of the LCA). Assuming asymmetric merge (yielding an ordered pair), there is no need for β to move, but if no movement occurs we expect the head α to precede its complement β (by (3)). In this situation, where β is the output of a separate derivation, passing through the sound interface before being enlisted in the numeration, β may have acquired a deviating (head-final) linear order. But the mechanism does not allow for α to come out as head-final, as α and β are merged in a regular syntactic derivation.

The mechanism does not preclude α (dominating β) being subject to reordering at the sound interface concluding the derivation in which α was created via merger of its head with β. Potentially this might yield an exception to the FOFC, if β itself happens to be head-initial. But it is predicted that this situation will be idiomatic, having the flavor of the exceptional. The empirical motivation of the FOFC, as discussed in Biberauer, Holmberg and Roberts (2008) and references cited there, however, is squarely based on observations of regular syntax.

I conclude that the system of structure-to-order conversion contemplated in this chapter provides a rationale for the FOFC: since head-final order is a lexical property, head-final phrases will occur at the bottom of the syntactic tree, and head-finality will never be a property of the main projection line, which is a function of merge, not of the lexicon.

6.4 Head-finality in a head-initial language: a closer look

Let us call head-finality as the result of movement *pseudo-finality*. Kayne (1994) conjectured that all head-finality is pseudo-finality. The discussion in this subsection acknowledges that some head-finality may be pseudo-finality, but that there is also a large amount of real head-finality which is brought about by morphology at the sound interface. This raises the question how to distinguish the two types of head-finality.

In my earlier work on the syntax of Dutch (Zwart 1993, 1994) I tried to argue that deviations from head-initial syntax were caused by leftward movement of various elements to specifier positions in the functional domain. Some of the movements proposed, in particular the object shift movement affecting the position of definite and specific indefinite objects, were already assumed in the standard literature (cf. Vanden Wyngaerd 1989). Other movements, involving leftward shift of nonspecific indefinite objects, particles, and secondary predicates, were novel and somewhat suspect, in that the elements involved remained adjacent to the verb. One possibility is that the earlier analysis erred in conflating the two types of head-finality, and that the OV order without obligatory object-verb adjacency involves pseudo-finality (caused by movement), whereas the complements left-adjacent to the verb owe their surface position to linearization at the sound interface.

To illustrate the basic facts, Dutch shows an asymmetry between nominal and clausal verb complements, such that nominal complements precede, and clausal complements follow the verb in clause-final position (example sentences involve embedded clauses with the verb in its base position; in main clauses, the verb is realized in the 'verb second' position, following the first constituent).

(13) *nominal vs. clausal complements in Dutch*

a.	..dat	Jan	**die**	**dingen**	niet	beweert	
	COMP	John	DEM:PL	thing:PL	NEG	claim:3SG	
	'..that John does not claim those things.'						
b.	..dat	Jan	niet	beweert	**dat**	**het**	**regent**
	COMP	John	NEG	claim:3SG	COMP	it	rain:3SG
	'..that John does not claim that it is raining.'						

In (13a), the nominal complement *die dingen* 'those things' precedes the verb (obligatorily), and is separated from it by the negation element *niet* 'not'. In (13b), the clausal complement *dat het regent* 'that it is raining' follows the verb (obligatorily). Assuming that the complement of the verb is generated as a sister to the verb, the position of *die dingen* 'those things' in (13a) must be the result of leftward movement (object shift), and although traditionally the base position of the object is taken to precede the verb (e.g. Koster 1975), a possible extension of that analysis is that the object originates in the position to the right of the verb, occupied by the clausal complement in (13b) (thus Zwart 1994).

Other elements obligatorily preceding the verb (in clause-final position) include verbal particles (14a), secondary predicates (14b), and nonspecific indefinite objects (14c). In most dialects of spoken Dutch, the past participle also precedes the clause-final auxiliary (14d), though many patterns occur across Dutch dialects, especially if larger verb clusters are taken into account (cf. Zwart 1996, Barbiers et al. 2008).

(14) *elements left adjacent to the verb in Dutch*

a. ..dat Jan die dingen niet **op** schrijft
COMP John DEM:PL thing:PL NEG up write:3SG
'..that John does not write those things down.'

b. ..dat Jan het hek niet **rood** verft
COMP John the fence NEG red paint:3SG
'..that John does not paint the fence red.'

c. ..dat Jan zelden **een boek** leest
COMP John rarely a book read:3SG
'..that John rarely reads a book.'

d. ..dat Jan het boek niet **gelezen** heeft
COMP John the book NEG read:PART have:3SG
'..that John didn't read the book.'

The bold face elements in (14) cannot be separated from the finite verb by negation, adverbs, and the like. The only evidence that their placement might be the result of movement offered in Zwart (1994:400) is that stranded prepositions of adjunct prepositional phrases (PP) may break up the adjacency. Compare (15a), with a full adjunct PP and (15b) with a stranded adjunct preposition.

(15) *stranded preposition breaking up adjacency*

a. ..dat Jan het hek **met die kwast** rood verft
COMP John the fence with DEM brush red paint:3SG
'..that John paints the fence red with that brush.'

b. de kwast waar Jan het hek (mee) rood (**mee**) verft
the brush REL John the fence with red with paint:3SG
'the brush that John paints the fence red with'

The position of the secondary predicate *rood* 'red' in (15b), preceding the stranded preposition *mee* 'with' (at least optionally), was taken to indicate that the secondary predicate *rood* had been moved to the left, on the assumption that (a) adjunct PPs are outside the verb phrase, and (b) the stranded preposition could not have been lowered. However, the distribution of stranded prepositions is not well understood, and in the present context it might be assumed that the position of the stranded preposition in linear order is the result of linearization at the sound interface, as stranded prepositions are typically light elements (cf. Zwarts 1997). If so, no evidence for leftward movement of the boldface elements in (14) remains.

Could the head-finality in (14) be the effect of linearization at the sound interface? Some suggestion that this might be right is provided by the observation that idiomaticity in Dutch is sensitive to the left-right asymmetry as reflected in head-final vs. head-initial linear order. For example, whereas adjunct and complement PPs can freely appear to the right of the verb in final position, verb-PP idioms require the PP to precede the verb (cf. Veld 1993:148):

(16) *verb-PP idioms require head-final order*

a. ..dat Jan de pijp (aan Marie) geeft (aan Marie)
COMP John the pipe to Mary give:3SG to Mary
'..that John hands the pipe to Mary.'

b. ..dat Jan de pijp (aan Maarten) geeft (*aan Maarten)
COMP John the pipe to Marten give:3SG to Marten
'..that John quits.'
idiom: *de pijp aan Maarten geven* = to quit

In the model of grammar considered here, idioms are construed in a separate subderivation, and pass through the interfaces before being enlisted in the numeration for the next derivation. They are, then, 'lexical elements' with special sound-meaning properties acquired when passing through the interface separating two derivation layers (see section 6.1). The head-final order of the idiom *de pijp aan Maarten geven* 'quit' (and in fact of all verb-PP idioms) may then be the outcome of linearization at PF.

I would like to suggest that all head-final constructions in (14) are construed in separate derivations, and hence are 'lexical' in the sense of section 6.1 (i.e. they are 'complex predicates', cf. Neeleman 1994). First, verb-particle combinations (14a) are invariably highly idiomatic (e.g. *op bellen* [up ring] 'phone', *uit vinden* [out find] 'invent', *in dikken* [in thick] 'thicken', *aan vallen* [on fall] 'attack', *voor stellen* [fore put] 'propose'), and the verb cannot be fronted without pied piping the particle:

(17) *no particle stranding under verb fronting*

a. * Geschreven heeft Jan dat niet op
write:PART have:3SG John DEM NEG up
intended: 'John did not write that down.'

b. Op geschrevenheeft Jan dat niet
up write:PART have:3SG John DEM NEG
'John did not write that down.'

If the verb-particle combination is the output of a separate derivation, it is expected that the verb and the particle cannot be separated in the context of the next derivation in which the combination appears. The verb and the particle may be separated under verb second, where the finite verb moves to the position following the first constituent in main clause, but this is not an operation of narrow syntax if Chomsky (2001:37f) is correct.

The verb and the particle may also be separated by other material belonging to a verb cluster, as in *op heeft geschreven* [up has written] ‘has written down’, but the logic of the analysis entails that the entire verb cluster is to be the output of a separate derivation, and hence the separation may be the effect of linearization as well. In fact, the combination of a perfective participle (like *geschreven*) and the auxiliary *hebben*, originally a verb of possession, has come to be used to refer to a relative past tense as the outcome of a grammaticalization process. As a result, the verb-particle combination in (14d) also has idiosyncratic sound-meaning properties that might be taken to betray construction in a separate derivation.

The fact that verb clusters in Continental West-Germanic dialects display a wide variety of linear orders, both across and within dialects, has been taken to suggest that the derivation of these clusters requires special rules, such as reanalysis (Haegeman and Van Riemsdijk 1986), flipping (i.e. inversion of sister nodes, Den Besten and Edmondson 1981:43) and rightward movement of heads (‘verb raising’) and larger projections of the verb (‘verb projection raising’)(Evers 1975). Of these, reanalysis is arguably a process not of narrow syntax but of the sound interface (Zwart 2006; cf. Zubizarreta 1985:286, stating in this context that “the grammar does allow for mismatches between morphophonology and morphosyntax”). Flipping is clearly not merge, so it, too, must be considered a function of linearization at PF. We know that stylistic injunctions against ‘German sounding’ patterns play a role in determining auxiliary-participle orders in written and carefully spoken Dutch (Stroop 1970:252), indicating that structure-to-order conversion is not free from influence by stored knowledge in this domain. The concept of rightward movement has no status assuming bare phrase structure theory, so the directionality aspect of it must come in only at the sound interface.

For secondary predicates (14b), a complex predicate analysis has been proposed and extensively argued for by Neeleman (1994). Semantically, the verb and the secondary predicate form a tight connection. When the secondary predicate is a PP, we again observe that the PP cannot be extraposed, just like PPs that are part of a verbal idiom (cf. (16)).

(18) *no extraposition of secondary predicate PPs*

a. ..dat Jan in de sloot sprong
COMP John in the ditch jump:PAST.SG
$^{\mathrm{OK}}$adverbial reading: ‘..that John was jumping [sc. up and down] in the ditch.’
$^{\mathrm{OK}}$secondary predicate reading: ‘..that John jumped into the ditch.’

b. ..dat Jan sprong in de sloot
COMP John jump:PAST.SG in the ditch
$^{\mathrm{OK}}$adverbial reading: ‘..that John was jumping [sc. up and down] in the ditch.’
*secondary predicate reading (‘..that John jumped into the ditch.’)

As the absence of the resultative (secondary predicate) reading of the postverbal PP in (18b) shows, secondary predicate PPs behave like idioms. Non-PP secondary predicates also readily lend themselves to the formation of idioms with the verb (e.g. *zwart maken* [black make] ‘badmouth’, *om de tuin leiden* [around the garden lead] ‘mislead’, etc.).

Nonspecific indefinites (14c) lose their nonspecific interpretation as soon as they are not adjacent to the verb (again, not counting the effect of ‘verb second’ placement of the finite verb in main clauses). Thus, in (14c') the leftward shifted indefinite object acquires a generic reading.

(14) c'. ..dat Jan **een boek** zelden leest
COMP John a book rarely read:3SG
'..that John rarely READS a book.'
(i.e., what John rarely does to a book is read it.'

This indicates that the combination of the verb and the adjacent indefinite acquires idiosyncratic sound-meaning properties, which we took to be indicative of construction in a separate subderivation.

While firm conclusions cannot be drawn at this point, the observations may be taken to indicate that head-final linear order in Dutch, overriding the automatic structure-to-order conversion, may be a function of the punctuated character of the derivation, where elements construed in one derivation acquire idiosyncratic sound-meaning properties (including deviating linear order) while passing through the interfaces before being enlisted in the numeration for the next derivation layer. It would be interesting to consider the question how many movements (and movement types) may be dispensed with when the potential of linearization at the interface between derivation layers is more fully taken into consideration.

6.5 Conclusion: head-finality as a linguistic sign

If I am correct that deviating (head-final) linear order originates at the sound interface separating two derivation layers, the following holds:

(19) Head-finality is a linguistic sign, signaling derivation layering.

The function of reordering at the sound interface might be to brand complex, derived structures as single items for use in a further derivation. Put differently, linear order identifies structures as being either open-ended (head-initial) or sealed off (head-final). Using head-final linearization to signal derivation layering is clearly not obligatory, but the crucial observation is that only head-final orders can be argued to perform such a signal function.

If the relation between linearization and derivation layering suggested in (19) is real, linear order is intimately connected with a fundamental property of the faculty of language, recursion (cf. Hauser, Chomsky, and Fitch 2002). To be clear, I understand 'recursion' slightly different from what is standard. The operation merge is standardly taken to be recursive in that the output of merge may be subject to further operations merge. However, each subderivation may just as well be taken to be iterative rather than recursive, for instance if each step of the derivation transfers a single element from the resource to the workspace (see section 4), or splits a single element off from the resource (see section 5). What is unquestionably recursive, though, is the process whereby the output of one derivation functions as a single item in the next derivation, and this is how I understand recursion here (cf. Hofstadter 2007:83, who identifies the recursive capacity of treating complex concepts as single packets, to be combined with other concepts ad infinitum, as a species-specific property of human cognition). If (19) is correct, linear order signals recursion in this sense.

In this connection, we may understand the memory limitation on center-embedding discussed in Yngve (1961) as reflecting a limit on the number of derivational loops that can be kept track of in the context of a single utterance. Inasmuch as the concept of center-embedding refers to the linear order of a branching and a non-branching category, it has no status in the minimalist approach to structure building (there is no 'center'). But a 'left branch element' (an adjunct or a specifier) must be merged with the object under construction (or split off from the resource under the approach of section 5) as a single whole, and therefore it has to be the output of a separate derivation.

7. Conclusion

In this chapter I have discussed the relation between structure and linear order in the minimalist approach to syntactic theory.

The general idea of Kayne (1994) that linearization is a function of structural asymmetry among syntactic nodes can be maintained in the bare phrase structure theory of Chomsky (1995), if we take the history of the derivation into account.

On its simplest definition, merge is the same at each step of the derivation, i.e. first merge should have no special properties. This is achieved if merge is taken to be an operation transfering one element at a time from a resource to a workspace (the object under construction). Simplifying even more, and adopting a top-down derivational approach, we can take structure to result from an operation that splits off one element at a time from the resource ('split-merge') until the resource is empty. Either way, sister pairs are not sets but ordered pairs, and the set of elements merged/split off is a nest, which is equivalent to an ordered n-tuple.

This allows us to consider structure-to-order conversion as the trivial equivalence relation in (3) (where material between slashes is ordered in time, i.e. realized one after the other in sound or gesture).

(3) *Structure-to-order conversion*

$\langle \alpha, \beta \rangle = / \alpha\ \beta /$

Deviations from (3) can be of two types: (i) pseudo-finality, which is the result of movement, or (ii) real finality, which is the outcome of a morphological process at the sound interface PF, where syntactic structures are replaced by linear strings. Movement rearranges the hierarchical relations among elements, and is therefore, on the system considered here, expected to lead to a change in linear order. The only real head-finality, then, is of type (ii), i.e. is essentially lexical.

In this context, I defined 'lexical' in relation to the process of derivation layering. A derivation is layered when an element in its numeration is the output of a previous derivation. The output of a derivation passes through the interfaces before being enlisted in the numeration of a subsequent derivation. I have argued that head-final order may result from a marked structure-to-order conversion at the sound interface, as part of a set of idiosyncratic sound-meaning properties acquired at that particular juncture. I have suggested that head-final order is a sign indicating that a complex element (the output of a derivation) is to be considered as a single item in the context of the next derivation.

Capitalizing on the architecture of layered derivations allows one to redefine the concepts ‘syntactic’ and ‘lexical’ in relation to derivations, such that an element may be syntactic in one derivation layer (i.e. created by merge) and lexical in the next (i.e. treated by merge as a single item). Since ordering does not exist in narrow syntax, it must be brought in at the sound interface. The layered derivation approach entails that linearization is not a once-only process, but may be interlaced with stages of structure building in which linear order plays no role (cf. Uriagereka 1999). Linear order, then, may be frozen in ‘packets’ manipulated by merge in the context of subsequent derivation layers.

Once this is understood, it becomes possible to reconsider analyses that have been proposed to derive head-final order in narrow syntax. It will be clear that syntactic operations that seem to lack independent motivation are the first suspects. I have tried to argue that some aspects of Dutch head-final order, unsuccessfully described in terms of movement in previous analyses, may indeed be explained as a function of linearization at the interface.

Further research might be directed at those instances of head-final order which are currently derived via the somewhat suspect process of ‘roll-up movement’ (turning [A [B [C]]] into [[[C]+B]+A] via successive incorporative movements). Inasmuch as the model for roll-up movement was provided by the distinctly morphological process of inflectional affixation (e.g. Baker 1985, Pollock 1989), a reduction to morphological linearization processes might not be unattractive for at least some of the cases.

What is missing from the present discussion is a more fundamental assessment of the factors entering into linearization at the sound interface. I have suggested that linear order is partly simply given by the morphological inventory, i.e. is stored knowledge of a fixed form-meaning pairing. But in addition, prosodic factors having to do with ‘lightness’ of elements (clitics, stranded prepositions, discourse particles) (cf. Anderson 1993) and ordering requirements particular to linkers (Maxwell 1984:253) are arguably involved, as may be other factors, and a more fundamental investigation of these factors and their interaction is needed for a complete understanding of the nature of structure-to-order conversion.

References

Ackema, Peter and Ad Neeleman (2004). *Beyond morphology: interface conditions on word formation*. Oxford: Oxford University Press.

Anderson, Stephen R (1993). 'Wackernagel's revenge: clitics, morphology, and the syntax of second position.' *Language* 69, 68-98.

Baker, Mark C. (1985). 'The Mirror Principle and morphosyntactic explanation.' *Linguistic Inquiry* 16, 373-415.

Bauer, Laurie. (2001). 'Compounding.' In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.), *Language typology and language universals*, 695-707. Berlin: Mouton de Gruyter.

Barbiers, Sjef, Johan van der Auwera, Hans Bennis, Margreet van der Ham, Gunther de Vogelaer, and Eefje Boef (2008). *Syntactic atlas of Dutch dialects, Vol. 2*. Amsterdam: Amsterdam University Press.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts (2008). 'Structure and linearization in disharmonic word orders.' *Proceedings of the West Coast Conference on Formal Linguistics* 26, 96-104.

Bobaljik, Jonathan D. (1995). 'In terms of merge.' *MIT Working Papers in Linguistics* 27, 41-64.

Caballero, Gabriela, Michael J. Houser, Nicole Marcus, Teresa McFarland, Anne Pycha, Maziar Toosarvandani, and Johanna Nichols (2008). 'Nonsyntactic ordering effects in noun incorporation.' *Linguistic Typology* 12, 383-421.

Chesi, Christiano (2007). 'An introduction to phase-based minimalist grammars: why move is top-down and from left-to-right.' *CISCL Working Papers on Language and Cognition* 1 (Studies in Linguistics), 38-75.

Chomsky, Noam (1970). 'Remarks on nominalization.' In R.A. Jacobs and P.S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184-221. Waltham: Ginn and Co.

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

Chomsky, Noam (1993). 'A minimalist program for linguistic theory.' In K. Hale and S.J. Keyser (eds.), *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, 1-52. Cambridge: MIT Press.

Chomsky, Noam (1995). 'Categories and transformations.' In *The Minimalist Program*, 219-410. Cambridge: MIT Press.

Chomsky, Noam (2001). 'Derivation by phase.' In M. Kenstowicz (ed.), *Ken Hale: a life in language*, 1-52. Cambridge: MIT Press.

Collins, Chris (2002). 'Eliminating labels.' In S.D. Epstein and T.D. Seely (eds.), *Derivation and explanation in the Minimalist Program*, 42-64. Malden: Blackwell.

Degener, Almuth (1998). *Die Sprache von Nisheygram im afghanischen Hindukusch*. Wiesbaden: Harrassowitz Verlag.

Den Besten, Hans and Jerrold A. Edmondson (1981). 'The verbal complex in Continental West Germanic.' *Groninger Arbeiten zur germanistischen Linguistik* 19, 11-61. http://gagl.eldoc.ub.rug.nl/root/1981-19/02/

Evers, Arnold (1975). *The transformational cycle in Dutch and German*. Dissertation, University of Utrecht.

Fortuny, Jordi (2008). *The emergence of order in syntax*. Amsterdam: Benjamins.

Fukui, Noaki and Yuji Takano (1998). 'Symmetry in syntax: merge and demerge.' *Journal of East Asian linguistics* 7, 27-86.

Haegeman, Liliane and Henk van Riemsdijk (1986). 'Verb Projection Raising, scope, and the typology of rules affecting verbs.' *Linguistic Inquiry* 17, 417-466.

Halle, Morris and Alec Marantz (1993). 'Distributed morphology and the pieces of inflection.' In K. Hale and S.J. Keyser (eds.), *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, 111-176. Cambridge, Mass.: MIT Press.

Hauser, Mark D., Noam Chomsky, and W. Tecumseh Fitch (2002). 'The faculty of language: what is it, who has it, and how did it evolve?' *Science* 298, 1569-1579.

Hofstadter, Douglas (2007). *I am a strange loop*. New York: Basic Books.

Holmberg, Anders (2000). 'Deriving OV order in Finnish.' In: P. Svenonius (ed.), *The Derivation of VO and OV*, 123-152. Amsterdam: Benjamins.

Jackendoff, Ray S. (1977). *X'-syntax*. Cambridge: MIT Press.

Jaspers, Dany (1998). 'Categories and recursion.' *Interface* 12, 81-112.

Kayne, Richard S (1992). 'Word order.' GLOW keynote lecture, Lisbon, April 13.

Kayne, Richard S. (1994). *The antisymmetry of syntax*. Cambridge: MIT Press.

Koster, Jan (1975). 'Dutch as an SOV language.' *Linguistic Analysis* 1, 111-136.

Kuratowski, Kazimierz (1921). 'Sur la notion de l'ordre dans la théorie des ensembles.' *Fundamenta mathematicae* 2, 161-171.

Langendoen, D. Terence (2003). 'Merge.' In A. Carnie, H. Harley and M. Willie (eds.), *Formal approaches to function in grammar: in honor of Eloise Jelinek*, 307-318. Amsterdam: John Benjamins.

Maxwell, Dan (1984). 'A typologically based principle of linearization.' *Language* 60, 251-285.

Mithun, Marianne (1984). 'The evolution of noun incorporation.' *Language* 60, 847-894.

Mithun, Marianne (1988). 'The grammaticization of coordination.' In J. Haiman and S.A. Thompson (eds.), *Clause combining in grammar and discourse*, 331-359. Amsterdam: Benjamins.

Moro, Andrea (2000). *Dynamic antisymmetry*. Cambridge: MIT Press.

Neeleman, Ad (1994). *Complex predicates*. Dissertation, University of Utrecht.

Phillips, Colin (2003). 'Linear order and constituency.' *Linguistic Inquiry* 34, 37-90.

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

Richards, Norvin (1999). 'Dependency formation and directionality of tree construction.' *MIT Working Papers in Linguistics* 34, 67-105.

Ries, John (1927[2]). *Was ist Syntax? Ein kritischer Versuch*. Prague: Taussig & Taussig.

Saito, Mamoru and Naoki Fukui (1998). 'Order in phrase structure and movement.' *Linguistic Inquiry* 29, 439-474.

Stroop, Jan (1970). 'Systeem in gesproken werkwoordsgroepen.' In J. Stroop, *Nederlands dialectonderzoek: artikelen uit de periode 1927-1982*, 247-264. Amsterdam: Huis aan de Drie Grachten, 1983.

Toyoshima, Takashi (1997). 'Derivational CED.' *Proceedings of the West Coast Conference of Formal Linguistics* 15, 505-519.

Uriagereka, Juan (1999). 'Multiple Spell-Out.' In S.D. Epstein and N. Hornstein (eds.), *Working minimalism*, 251-282. Cambridge: MIT Press.

Veld, Joop (1993). *Postverbal constituents in Dutch and Turkish*. Dissertation, University of Amsterdam.

Yngve, Victor (1961). 'The depth hypothesis.' *Proceedings of the American Mathematical Society* 12, 130-138.

Zubizarreta, Maria Luisa (1985). 'The relation between morphophonology and morphosyntax: the case of Romance causatives.' *Linguistic Inquiry* 16, 247-289.

Zwart, Jan-Wouter (1993). *Dutch syntax: a minimalist approach*. Dissertation, University of Groningen.

Zwart, Jan-Wouter (1994). 'Dutch is head-initial.' *The linguistic review* 11, 377-406.

Zwart, Jan-Wouter (1996). 'Verb clusters in Continental West-Germanic dialects.' In J.R. Black and V. Motapanyane (eds.), *Microparametric syntax and dialect variation*, 229-258. Amsterdam: John Benjamins.

Zwart, Jan-Wouter (2004). 'Unary merge.' Paper presented at the Tilburg University Staff Seminar, October 28.

Zwart, Jan-Wouter (2005). 'Some notes on coordination in head-final languages.' In J. Doetjes and J. van de Weijer (eds.), *Linguistics in the Netherlands 2005*, 232-241. Amsterdam: Benjamins.

Zwart, Jan-Wouter (2006). 'Over het enten van interpolaties.' *Tabu* 35, 163-180.

Zwart, Jan-Wouter (2008). 'Prospects for top-down derivation.' To appear *Catalan Journal of Linguistics*.

Zwart, Jan-Wouter (2009). 'Relevance of typology to minimalist inquiry.' *Lingua* 119, 1589-1606.

Zwarts, Joost (2007). 'Lexical and functional prepositions.' In D. Haumann and S.J. Schierholz (eds.), *Lexikalische und grammatische Eigenschaften präpositionaler Elemente*, 1-18. Tübingen: Max Niemeyer Verlag.