The Asymmetry and Antisymmetry of Syntax[1]

A Relational Approach to Displacement

Justin Malčić

2019

Part IIB Dissertation
Linguistics Tripos
University of Cambridge

9,963 words

[1]Special thanks go to Theresa Biberauer for her invaluable comments and support while supervising this dissertation. Any remaining errors are my own.

Abstract

In both syntax and phonology, it has long been observed that significant restrictions exist on displacement. One such restriction ensures that displacement leads to sequences of elements which are in some sense contiguous, formalised in syntax in the concept of Feature Geometry-based Relativised Minimality by Starke (2001) and Contiguous Agree by Nevins (2007), and in Autosegmental Phonology by the Line-Crossing Prohibition (originating in the Well-formedness Condition in Goldsmith 1976).

I argue that effects of this type, which have been called *Contiguity Effects*, are best captured by taking displacement to involve total weak orders of elements in the sense of Order Theory. Building on work taking the LCA to hold throughout the derivation, I argue that precedence relations may be the basis of phrase structure, though without claiming that linearisation is necessary for LF (as for example suggested in Kayne 2013). I then develop this approach to show that Order Theory provides useful axioms for both phrase structure and displacement, and that the existence of displacement is expected given the use of Order Theory.

The extent of 'antisymmetry' in syntax has been a major issue for phrase structure theories ever since Kayne's (1994) original proposal for the Linear Correspondence Axiom. The requirement for antisymmetry in linearisation seems to follow from the nature of the phonological component, which seems to show a different kind of recursion to that displayed in syntax. However, it is not at all clear why the only kind of order (in the Order-Theoretic sense) relevant to syntax is specifically a *strict* total order (i.e. based on asymmetric relations).

It is important to highlight here that what Kayne (1994 et seq.) calls 'antisymmetry' is referred to outside of linguistics as 'asymmetry', being based on asymmetric—not antisymmetric—relations. For clarity, symmetry holds when the existence of a relation from a to b implies its own inverse (from b to a). Asymmetry is the opposite: the existence of a relation from a to b implies the lack of a relation from b to a. Antisymmetry, by contrast, holds if a symmetric relation implies equality (so for example the equality relation is both symmetric and antisymmetric). These definitions are given in (1).

(1)

Symmetry	**Asymmetry**	**Antisymmetry**
$aRb \iff bRa.$	$aRb \Rightarrow \neg bRa.$	$aRb \wedge bRa \Rightarrow a = b.$

Two of these relations have corresponding orders: asymmetric relations form strict orders, and antisymmetric relations weak orders (symmetric relations only lead to a lack of ordering). In (2a), β is ordered before α as there is an asymmetric relation from β to α (indicated by the arrow). In (2b) δ is ordered before γ, but in addition δ is equal to δ, and γ is equal to γ, because for each node there is a reflexive relation.

As can be seen, the key difference between asymmetric and antisymmetric relations is reflexivity: asymmetric relations are necessarily irreflexive, but antisymmetric relations may be reflexive. And finally, (2c) shows that symmetric relations cannot be used for ordering, since there is no principled way to determine based on the relations that ζ/θ comes before or after ε/η.

(2)

a. **Strict order**	b. **Weak order**	c. **No order**
		$\zeta \longleftrightarrow \varepsilon$
	$\circlearrowright \quad \circlearrowright$	$\circlearrowright \quad \circlearrowright$
$\beta \longrightarrow \alpha$	$\delta \longrightarrow \gamma$	$\theta \longleftrightarrow \eta$

So given the prominence of asymmetric c-command relations due to Kayne's (1994) LCA, and given Moro's (1997) suggestion that symmetric c-command is also crucial, it is worth asking whether *antisymmetric* relations play any role in syntax. More pointedly, why is (2a) available to the Language Faculty but not (2b)? To my knowledge the possibility that it *does* play some kind of role has not been extensively explored

(perhaps in part due to the terminological confusion). And even if it plays no role, it still needs to be ruled out in some principled way.

While the ability to extend the Phrase Marker using material internal to it is certainly not ruled out by Merge, Merge does not of itself predict that the two resulting occurrences of material should be identified with each other (especially given the intervening distances often involved, and, potentially, phase boundaries). Recalling that under antisymmetry, $aRb \wedge bRa \Rightarrow a = b$, weak orders appear able to capture this effect, and if so then there ought to be apparent idiosyncrasies of displacement resulting from this. In section 2, I show that well-known contiguity effects observed in syntax and phonology are indeed predicted by this requirement. But first, I review the role of Order Theory in syntax, and argue that it is much more wide-ranging than commonly assumed.

1 Evidence for asymmetry in syntax

The use of Set Theory as the basis of phrase structure originates in Chomsky's (2014/1995 p. 223) definition of Merge, a function taking two Syntactic Objects α and β, which returns the set of these elements $\{\alpha, \beta\}$. The original rationale for this was that this (unordered) set is the 'simplest object' able to be formed:

> The simplest object constructed from α and β is the set $\{\alpha, \beta\}$, so we take K [the output of Merge of α and β—J.M.] to involve at least this set. (Chomsky 2014/1995 p. 223)

This definition of Merge, termed Set Merge, can be distinguished from another kind yielding ordered pairs, known as Pair Merge, introduced in Chomsky 2004 pp. 117-118. It is explicit that Pair Merge is more complex than what has been termed 'simple Merge' (Chomsky 2013 p. 42), and as long as Set Merge exists, Phrase Structure remains ultimately defined in terms of Set Theory. This becomes especially clear in Chomsky 2013; 2015 where labelling is fully determined by the Labelling Algorithm, occurs once per phase, and is strictly independent of the phrase structure component. In this case the function of Set Merge really is just to form sets comprising the two arguments to the function.

There is an apparent symmetry here, in that $Merge(\alpha, \beta)$ does not directly imply that α or β is subordinate to the other—it is the Labelling Algorithm that is sensitive to the head-phrase distinction in, for example, $\{\alpha, \beta\text{P}\}$. Indeed, on this basis Chomsky (2013; 2015) proposes that symmetry forces movement in order to meet the requirements of the Labelling Algorithm. But there is an important question as to whether this symmetry is real or just a misleading edge case, because only in the merging of two phrases does there appear to be real symmetry between the function's arguments: there is clear asymmetry between heads and phrases, and in the case of two heads,

Chomsky (2013 p. 47) suggests that a root-categoriser distinction is enough to ensure the complex is labelled by the categoriser. In the following I suggest that merging of two phrases also exhibits a distinction, and that the apparent symmetry is misleading.

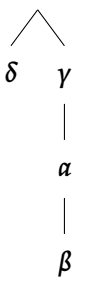
1.1 The status of c-command

Another way to characterise a symmetric combinatorial operation such as Merge would be to have the operation establish an abstract symmetric relation between two elements α and β. And indeed representationally for the structure $\{\alpha, \beta\}$ produced by Chomsky's (2014/1995 p. 223) version of Merge, α would standardly c-command β and vice versa. On this basis Epstein (1999) proposes that the effect of Merge is precisely to establish such c-command relations derivationally, obviating the need to compute them representationally as under the standard view.

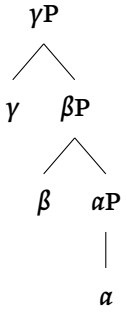
But even if c-command is not as intimately connected to Merge as this, its apparent relevant to extremely disparate syntactic phenomena (binding, agreement, movement, etc., and for many other examples see Uriagereka 2011 p. 125) suggests that it must be fundamental to the system in some non-trivial way. Indeed, Frank and Vijay-Shanker (2001) show that taking c-command as a primitive restricts possible structures to a subset of those able to be linearised with Kayne's (1994) Linear Correspondence Axiom. Specifically, structures which can be defined in terms of c-command have no nodes joined to two others by unary branching, structures which the LCA of itself does not rule out, since it addresses only terminals. Observe that the c-command relations in (3a) and (3b) are identical: crucially δ asymmetrically c-commands both β and α in both cases, no other node c-commands them, and they do not c-command any node. Hence because c-command is not extensional over these structures, taking c-command as a primitive enforces structures such as (4), which is of course LCA-compatible. And note also that a single instance of unary branching, needed for First Merge, is *not* ruled out.

(3) a.

b.



(4)



But supposing c-command is indeed the basis of phrase structure, it is still unclear why this particular relation should play such a major role in syntax. As the closure of a

symmetric relation, sisterhood, and an asymmetric relation, dominance[1], c-command is a rather curious relation, yet as Epstein (1999) points out, c-command was proposed very early (effectively in Klima 1964 though his in-construction-with relation is really the inverse of c-command), and has outlasted many other proposed relations, notably m-command and Government (for a review and formalisation of these and many other relations, see Barker and Pullum 1990), suggesting that its formulation is broadly correct. However, as Bruening (2014) points out, there have been a number of proposals over the years to use a slightly different relation to c-command, namely precede-and-command, which instead combines two asymmetric relations, precedence and dominance, resolving the contrast between symmetric sisterhood and asymmetric dominance. The other possibility for simplifying c-command would be to attempt to reduce it to only precedence or dominance, by causing these two relations to coincide.

In fact, under the LCA, asymmetric c-command already coincides with both precedence and dominance to a large extent, suggesting that the nature of linearisation may be instructive in understanding what c-command really is. Uriagereka (2011 pp. 141-142), for example, has suggested that the existence of c-command in syntax follows from its role in linearisation under Kayne's (1994) LCA as defined in (5). But it has been argued—by Chomsky (2014/1995 pp. 307-313) among others—that the LCA only holds at PF, given the lack of convincing effects due to linear order that would be otherwise expected at LF. Admittedly Kayne (2013) has claimed that for instance, the cross-linguistic predominance of forward over backward pronominalisation could be explained by the effect of linear order at LF[2]. But given that so many core cases of binding famously show a sensitivity to c-command over linear order, it seems wise to maintain that linearisation does not occur at LF, and find another way to explain the pronominalisation facts.

(5) **Linear Correspondence Axiom (Uriagereka 2011 p. 84)**

a. Base:
When *x* asymmetrically c-commands *y*, *x* precedes *y*.

b. Induction:
If a non-terminal *X* dominates a terminal *y*, and *X* is linearized with regard to terminal *z*, then *y* is linearized with regard to *z*.

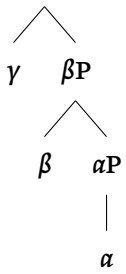
Suppose therefore that linearisation does not occur at LF, which is not sensitive to linear order. In this case, Uriagereka's (2011) explanation for the presence of c-command in syntax—that it follows from the need to linearise structure—will have no bearing of itself on LF. Indeed it might be expected that phenomena at LF and PF would make use of entirely different relations. But this appears not to be the case: the effects

[1]Assuming proper (irreflexive) dominance.
[2]Linear order because, given Principle B, restrictions on the relative ordering of pronouns must be defined in some other way than c-command.

of c-command are as widespread at LF as at PF, notably in determining scope and in quantifier raising, hence it seems clear that c-command is required for both PF and LF phenomena. Uriagereka (2011 pp. 142-143) therefore suggests making phenomena at LF parasitic on c-command, as it already exists in the Narrow Syntax given the need for externalisation at PF. This goes against much Minimalist thinking which views externalisation as a somewhat peripheral aspect of language (Chomsky 2013 p. 36), and arguably replaces the so-called 'LF bias' with a 'PF bias', in that properties of LF follow from the needs of PF. By contrast, for Chomsky (2014/1995) the need to linearise at PF arises due to the inability of PF to handle the kinds of structures in the Narrow Syntax and at LF.

However, while the claim that properties of LF follow from the requirements of PF is open to dispute, there does seem to be evidence that (subparts of) structures able to be defined in terms of c-command are easier to parse. Uriagereka (2011) observes that structures such as (6) can be parsed by Finite State Automata, since the rules in a Regular Grammar are of the form $S \rightarrow aA$, $S \rightarrow Aa$, or $S \rightarrow a$. In particular, the structure in (6) could be produced by a Right Linear Grammar, which only has rules of the form $S \rightarrow aA$ and $S \rightarrow a$. This is significant because of the position of Regular Grammars on the Chomsky Hierarchy: Regular Grammars are the most restricted, and consequently have the lowest time complexity in parsing. Rules for the structure in (6) are given in (7) (where non-terminals are distinguished by the suffix P rather than complete capitalisation).

(6)

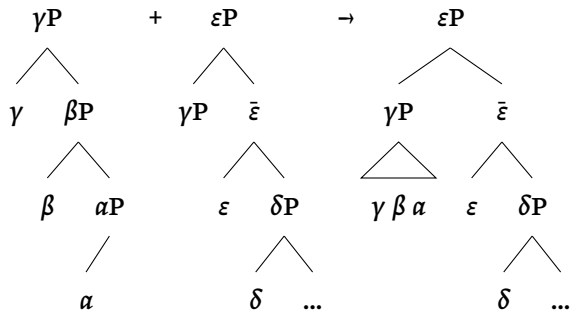


(7) $\{\gamma P \rightarrow \gamma\ \beta P, \beta P \rightarrow \beta\ \alpha P, \alpha P \rightarrow \alpha\}$

Note however that taking c-command to be a primitive following Frank and Vijay-Shanker (2001) does not automatically lead to FSA-equivalent structures, because specifiers are not ruled out, and FSAs on their own cannot accommodate specifiers as normally conceived: if for example γ were a specifier γP, then the rule $\gamma P \rightarrow \gamma\ \beta P$ would need to be replaced with $\beta P \rightarrow \gamma P\ \beta P$—but rules of the form $S \rightarrow AB$ would (minimally) require a Context-Free Grammar and a corresponding Pushdown Automaton. But adopting a Context-Free Grammar with rules of the form $S \rightarrow AB$ would predict as much left-branching as right-branching in phrase structure, since both options are available to the system. Not only is this emphatically not the case for specifiers, but abstracting away from movement it may also be the case for complements, if Kayne

(1994) is right. This is especially relevant given approaches which posit significantly lengthened clausal spines such as Cartography and Nanosyntax, since the asymmetry in the directionality of branching significantly increases, resulting in sections of the phrase marker with no left-branching, able to be modelled with Regular Grammars and Finite State Automata. So discounting specifiers, empirical evidence and processing concerns would suggest that the kind of structure in (6) may be on the right track.

(8)



Accommodating specifiers is a challenge because according to the definition in (5) they must appear to the system to be simplex nodes, in order to be related by asymmetric c-command to other terminals in the main clausal spine. This is overcome in Kayne's (1994) original approach (and also in Frank and Vijay-Shanker 2001) by utilising the segment-category distinction introduced by May (1985) and Chomsky (1986), but the distinction is highly stipulative. Uriagereka (2011) proposes instead that a phrase marker involving complex specifiers is actually linearised in multiple cycles, known as Multiple Spell-out. Specifiers are linearised from most to least embedded, and after each cycle the specifier linearised becomes opaque. The linearised specifier is then inserted into a progressively larger structure by matching the root node of the specifier in question with the corresponding terminal node in the spine which hosts the specifier, as shown in (8), where the triangle labelled (γ, β, α) indicates the opacity. This is formalised by what Uriagereka (2011 p. 121) terms the Linear Correspondence Theorem, given in (9). So in addition to Uriagereka's (2011 p. 84) formulation of the LCA in (5), which is effectively covered by (9a), there is a further step which 'flattens' any specifiers into a single strict total order, given in (9b).

(9) **Linear Correspondence Theorem (Uriagereka 2011 p. 121)**

a. Base:
A phrasal structure K that is weakly equivalent to the output string χ of a Finite-State automaton is externalized by directly interpreting χ as a phonetic string.

b. Induction:
A phrasal structure K that cannot be externalized as in (9a), because its sub-components L and M are weakly equivalent, respectively, to the output

strings χ and ψ of different Finite-State automata, can be externalized by (9b-i) and (9b-ii):

i. applying (9b) separately to χ and ψ, as effectively as possible,

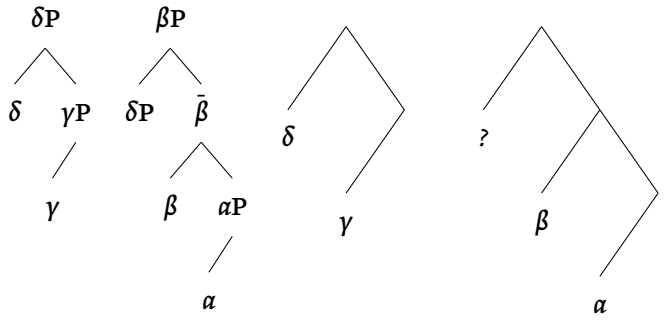
ii. addressing L, linearized as χ, as a tributary current [string equivalent to the output string of a Finite-State automaton embedded in another such string—J.M.] of M, linearized as ψ.

Multiple Spell-Out therefore suggests that the hypothesis that linearisation addresses strings weakly equivalent to Finite State Automata can be maintained, provided that FSAs can be recursively embed within these strings. Now observe that if the system always manipulates the kind of structure shown in (6), as the Multiple Spell-Out approach would suggest, then for terminals hierarchy and order will always coincide with asymmetric c-command, and only the presence of phrasal nodes (and more specifically the apparent need for labelling) supports the traditional distinction between dominance and precedence. On the other hand, if phrase structure is based on a single asymmetric relation, then the phrase marker can be conceived as a single strict order of terminals based on this asymmetric relation, mapping to hierarchy, and through an analogue to (9b-ii) also to linear order. This is set out in the following section.

1.2 Phrase markers as strict orders

As Chomsky (2013 p. 37) has observed, projection or labelling appears to be a theory-internal notion (even if externalisation sometimes manifests a sensitivity to category). One possibility advanced in Chomsky (2013; 2015) is that labelling is determined towards the end of the phase by the Labelling Algorithm, an instance of Minimal Search, so that it is no longer a property inherent to phrase markers. Chomsky (2015 p. 6) then advocates abandoning trees because of the absence of labelled nodes in representations, but for expository purposes I will use an unlabelled tree diagram to represent the FSA-equivalent structures seen above, as shown in (10).

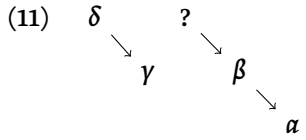
(10)

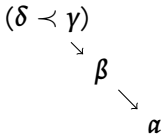


The Labelling Algorithm in Chomsky 2013; 2015 works as follows: given a phrase

comprising a head and a phrase {H, XP}, the head H labels the containing phrase. For a phrase containing two phrases {YP, XP} there are two options: either it is labelled by the closest visible head, Y or X (crucially, copies are somehow invisible to the algorithm), or by a feature common to Y and X. While this solution has interesting consequences in forcing movement and successive cyclicity, specifiers might never label the phrase containing them because, as embedded FSAs, they are opaque to the Labelling Algorithm—much as Chomsky (2013 p. 47) suggests that the root–categoriser distinction leads only categorisers to contribute labels. The first head inside the specifier will therefore not be able to provide a label for the phrase immediately containing the specifier (the labelling of the specifier itself is addressed below). If this is correct, then the root of the unlabelled tree diagram corresponding to βP in (10) would be labelled by β following Minimal Search. The next-highest non-terminal will also be labelled by β, and the following by α.

So if labelling is predictable in this way, and phrase markers therefore contain no labels for non-terminals, are non-terminals strictly necessary? Consider that from (11), the strict order derived from the asymmetric c-command relations in (10), the original structure in (10) could be derived by adding a non-terminal for each node and labelling it using the Labelling Algorithm. Moreover, in adopting structures such as (11), a single abstract asymmetric ordering relation would be axiomatic in phrase structure, and hence (asymmetric) c-command would no longer need to be defined in graph theoretic terms—the abstract ordering relation would then map to order and hierarchy in externalisation. Following Minimalist concerns, it would therefore seem desirable to attempt to dispense with non-terminals in favour of strict orders of terminals, and find a way to derive any necessary distinction between heads and phrases from strict orders themselves.

(11)


(12)


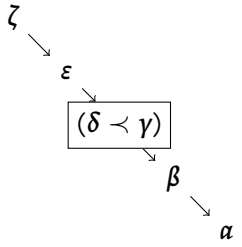
Before doing so, an important but unresolved problem is what Uriagereka (2011 p. 75) terms the Address Issue, given in (13). The difficulty is that if the specifiers are represented as separate orders as in (11), they will need to be composed with the main order, e.g. in Spell-out—but if labelling occurs late in the phase, then labels cannot be used to refer to one order from another. Since there is no restriction on the type of items which stand in an order, one possibility is that the node labelled ? in (11) should have the value of the entire order it represents, as shown in (12), which corresponds to the notion of conservative Spell-out of specifiers in Uriagereka 2011 p. 170. Since this is clearly not a head, the Labelling Algorithm will consider this object opaque, just as in the case of roots.

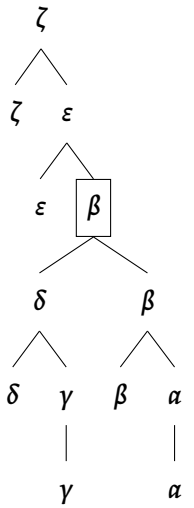
(13) **Address issue**

> Whenever a phrase-marker *K* is divided into complex sub-components *L* and *M*, for *K* to meet LCA conditions of multiple Spell-out, the daughter phrase-marker *M* that spells-out separately must correspond to an identical term *M* within *K*.
>
> (Uriagereka 2011 p. 75)

An interesting consequence of using Minimal Search for labelling as Chomsky (2013; 2015) suggests is that the ultimate root of the phrase marker is not labelled (Chomsky 2015 p. 6)—but since this node by definition will not participate in operations on the phase marker, a label may be redundant—as has been suggested in the past by e.g. Emonds (2004) and Blümel (2017). Extending this idea, since labelling is an instance of Minimal Search, it could occur as needed for any operation requiring labels, rather than just at a particular point in the phase. Given the top-down direction of Minimal Search, the first node given a particular label will be always be the first node of a full phrase. To illustrate this point, consider (14), which shows the strict order and equivalent tree. If the head ζ is searching for some phrase headed by β, it will find $(\delta \prec \gamma)$ first. Since $(\delta \prec \gamma)$ cannot label a phrase, the label from the next node will be used, and the phrase containing $(\delta \prec \gamma)$ and everything following it will be labelled by β, as the tree shows. So the phrase identified by ζ will be a strict order containing all the nodes starting at $(\delta \prec \gamma)$, as expected.

(14)





A major issue with this approach is that the nodes corresponding to specifiers need to have a category e.g. for selection, but this is not clear from their labels, e.g. $(\delta \prec \gamma)$—the label could presumably be provided by Minimal Search starting at the first node of the specifier's own embedded order, but this would require an operation to enter embedded orders to begin the search. Yet if this ability to enter embedded orders were automatic, then the proposed opacity of specifiers when labelling phrases

containing them could not be maintained. Since the ultimate focus here is on weak orders, and since it seems likely that any solution will require the representations posited to be made more complex, I leave this issue open.

Nonetheless it is important to emphasise that this is by no means the first approach to suggest that non-terminals should be abandoned—in fact it has been suggested that Bare Phrase Structure itself, when simplified as far as possible, ultimately leads to a dependency structure, which of course lacks phrasal nodes (for a review of this work see Osborne et al. 2011 and the works cited there). Brody's (2000 pp. 39-41) *Telescope* is essentially similar, although he does not link it explicitly to dependency structures. More recently, Bowers (2018) has argued for an approach to structure based around relations, though his approach does not assume the primacy of strict orders argued for above. The approach detailed so far also implies that Merge is asymmetric, establishing an asymmetric relation between two arguments—this too is by no means a new idea, Zwart (2011) and Kayne (2013) being two rather different approaches implementing asymmetric Merge.

A more unusual feature of this approach is that it implies that the gap between structure in the Narrow Syntax and linearised structure is far smaller than normally assumed. The only function of linearisation under this view is to flatten recursively embedded specifiers into a single total order, linearising an existing order which has non-tail recursion. Because this approach also suggests that, abstracting away from the linearity of linearised structure, there is no distinction between order and hierarchy, then the Narrow Syntax, PF, and LF can all deal with the same representations, which make use of the same abstract asymmetric ordering relation. This effectively allows the requirements of the LCA to hold throughout the derivation as Kayne (1994) originally suggested (and contra e.g. Chomsky 2014/1995), but without any claim that the kind of restrictions the LCA makes on, e.g. branching are due to the need to linearise: it is instead due to the axiomatic use of an asymmetric ordering relation. This relation is moreover far less mysterious than c-command in that it is axiomatic, as opposed to being defined in terms of Graph Theory, and maps to both order and hierarchy and hence is in evidence at both interfaces.

The fact that the syntax then manipulates strict orders based on this relation is also significant. From a biolinguistic point of view, strict orders would seem plausible as psychological realities due to the importance of sequencing both in humans and other species. The particular kind of sequencing that strict orders relate to has been tested experimentally for primates (Samuels et al. 2017 p. 539), suggesting it would at least be available to be recruited by language. And while it might be argued that the ordered pairs representing the relations upon which orders depend can be defined in terms of sets (e.g. Kuratowski's definition: $(a, b) := \{\{a\}, \{a, b\}\}$) and hence that relations are not sufficiently primitive to be used axiomatically in this way, this is only relevant if ordered pairs (and more generally tuples) are non-atomic as psychological realities, which is not a necessary assumption.

But most importantly, it highlights the question of why weak orders do not play any role in syntax, since discounting the distinction between total and non-total orders, orders are either strict or weak. Recalling the difference between asymmetry, $aRb \Rightarrow \neg bRa$, and antisymmetry, $aRb \wedge bRa \Rightarrow a = b$, weak orders may be used to bring about displacement, since symmetry in a weak order would cause two items to appear indistinct. The relation used to derive such orders would be an antisymmetric version of the abstract asymmetric ordering relation described above, only differing with regard to reflexivity and antisymmetry. If antisymmetry (in this non-Kayninan sense) is what allows displacement, then the properties of weak orders ought to create apparent idiosyncrasies, and in the following section I detail the consequential predictions made for possible and impossible displacement.

2 The role of antisymmetry

Beyond allowing the extension of an existing structure using a subpart of this same structure, formalisms modelling displacement such as Internal Merge, Agree, and Autosegmental spreading are substantially different. Despite this, a strikingly similar condition on displacement has been proposed for all three, ensuring that displacement results in sequences of elements which are in some sense contiguous, formalised in syntax in the concept of Feature Geometry-based Relativised Minimality by Starke (2001) (and various other more recent proposals) and Contiguous Agree in Multiple Agree by Nevins (2007), and in Autosegmental Phonology by the Line-Crossing Prohibition (which originates in the Well-formedness Condition of Goldsmith 1976).

In each case, the 'repetition' in some sense of part of the structure leads to a tension between identity and distinctness. These repeated bits of structure, which can be termed occurrences following Chomsky (2004 p. 112), are to some degree identical, given the fact that part of the structure is copied, yet also distinct, since it is possible to individuate specific copies. In the following section, I claim that this tension between identity and distinctness is caused by conflicting ordering information, as Syntactic Objects elements are simultaneously part of a single strict order and multiple weak orders. There is not substantial conflict between orders, but what conflict there is allows items to simultaneously appear identical and distinct depending on the order considered. I firstly clarify the distinction between identity and distinctness of occurrences in displacement, before showing how properties of weak orders allow these facts to be captured.

2.1 Properies of displacement

A similar condition on displacement has developed in three formalisms which model it: in Minimalism Feature Geometry-based Relativised Minimality (Starke 2001) and

Contiguous Agree (Nevins 2007), while in Autosegmental Phonology the Line-crossing Prohibition, originally part of the Well-formedness Condition introduced by Goldsmith (1976). These effects are as follows: in syntax, movement is blocked when a chain of copies which share a particular feature is interrupted by an intervening element with the shared feature, as in (15) (Starke 2001), Multiple Agree searching from a higher probe for a goal with a marked or contrastive feature value cannot skip intervening unmarked or non-contrastive goals as in (16) (Nevins 2007), and in phonology spreading cannot take place across an intervening element already linked to the tier for which the spreading is taking place as in (17) (Goldsmith 1976).

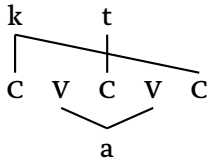
(15) * **[Quant$_1$][Quant$_2$][Quant$_1$]**

How much fun is she not having <how much fun>?
[Quant] [Quant] [Quant]

(16) * **<2,1> with probe search relativisation for marked [Auth] (= [+ Auth])**

v Maria tie- m- a prezentat
Maria 2-dat 1-acc has introduced
[*u*Auth] [– Auth] [+ Auth]
'Maria has introduced me to you.'
(Romanian; adapted from Nevins 2007 p. 297)

(17) * **C$_1$C$_2$C$_1$**



In each case, elements arising from displacement must form a contiguous chain with the source element: for Relativised Minimality displacement must form a contiguous chain of occurrences based on the features causing movement, for Contiguous Agree again a contiguous chain of occurrences must be formed based on the feature involved in agreement, and for autosegmental spreading in phonology, the node which is spreading must be linked to a contiguous stretch of segments of a particular type in another tier, such as consonants in (17). Contiguity in this sense is not with respect to the surface string: rather, it is either maintained that the locality is relativised according to a particular feature, or that locality is in fact strict, and the fact that only certain elements are relevant for the definition of a contiguous chain supports the postulation of distinct tiers of elements, again according to a particular feature. But while the conditions specific to Movement, Agree, and spreading given above do capture these phenomena, the extent to which they are similar suggests a deeper reason for why displacement has this characteristic, which presumably would have to be very basic.

Another, more obvious effect of displacement, is that what appears to be the same element (possibly as minimal as a feature value) appears to exist independently in distinct positions. This would seem to be contradictory, and a range of options exists to account for this phenomenon, most implying it is an illusion at some level: one possibility is a copying operation producing a new element identical to the original, which is the basis of the Copy Theory of Movement, and Chomsky's (2000) approach to Agree. Under this approach, a key issue is just how 'identical' the copies are to the original element. To some extent they must be distinct, since if they were entirely indistinguishable they would be impossible to order relative to one another, and hence impossible to linearise (Nunes 2004 pp. 15-17). But if they were interpreted as entirely distinct, then once the copying operation had occurred, there would be no reason for the items to be construed as related unless they shared some property, such as an index, or were interpreted for some purposes as identical (e.g. interpretation) and for others (e.g. linearisation) as distinct.

Another approach views displacement as a referencing operation, a position taken in Unification within HPSG, and also some approaches to agreement within Minimalism (notably Pesetsky and Torrego 2007). A referencing operation introduces a distinction between the true element (i.e. the source in a copying approach), and the references to this element, but this distinction is not necessarily clearly visible to the system. The Trace Theory of Movement notably adopted this approach, but in this case the 'true element' moved from its original position, leaving a trace as a reference. Traces *were* interpreted as visibly different to the moved element from the perspective of the system for various purposes, but this is not a consequence of the general approach itself. However, in all cases any change to the true element should be always reflected in all of the references (and vice versa), as opposed to a copying approach where copies are basically distinct. A related approach, and the position of Autosegmental Phonology, is that all objects in the system are reached through references, captured in this formalism by association lines between different tiers.

Both these approaches face the same fundamental problem that identity and distinctness are contradictory, but apparently both required—as seen in the way that the copying approach requires indices or variable interpretation, whereas the referencing approach often involves making a distinction between a single true element and the references to account for differing behaviour (e.g. with regards to (non-)pronunciation), or treating every occurrence of the element as an independent reference which can behave autonomously. To further complicate this, consider the Inclusiveness Condition in (18).

(18) **Inclusiveness**

> A "perfect language" should meet the condition of inclusiveness: any structure formed by the computation (in particular, π and λ) is constituted of elements already present in the lexical items selected

> for *N*; no new objects are added in the course of computation apart from rearrangements of lexical properties (in particular, no indices, bar levels in the sense of X-bar theory, etc. ...)
>
> (Chomsky 2014/1995 p. 209)

This rules out capturing the identity aspect of displacement with indices or references, since references will need to be added in the derivation after the true element is added. Variable interpretation would seem a more viable option, though defining the basis for this variability in a principled way poses a challenge. On the other hand, Chomsky (2014/1995 p. 351) suggests that: 'with sufficiently rich formal devices (say, set theory), counterparts to any object (nodes, bars, indices, etc.) can readily be constructed from features'. An issue with using Set Theory to construct counterparts to indices in particular is the danger of inadvertently axiomatising the natural numbers using Set Theory, just as the mathematician Peano did—using these sorts of set-theoretic indices would imply that the system can count, going against the general assumption that if grammars were able to count in this way, there would be less structural dependence in language. But another formal device which could be used to construct a counterpart to the *effect* of indices is Order Theory, and in particular weak orders, given the property of antisymmetry.

2.2 Properties of weak orders

It has already been argued above that phrase structure is based around an abstract asymmetric ordering relation, which leads the syntax to manipulate strict orders of nodes corresponding to terminals in a tree, but at the time no argument was made for selecting a strict order over a weak order. An important property of asymmetric relations is irreflexivity, preventing the same element from appearing twice in a single order based on an asymmetric relation. Assuming asymmetric relations form the basis of phrase structure, apparently identical nodes in strict orders must therefore be illusory, a point I return to below. For now though, asymmetry seems to account for the distinctness side of displacement, since as mentioned above occurrences need to appear to be distinct at some level in order to be linearised relative to one another (Nunes 2004 pp. 15-17). However, it was also established that displacement involves a degree of identity: importantly, antisymmetric relations do not enforce distinctness as they are not irreflexive, and as already mentioned actually enforce identity when symmetric. And since, abstracting away from whether orders are total, they are either strict or weak, the language faculty would be expected to make use of both options absent some principled way of ruling out one or the other.

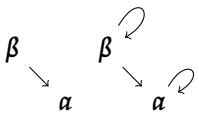
One possibility to resolve this tension between identity and distinctness would therefore be to posit that nodes exist in both strict and weak orders simultaneously, and that whether a node is judged identical to or distinct from another depends on the

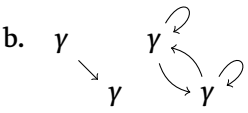
order considered. Nodes will always be judged distinct based on the strict order, due to the irreflexivity of asymmetric relations, but this is not necessarily true of weak orders, which may or may not show nodes to be identical. In other words strict orders exist alongside weak orders throughout the derivation, and the slightly differing behaviour of the two sometimes leads to a tension between identity and distinctness. This is a Minimalist approach in a way that distinguishing D-structure and S-structure is not—the reason why the language faculty would putatively make use of these orders is because they follow from the use of asymmetric and antisymmetric relations, which represent the only two available options, which are equally simple. The distinction is more reminiscent of that between Internal and External Merge, which also represents a case where the Language Faculty makes use of the two available options it is given.

(19) **Ordering possibilities and identity and distinctness**

Operation	Strict order	Weak order
External Merge	Distinct	Distinct
Internal Merge	Distinct	Identical

Indeed, (19) shows how this two-way distinction between orders would yield the difference in the effect of External and Internal Merge on the phrase marker. Crucially, Internal Merge establishes an extra antisymmetric relation when compared to External Merge, and this results in the treatment of identical occurrences as copies. This is shown graphically in (20): for External Merge as in (20a) there is one antisymmetric relation from β to α, whereas for Internal Merge as in (20b) there is both an antisymmetric relation from the first node γ to the second node γ and also from the second node γ to the first node γ. This is the is the full range of possibilities for establishing relations given two occurrences. One will always be ordered relative to the other in the strict order, but it may or may not be the case that the two occurrences will be interpreted as copies based on the weak order.

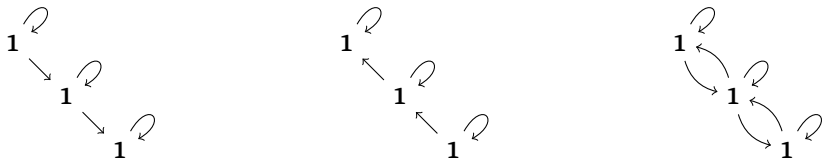
(20) a. 

b. 

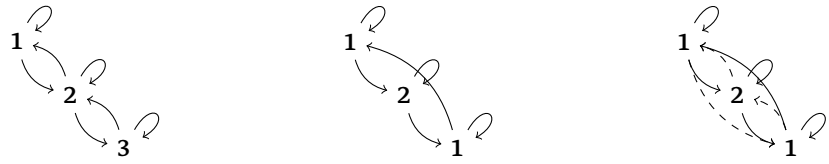
To introduce some shorthand now that two different relations are being used, the precedence relation $\prec$ stands for the abstract asymmetric ordering relation corresponding to asymmetric c-command, (non-linear) precedence, and dominance, so $\beta \prec \alpha$ indicates that β precedes α. $\beta \preceq \alpha$ is the equivalent antisymmetric relation indicating that β precedes or is equal to α. Obviously $\beta = \alpha$ indicates that β equals α. The difference between the $\prec$ and $\preceq$ relations is exactly that which distinguishes the less than ($<$) relation from the less than or equal to ($\leq$) relation, which are asymmetric and antisymmetric respectively.

With this in mind, it should now be possible to characterise chains. In (20b) the symmetric antisymmetric relation between γ and γ means that by antisymmetry $\gamma = \gamma$. This situation corresponds to the usual notion of a chain. But unlike the use of indices, there are limits on what chain formation is possible assuming that chains represent equality relations between nodes in weak orders. Consider first the order $1 \leq 1 \leq 1$ shown in (21a). The inverse of $\leq$ is $\geq$. Clearly it is also true that $1 \geq 1 \geq 1$ as in (21b). So since these relations are antisymmetric and the orders weak, $1 = 1 = 1$ as in (21c). But antisymmetry also appears to rule out a large number of illicit chains, because equality, as opposed to arbitrary indexing, will not allow the formation of chains of non-identical syntactic objects. Thus the order in (22a) is contradictory. Clearly $1 \neq 2 \neq 3$.

(21) a. $1 \leq 1 \leq 1$ b. $1 \geq 1 \geq 1$ c. $1 = 1 = 1$



(22) a. * $1 = 2 = 3$ b. * $1 \leq 2 \leq 1, 1 = 1$ c. * $1 = 2 = 1$



Nor is this the only restriction antisymmetry places on chains. Consider the graph in (22b). This is intended to show that $1 \neq 2$ and $2 \neq 1$ but $1 = 1$, yet by transitivity we also have the relations shown by the dashed lines in (22c). So this would actually show that $1 = 2 = 1$, which is false. Orders must be monotonic: in the case of strict orders this follows from the irreflexivity of asymmetric relations, and in the case of weak orders from antisymmetry. Because of monotonicity there can be no intervener between two nodes in a chain, where an intervener is a syntactic object distinct from the adjacent objects in the relevant sense. In the case of Internal Merge the relevant sense is usually the whole node, in the case of Agree, one or more features (though this will be made more precise below). Hence the properties of weak orders predict that all chains must be contiguous.

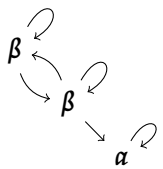
Before showing how this approach works derivationally, it is useful to compare some existing approaches to chains. Under the contextual distinctness approach of Martin and Uriagereka (2014), copies and repetitions are distinguished based on context alone: repetitions within a single phase are impossible, and instead interpreted as copies. For A-chains at least, occurrences across two phases would be interpreted

as repetitions—though note that, if Legate (2003) is right that even A-chains for the subjects of passives and unaccusatives cross a phase boundary, it will be impossible to form A-chains for these. An even more obvious issue for this approach is A-bar chains, since these uncontroversially span multiple phases, yet involve copies, hence in these cases the chain will need to remain after the transfer of a phase. If on the other hand the phase edge is used to allow chains to span phase boundaries, then there needs to be a way to recreate full chains from the partial chains transferred at each phase. Martin and Uriagereka (2014 pp. 175-176) suggest for A-bar chains that existing differences between Internal and External Merge should be extended:

> ...whereas external merge creates a new term, or syntactic object, in a phrase marker, internal merge works differently in that it, in some sense, "stretches" the very same syntactic object across two different syntactic contexts. (Martin and Uriagereka 2014 p. 175)

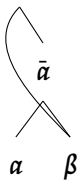
This 'stretching' is what leads to obligatory formation of a chain. This clearly captures the tension between identity and distinctness mentioned above, and is explicitly recognised in the characterisation of Internal Merge: 'there is only one occurrence to speak of, albeit one that exists simultaneously in multiple syntactic contexts'. So here, the difference between copies and repetitions reflects the difference between the effect of Internal and External Merge on the phrase marker (though the authors do not explicitly abandon the phase-based characterisation for A-chains). This corresponds to the distinct effects of these operations on the phrase marker posited in (19). Martin and Uriagereka (2014 pp. 172-173) further suggest that the appearance of a chain is illusory: the system is fooled into interpreting only one syntactic object when two have been presented, because the two objects look identical to each other, and are presented consecutively. Hence a syntactic object is interpreted as distinct if it cannot be collapsed with another adjacent identical object.

This too corresponds to what has been presented above, though in a less obvious way: equality relations between distinct nodes in weak orders such as (21c) imply that the system is fooled into treating distinct nodes as the same because a reflexive relation on a single node, as in the reflexive relation on the first node in (23), and a relation from one identical node to another, as in the relation between the first and the second node in (23) cannot be formally distinguished. Chain formation is therefore automatic, an issue I return to shortly.

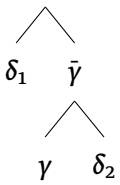
(23) 

Collins and Groat (2018 pp. 2-4) note that Multidominance could also be used to achieve the effect of chain formation. In this case, a single node dominated by multiple mothers (as in (24a) appears in multiple positions in the structure, as opposed to the equivalent structure without multidominance in (24b)).

(24) a.



b.



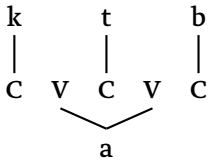
These structures conflate the strict orders and weak orders posited above by attempting to show identity and distinctness in the same structure. As can be seen, this comes at the expense of the distinctness aspect of displacement, since occurrences of β in the structure in (24a) can no longer be distinguished. And as Citko (2011 pp. 135-141) notes, the real hurdle for multidominance structures is linearisation, because such structures cannot be straightforwardly linearised with the LCA. There are various ways to overcome this, some outlined by Citko (2011), and another proposed by Gračanin-Yuksek (2013). Discounting approaches abandoning the LCA, these usually involve modifying either the multidominance structures, or the definition of the LCA. However, fundamentally, the impression created is that multidominance structures are unavoidably awkward to linearise, and the persuasive simplicity of linearisation with the LCA cannot be retained. So given the linearisation difficulties encountered by multidominance due to its emphasis on identity in displacement at the expense of distinctness on the one hand, and the difficulty of showing identity without indices in approaches such as Bare Phrase Structure, on the other, it seems warranted to adopt two distinct representations, weak and strict orders, each accounting for the identity and the distinctness aspect of displacement respectively.

In claiming that strict orders and weak orders coexist in this way, a clear analogy can be made with the multiple tiers of Autosegmental Phonology, as shown in (25)[3]. In this formalism, a timing tier with CV skeleta indicating order of pronunciation of segments can be distinguished from other tiers corresponding to features, in this case vowels and consonants. Nodes from these latter tiers are linked to nodes on the timing tier via association lines, and multiple linking allows displacement to be captured. Here /a/ is multiply-linked to two V nodes, which allows /a/ to be represented as a single morpheme even if it surfaces in two distinct positions in the surface string. Notice also that the association lines between consonants in the consonant tier only link to nodes in the timing tier designated as consonants, and the same for vowels—

[3]Here I abstract away from the use of root nodes, choosing between an X-slot or moraic analysis, and Feature Geometry for expository purposes.

similarly, the weak orders in which syntactic chains form contain subsets of the nodes in the strict order, with membership of each subset based on the presence of a feature.

(25) **/katab/**



The timing tier in (25), made up of CV skeleta able to host features, corresponds to the strict orders above, made up of syntactic objects containing features. The other tiers correspond to the various kinds of chains of structural positions which show Relativised Minimality effects (there are various suggestions for these; for now I follow Rizzi 2001; 2011), such as head chains and chains based on feature classes. The resulting orders are in (26).

(26)

Order	Membership criteria
Strict	None (all nodes)
Weak	Head positions
Weak	Argumental features: Person, Number, Gender, Case,...
Weak	Quantificational features: *wh*, Foc, Neg, Measure, Frequency
Weak	Modifier features: Evaluative, Evidential,... Manner, Measure, Frequency, Neg,...
Weak	Topic

(adapted from Rizzi 2011)

In summary, under this view the existence of Relativised Minimality, Contiguous Agree, and the Line-crossing Prohibition reduces to a weak ordering effect, and the conflicting requirements of identity and distinctness in displacement follow from the possibility for elements to be in at least two orders, which may give contradictory ordering information. Strict orders are used to extend the phrase marker, may be embedded inside each other (at least in syntax), and contain timing information used by linearisation at PF, as well as information used for determining scope and binding at LF. Weak orders on the other hand allow displacement to occur because they permit the formation of equality relations between nodes, and only contain a subset of the nodes in the strict order. Having focused on the representations, it is now possible to show how this approach works derivationally.

3 Schematic derivations

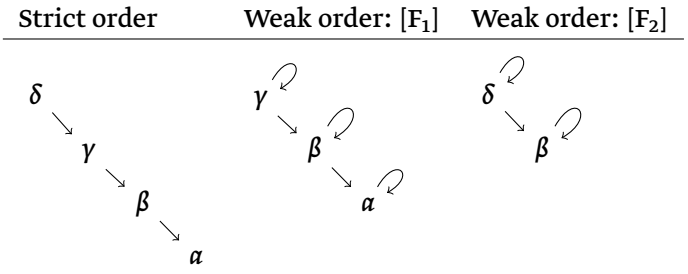
Having focused on the representational side of asymmetry and antisymmetry, it is now time to deal with more derivational issues. The intention is that by adopting the

representations proposed above, the derivation can proceed much as usually conceived. I first show how some schematic derivations would work, beginning with the effect of External Merge and Internal Merge on the Phrase Marker, in particular the fact that chain formation is automatic given weak orders, before turning to Agree. I then consider how Multiple Spell-out and phases impact on the approach developed, and the relationship between cyclicity and Order Theory more generally.

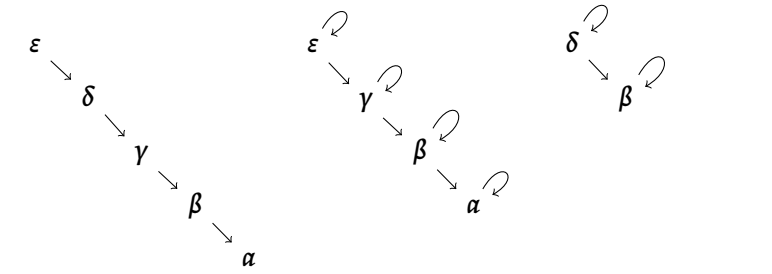
3.1 Operations on the phrase marker

Consider first External Merge. In (27) besides the strict order there are two weak orders, each corresponding to a feature shared by a subset of nodes in the strict order, (e.g. δ and β have $[F_2]$). Suppose as in (28) a new item ε is merged with the existing structure in (27), and that ε has $[F_1]$. Merge will need to establish a precedence relation between ε and δ in the strict order first, and from this the ordering in the weak order can be derived. So in the strict order, ε will come to be adjacent to δ. Since ε has $[F_1]$, it also comes to be adjacent with γ in the weak order for $[F_1]$, as γ was the first node in this order in (27), and because by transitivity ε precedes γ in the strict order. Because ε and γ are distinguishable as nodes, ε is merely ordered before γ as in the strict order. Since ε does not have $[F_2]$, it does not appear in the weak order for $[F_2]$. Linearisation would straightforwardly yield ($\varepsilon \prec \delta \prec \gamma \prec \beta \prec \alpha$) from the strict order, and there is no conflict in ordering between the strict order and any of the weak orders.

(27)



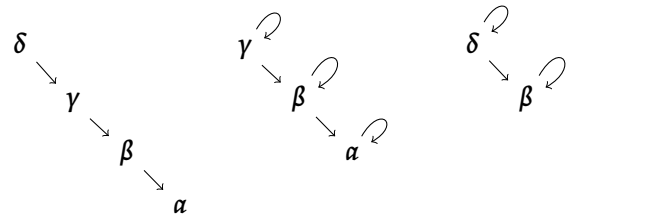
(28)



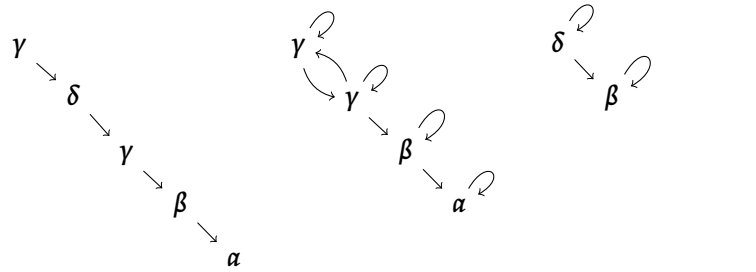
Now consider Internal Merge. Suppose that this time γ, part of the existing structure in (29) is merged with the overall structure, as in (30), and that γ again has

[F_1]. Again, in the strict order, γ will come to be adjacent to δ—assuming for now that in strict orders every node is taken to be distinct to avoid violating irreflexivity. Additionally, since γ has [F_1], it also comes to be adjacent to the other γ in the weak order for [F_1]. But because the two γ nodes in the weak order are identical, the relation from γ to γ in the strict order leads to the vacuous establishment of a relation from (some) γ to (some) γ, i.e. a relation identical to the existing reflexive relation on γ. This results in a symmetric relation in the weak order between the two γ nodes, as shown in (30), as they cannot be distinguished. Importantly, based on the relations and the nodes, the structure in (30) is identical to the structure in (31), because the relation from the first node γ to the second node γ is formally identical to a reflexive relation on a single node γ, and therefore the system perceives only a single γ in the weak order for [F_1] as in (31). So copy formation is obligatory, and arises from a kind of 'collapsing' of the phrase marker, as Martin and Uriagereka (2014 pp. 172-173) suggest.

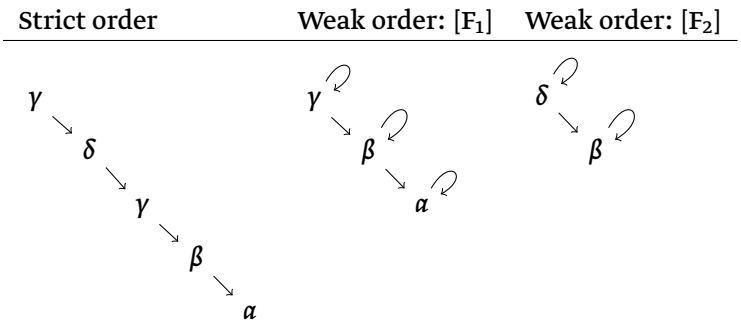
(29)



(30)



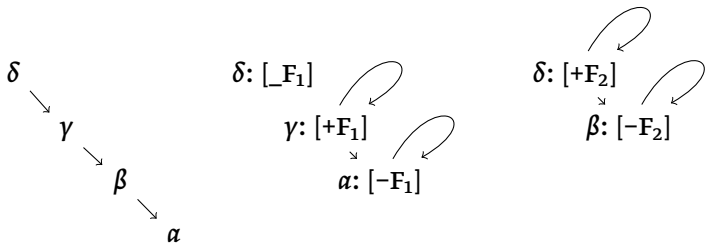
(31)



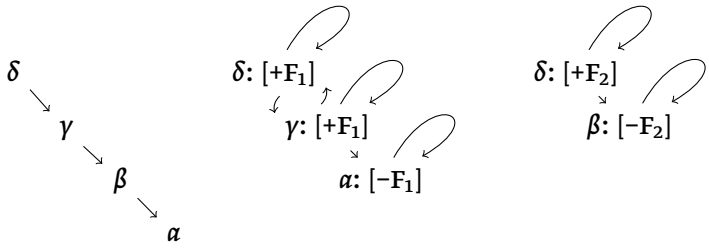
This approach can now be extended to Agree. Unlike External and Internal Merge, Agree does not extend the phrase marker by adding a node to the strict order.

Additionally, whereas Internal Merge creates copies of syntactic objects, Agree can be seen as creating copies of feature values. It is these feature values which stand in weak orders, though for expository purposes the node label also shows the feature which the value corresponds to and the node in the strict order which the feature corresponds to. I also assume following Pesetsky and Torrego (2007 pp. 269-274) that both interpretable and uninterpretable features may be valued or unvalued, and hence when showing feature valuation I do not distinguish interpretable and uninterpretable features. In (32a), δ has an unvalued [F_1] feature, meaning that at this point there is no node for δ in the weak order for the feature value of [F_1] (shown with the underscore). In (32b), the value from γ is copied and, due to identity, an equality relation is established between the copy and the original feature value, which the system interprets as the structure in (32c).

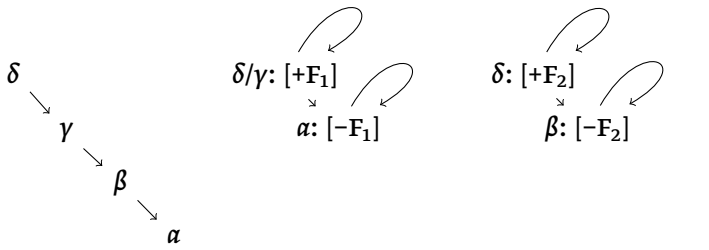
(32) a.



b.



c.



While this example works for downward Agree, where the unvalued probe is higher than the valued goal, upward Agree might seem more challenging, because the insertion of nodes in the middle of an existing order would appear akin to a No

Tampering violation. Consider for example the weak order based on the ≤ relation in (33). Inserting 0 in the order in (33) before the least element 1, as in (34) will require the establishment of a single precedence relation 0 ≤ 1 from which all others (0 ≤ 2, 0 ≤ 4, 0 ≤ 5) can be derived. On the other hand, inserting 3 between 2 and 4 as in (35) will minimally require two more relations, 3 ≤ 4 and 2 ≤ 3.

(33) 1 ≤ 2 ≤ 4 ≤ 5

(34) 0 ≤ 1 ≤ 2 ≤ 4 ≤ 5

(35) 1 ≤ 2 ≤ 3 ≤ 4 ≤ 5

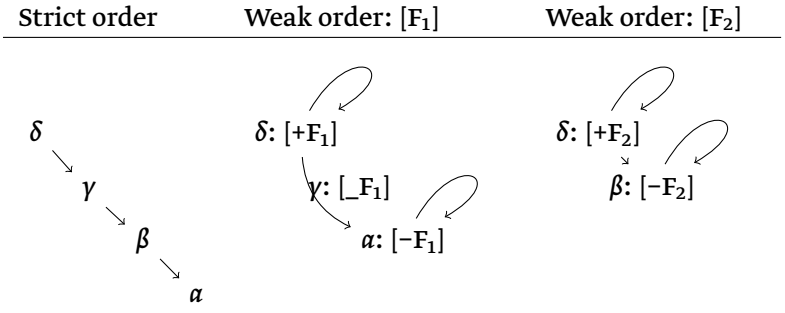
It would not be possible to get a total order if one relation were established—there would either be no ordering relation between 2 and 3, or between 3 and 4, yet neither External nor Internal Merge as defined above establishes more than one relation within a single order. Note however that neither (36) nor (37), where material is copied within the phrase marker, requires the establishment any non-vacuous relations—in (36) the relation 2 ≤ 2 between the two resulting identical nodes follows from the existing reflexive relation 2 ≤ 2, and in (37) the same is true of 4 ≤ 4. The insertion of a node in the middle of a weak order should therefore be unproblematic provided it is a copy of an adjacent node, as expected under Agree.

(36) 1 ≤ 2 ≤ 2 ≤ 4 ≤ 5

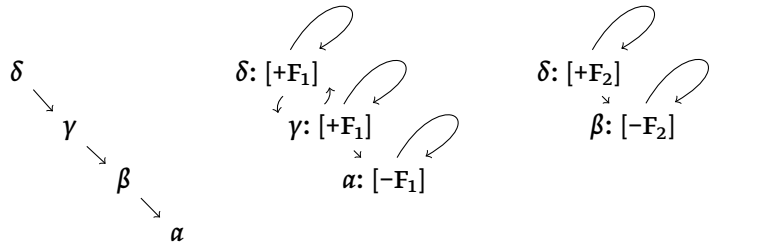
(37) 1 ≤ 2 ≤ 4 ≤ 4 ≤ 5

The situation before upward Agree takes place is illustrated in (38). In (39), copying the value of $[F_1]$ from δ to γ results in these two nodes being indistinguishable, and hence due to identity, an equality relation is established between the values of $[F_1]$ for δ and γ. Note that the relation from $[+F_1]$ to $[-F_1]$ is maintained, but is shown in (39) from the value of $[F_1]$ for γ because the relation from the value of $[F_1]$ for δ can be derived by transitivity. So despite initial appearances, all is well for upward Agree.

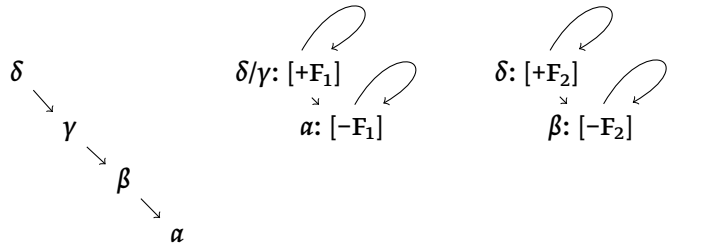
(38)



(39)



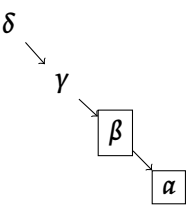
(40)



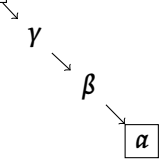
3.2 Embedded strict orders and Phases

Recall that for Internal Merge, which creates identical nodes in strict orders, every node must be taken to be distinct to avoid violating irreflexivity. While nodes could be distinguished based on their relations for example, this would introduce an undesirable distinction from weak orders, which must use node labels to determine identity for the effects on displacement posited above to be derived. However, if specifiers are embedded strict orders as suggested above, then the first instance of phrasal movement will satisfy the irreflexivity requirement through embedding, as shown in (41a). And if Harizanov and Gribanova (2018) are correct in their distinction of syntactic head movement from what they term 'amalgamation' (Harizanov and Gribanova 2018 pp. 24-32) (a PF operation with very different locality requirements), and that syntactic head movement involves movement to a specifier, then the first instance of head movement will also fail to violate irreflexivity, as in (41b). So where material moves off the clausal spine into a specifier, the lack of an asymmetric relation between the two resulting copies satisfies irreflexivity.

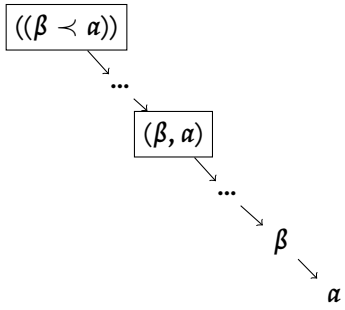
(41) a. $(\beta \prec \alpha)$

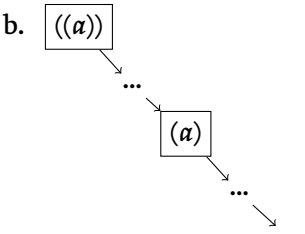


b. (α)



This will only work for cases involving two copies—once identical specifiers appear, then the irreflexivity problem resurfaces. One solution would be to take successive applications of Internal Merge to recursively embedded specifiers, reminiscent of an approach by Groat (2018) to distinguishing repetitions using recursively embedded sets (i.e. $\{\alpha\} \neq \{\{\alpha\}\} \neq \{\{\{\alpha\}\}\}$). This is shown for phrasal movement and head movement respectively in (42a) and (42b).

(42) a. 

b. 

The logic of this approach rests on the fact that the system treats specifiers as simplex nodes, and hence copying a specifier as in (42a) and (42b) is parallel to copying a head as in (41b). One possible concern about Groat's (2018) proposal is that it appears to reconstruct the natural numbers using Set Theory (i.e. $\{\alpha\}$ is α_1, $\{\{\alpha\}\}$ is α_2, $\{\{\{\alpha\}\}\}$ is α_3, ...), and hence distinguishing repetitions on this basis reduces to distinguishing indices. However, this would seem to imply the ability to enumerate and compare the amount of embedding for two occurrences in different parts of a phrase marker—a surprising operation when compared with others which seemingly observe structure dependence and hence do not resort to enumeration. By contrast, under the present approach the fact that specifiers are embedded strict orders follows from the need to treat these structures as non-complex for ordering purposes, and the suggestion that the embedding is recursive for successive instances of Internal Merge from the fact that specifiers are treated as non-complex by the system. The result is that no two items in the same strict order will be identical, satisfying irreflexivity.

While this is able to prevent irreflexivity violations in strict orders, a similar issue arises for weak orders: consider that if the approach taken above for Agree is correct, and features are binary, then the sequences of feature values in (43a) and (43b) are impossible. Clearly, it would be undesirable for this effect to hold over the whole phrase marker. Instead, it would seem more plausible that (43a) and (43b) could hold within phases, and this would appear to be supported by one of the conditions introduced by Nevins (2007 p. 291) on Multiple Agree, namely Contiguous Agree, given in (44). Importantly, Nevins (2007 p. 290) assumes that Multiple Agree, and by extension this condition, respects phase boundaries.

(43)

a. * [−F$_1$] [+F$_1$] [−F$_1$]

b. * [+F$_1$] [−F$_1$] [+F$_1$]

(44) **Contiguous Agree (Nevins 2007 p. 291)**

For a relativization *R* of a feature *F* on a Probe *P*, and $x \in$ *Domain*(*R*(*F*)), ¬∃*y*, such that *y* > *x* and *P* > *y* and $y \notin$ *Domain*(*R*(*F*))
"There can be no interveners between *P* and *x* that are not in the domain of relativization that includes *x*".

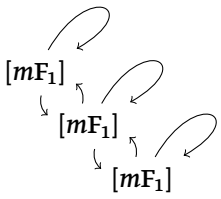
R(*F*) (the relativisation *R* of a feature *F*) accounts for cases where Agree ignores otherwise legitimate goals because their feature values are either unmarked or non-contrastive. The relativisation is supposed to affect the search, though the definition in (44) would seem to indicate that the search is also affected by values outside the relativisation. *R*(*F*) may also include all the values of *F*, and in theory it is possible to require marked and contrastive values in combination. If *R*(*F*) only has marked values, then a violation of (44) will appear as in (45a), where *m* designates marked values and *u* unmarked values, hence [*u*F$_1$] is an intervener. Similarly if *R*(*F*) only has contrastive values, then a violation of (44) will appear as in (45b), where *c* designates contrastive values and *n* non-contrastive values. (46a) and (46b), however, do not violate (44), as the use of weak orders would predict. (47a) and (47b) are obviously also fine. However, (48a) and (48b) both violate (44) because of the enlarged domain of relativization. So under the weak orders approach, (44) should follow from the properties of displacement and need not be stipulated.

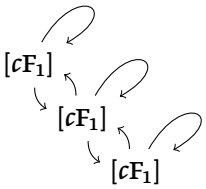
(45) a. * [*m*F$_1$] [*u*F$_1$] [*m*F$_1$]

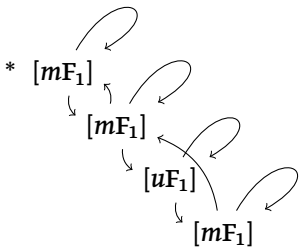
b. * [*c*F$_1$] [*n*F$_1$] [*c*F$_1$]

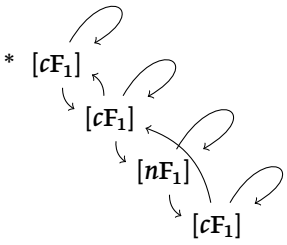
(46) a. [*m*F$_1$] [*m*F$_1$] [*u*F$_1$]

b. [*c*F$_1$] [*c*F$_1$] [*n*F$_1$]

(47) a. 

b. 

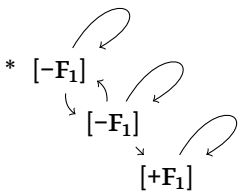
(48) a. * 

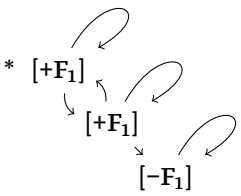
b. * 

The conditions on possible sequences of actual feature values (e.g. + and −) are even stricter, due to the second condition on Multiple Agree introduced by Nevins (2007 p. 291), given in (49). If $R(F)$ includes all the values of F, then (49) prohibits (43a) and (43b), but also (50a) and (50b). While weak orders themselves do not rule out (50a) and (50b), the fact that (43a) and (43b) are ruled out shows that antisymmetry is satisfied, even if vacuously. Hence, the assumption that Multiple Agree respects phase boundaries adopted by Nevins (2007) should also apply relatively straightforwardly to this account, given that the configurations ruled out by weak orders are a subset of those ruled out by the combination of Contiguous Agree and Matched Values.

(49) **Matched Values (Nevins 2007 p. 291)**

For a relativization R of a feature F, $\exists \alpha, \alpha \in \{+, -\}$,
$\forall x, x \in Domain(R(F)), val(x, F) = \alpha$
"All elements within the domain of relativization must contain the same value for the feature F being agreed with".

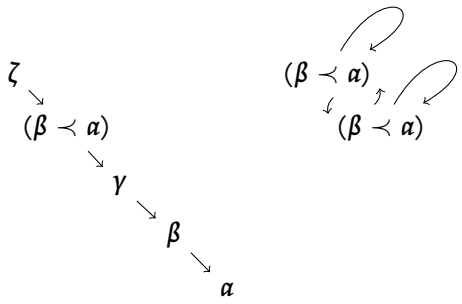
(50) a. * 

b. * 

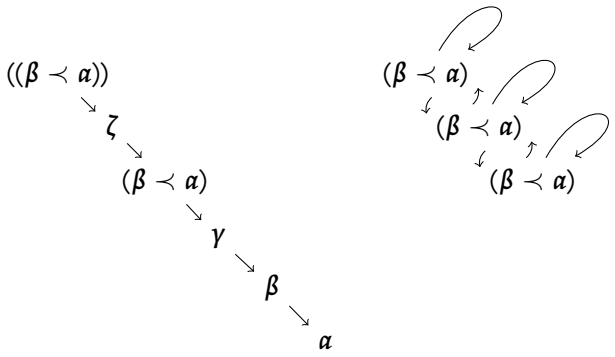
In view of this, whenever a phasal node is reached in the strict order, all succeeding nodes will need to undergo transfer, and the same will be true of corresponding nodes in weak orders. Note that this means that chains only exist within phases, hence successive-cyclic movement involves as many chains as there are cycles. Consider the

situation in (51a), where ζ is a phase head, and hence the elements it precedes undergo transfer at the end of its phase[4]. The elements in the weak order are all phrases, as represented by the embedded orders. $(\beta \prec \alpha)$ needs to move to a position in a higher phase, and hence in (51b) moves to the phase edge. Transfer occurs and the resulting structure is as in (51c)—$((\beta \prec \alpha))$ could then be displaced later, forming a new chain.

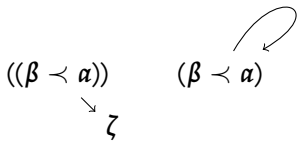
(51) a. Strict order — Weak order: $[F_1]$



b. Strict order — Weak order: $[F_1]$



c. Strict order — Weak order: $[F_1]$



With both the representational and derivational details now in place for asymmetry and antisymmetry, in the next section I show how antisymmetry accounts for the effects traditionally characterised separately as Relativised Minimality, Contiguous Agree, and the Line-crossing Prohibition.

[4]In the following structures I abstract aware from Anti-Locality considerations for expository purposes.

4 Antisymmetry and Contiguity Effects

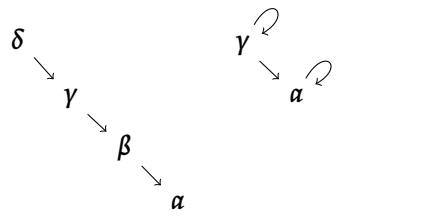
In deriving contiguity effects from antisymmetry, I start with Relativised Minimality, since here the effect of weak orders is broadly similar to what has traditionally been proposed. I then turn to Contiguous Agree, where I assume the previous discussion on the relevance of weak orders, and finally show that the Line-crossing Prohibition can also be derived under this approach.

4.1 Relativised Minimality

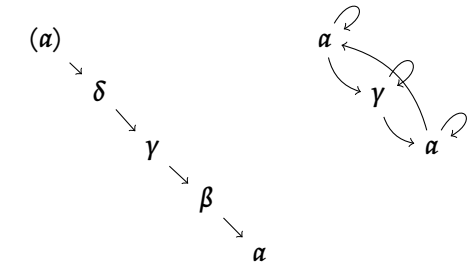
First consider Head Movement, starting from the structure in (52a), where the items in the weak order are all heads. Following Matushansky (2006), syntactic head movement creates specifiers. In (52b) head movement creates a specifier (α) for δ in the strict order, but this situation constitutes a Relativised Minimality violation which can be derived from the fact that the weak order violates antisymmetry (γ is implied to be the same as α). (52c) on the other hand, does not violate antisymmetry, as γ does not have [F_1] and hence is not present in the weak order.

Note that this account does not hinge entirely on Harizanov and Gribanova's (2018) view of Head Movement—what the weak order approach crucially assumes is that copies become distinct in strict orders as copying creates opacity, and that the locality of head movement can be captured in terms of the features relevant to the displacement. Hence, approaches which assume adjunction to minimal projections (which would have to involve embedded strict orders) are not ruled out under this approach, nor is the locality of Head Movement necessarily expected to be uniform (i.e. based on the same features).

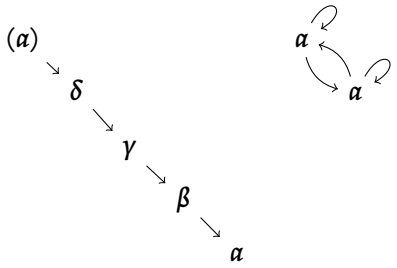
(52) a. Strict order Weak order: [F_1]



b. * Strict order Weak order: [F_1]



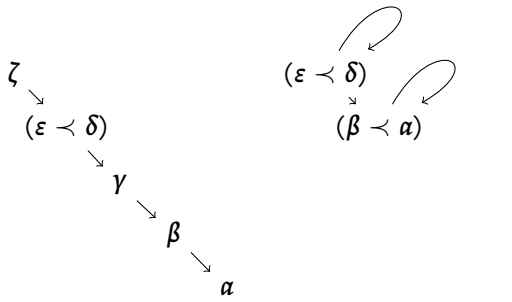
c. Strict order Weak order: [F_1]



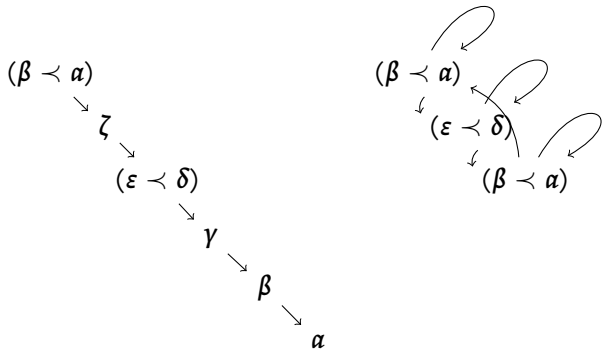
Next consider A- and Ā-Movement, starting from the structure in (53a), where the items in the weak order are phrases: $(\beta \prec \alpha)$ is the phrase starting at β in the strict order, and (ε, δ) the phrase starting at ε in the specifier in that order. The situation in (53b) where $(\beta \prec \alpha)$ becomes a specifier of ζ constitutes a Relativised Minimality violation following from the violation of antisymmetry in the weak order—$(\beta \prec \alpha)$ is implied to be the same as $(\varepsilon \prec \delta)$. If $(\varepsilon \prec \delta)$ were not present in the weak order, as in (53c), there would be no violation.

As should be clear, this is essentially parallel to the situation described for heads above, except that the items in the weak order are phrases rather than heads. What the case of phrasal movement highlights which is perhaps less obvious for head movement is that the nodes in weak orders do not necessarily correspond to nodes in the same strict order—in this case in (53a) $(\varepsilon \prec \delta)$ corresponds to nodes in the embedded strict order for the specifier, whereas $(\beta \prec \alpha)$ corresponds to nodes in the main order.

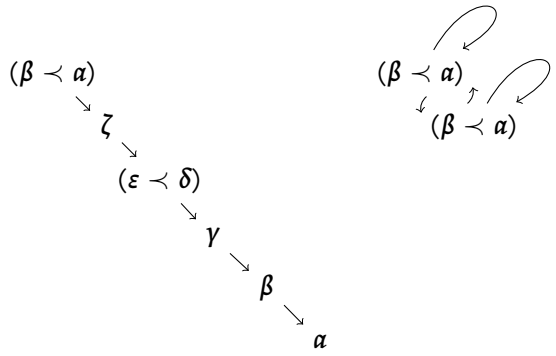
(53) a. Strict order Weak order: [F_1]



b. * Strict order Weak order: [F_1]



c. Strict order Weak order: [F_1]

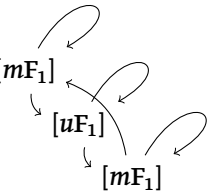


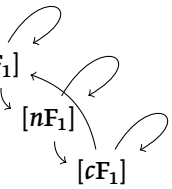
4.2 Contiguous Agree

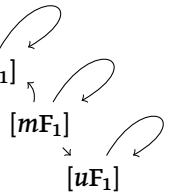
Weak Orders also derive a contiguity effect—Contiguous Agree—in cases where an unvalued probe searches for a valued goal. Where Nevins (2007) parameterises the search from a probe *P*, which may seek all feature values, only marked values, or only contrastive values, this be recast in terms of the kind of values that *P* is able to take. Presumably *P* fails to Agree with its goal if it finds no such values and is given a default value, as Preminger (2014) has argued (and which is also attested in phonology: see e.g. the example of Woleaian vowel harmony in Nevins 2010 pp. 39-45). If *P* must be marked, Contiguous Agree rules out sequences such as that in (45a), repeated below, where *m* stands for a marked value and *u* for an unmarked value. Likewise if *P* must be contrastive, a structure like (45b), repeated below, is ruled out. Both results follow from antisymmetry, because the sequences (−, +, −) and (+, −, +) violate antisymmetry[5], and for a binary feature where one value is marked or contrastive, marked/contrastive and unmarked/non-contrastive values must necessarily form contiguous chains. (46)

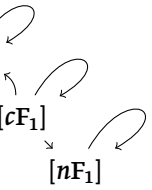
[5]This pattern also appears to occur at other levels of granularity than the phase—the Final-over-Final Condition (Sheehan et al. 2017) under its Kaynian interpretation would, over the extended projection, lack the (−, +, −) and (+, −, +) patterns for the presence of the ^ feature for example (though (+, −) is also ruled out).

shows licit alternatives where the marked or contrastive values are contiguous with the probe.

(45a) * 

(45b) * 

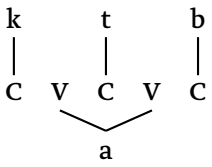
(46a) 

(46b) 

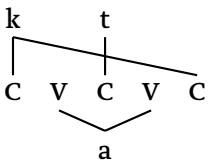
4.3 Line-crossing Prohibition

Finally, the Line-crossing Prohibition in phonology can also be derived from the use of weak orders for displacement. Suppose that strict orders in phonology can be identified with a timing tier (or root node tier in more recent work, e.g. Broselow 1996) and weak orders with the various tiers corresponding to the nodes of phonological Feature Geometry, as suggested above. In this case, it is unsurprising that the Line-crossing Prohibition should exist, since a restriction stating that displaced elements cannot cross over one another if they exist on the same tier, as in (54b) follows from the use of weak orders, as in (55)[6].

(54) a. **/katab/**

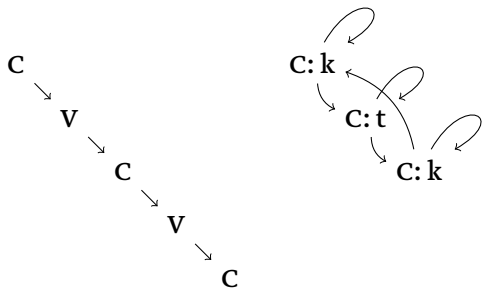


b. * **/katak/**



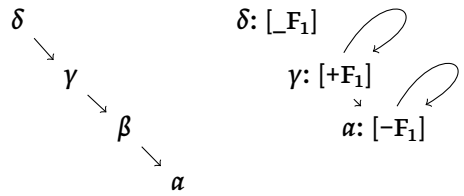
[6]This abstracts away from the problem of the identical nodes in the strict order; determining how best to label these is beyond the scope of the current work.

(55) * Strict order Weak order: [F_1]

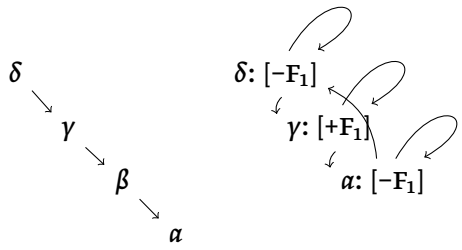


Note how similar these representations are to those used above for Agree. Nevins (2010) has indeed shown that many different kinds of vowel harmony can be understood as instantiating a search from an underspecified vowel for some other vowel, and then copying the value of the vowel found, as Agree does. Violations of antisymmetry which are traditionally captured by the Line-crossing prohibition are shown in (56b) and (57b). Note from the previous discussion that insertion of material into the middle of the weak orders will also be impossible, hence vowel harmony which affects an underspecified vowel surrounded by two other vowels targeted by the Agree operation will always result in the valuation of the central vowel by one or other of the adjacent vowels. And like Multiple Agree, Nevins (2010) shows that it is possible for the search to be relativised to a particular value (e.g. the marked value) of some feature *F*.

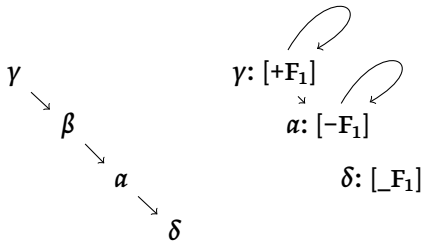
(56) a. Strict order Weak order: [F_1]



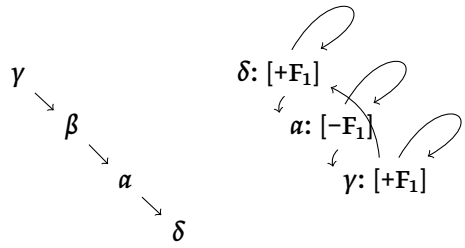
b. * Strict order Weak order: [F_1]



(57) a. Strict order Weak order: [F_1]



b. * Strict order Weak order: [F_1]



Before concluding, it is worth highlighting that the relationship between Order Theory and locality may well go beyond weak orders: since Multiple Spell-out predicts difficulties with subextraction (Uriagereka 2011), the recursive embedding of strict orders is similarly able to account for the opacity of these domains (at least when they *are* opaque). And in general, if Order Theory is fundamental to the system as has been proposed here, its relation to locality certainly needs further investigation.

5 Conclusion

It should be clear from the above that weak orders appear not to have received as much attention as they merit, given that their properties predict fundamental properties of displacement seen in both syntax and phonology. I have also argued that there is evidence to support a relational view of syntax, which, if adopted, resolves some of the architectural asymmetries introduced if weak orders are to model displacement, but strict orders only become relevant during linearisation—instead, phrase structure itself is based around strict orders, which explains among other things the prevalence of c-command. Most importantly however, the adoption of weak orders to model displacement allows a number of strikingly similar but usually separately stipulated locality conditions to be unified without stipulating some new metacondition—Relativised Minimality, Contiguous Agree, and the Line-crossing Prohibition all follow from the use of weak orders in displacement, which itself follows from the use of Order Theory in syntax and phonology.

References

Barker, Chris, and Geoffrey K. Pullum. 1990. A theory of command relations. *Linguistics and Philosophy* 13:1–34. URL http://dx.doi.org/10.1007/bf00630515. *Cited page 6*

Blümel, Andreas. 2017. Exocentric root declaratives: Evidence from V2. In *Labels and Roots*, ed. Leah Bauke and Andreas Blümel. Berlin, Boston: De Gruyter. URL http://dx.doi.org/10.1515/9781501502118-011. *Cited page 11*

Bowers, John. 2018. Relational Derivation. In *Deriving Syntactic Relations*, 6–48. Cambridge University Press. URL http://dx.doi.org/10.1017/9781316156414.002. *Cited page 12*

Brody, Michael. 2000. Mirror Theory: Syntactic Representation in Perfect Syntax. *Linguistic Inquiry* 31:29–56. URL http://dx.doi.org/10.1162/002438900554280. *Cited page 12*

Broselow, Ellen. 1996. Skeletal Positions and Moras. In *The Handbook of Phonological Theory*, ed. John A. Goldsmith, 175–205. Oxford, UK: Blackwell Publishing Ltd. URL http://dx.doi.org/10.1111/b.9780631201267.1996.x. *Cited page 34*

Bruening, Benjamin. 2014. Precede-and-command revisited. *Language* 90:342–388. URL http://dx.doi.org/10.1353/lan.2014.0037. *Cited page 6*

Chomsky, Noam. 1986. *Barriers*, volume 13. Cambridge, Mass: MIT Press. *Cited page 8*

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagareka, 89–155. Cambridge, Mass.: MIT Press. *Cited page 15*

Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In *Structures and beyond*, ed. Adriana Belletti and Guglielmo Cinque, 288. Oxford, UK: Oxford University Press. *Cited pages 4 and 13*

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49. URL http://dx.doi.org/10.1016/j.lingua.2012.12.003. *Cited pages 4, 5, 7, 9, 10, and 11*

Chomsky, Noam. 2014/1995. *The Minimalist Program*. Cambridge, Mass.: The MIT Press. URL http://dx.doi.org/10.7551/mitpress/9780262527347.001.0001. *Cited pages 4, 5, 6, 7, 12, and 16*

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, Strategies and Beyond*, ed. Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 1–16. Amsterdam: John Benjamins Publishing Company. URL http://dx.doi.org/10.1075/la.223.01cho. *Cited pages 4, 9, and 11*

Citko, Barbara. 2011. Multidominance. In *The Oxford Handbook of Linguistic Minimalism*, ed. Cedric Boeckx, 119–142. Oxford, UK: Oxford University Press. URL http://dx.doi.org/10.1093/oxfordhb/9780199549368.013.0006. *Cited page 20*

Collins, Chris, and Erich Groat. 2018. Copies and Repetitions lingbuzz/003809. URL http://ling.auf.net/lingbuzz/003809. *Cited page 19*

Emonds, Joseph. 2004. Unspecified Categories as the Key to Root Constructions. In *Studies in Natural Language and Linguistic Theory: Peripheries*, 75–120. Dordrecht: Springer Netherlands. URL http://dx.doi.org/10.1007/1-4020-1910-6_4. *Cited page 11*

Epstein, Samuel David. 1999. Un-Principled Syntax. In *Working Minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 317–347. Cambridge, Mass.: The MIT Press. URL http://dx.doi.org/10.7551/mitpress/7305.003.0014. *Cited pages 5 and 6*

Frank, Robert, and K. Vijay-Shanker. 2001. Primitive C-Command. *Syntax* 4:164–204. URL http://dx.doi.org/10.1111/1467-9612.00043. *Cited pages 5, 7, and 8*

Goldsmith, John Anton. 1976. Autosegmental phonology. Doctoral Dissertation, MIT, Cambridge, Mass. *Cited pages 2, 13, and 14*

Gračanin-Yuksek, Martina. 2013. Linearizing multidominance structures. In *Challenges to Linearization*, ed. Theresa Biberauer and Ian Roberts, 269–294. Berlin, Boston: De Gruyter. URL http://dx.doi.org/10.1515/9781614512431.269. *Cited page 20*

Groat, Erich M. 2018. ∅: A Minimalist Diacritic. *Cited page 27*

Harizanov, Boris, and Vera Gribanova. 2018. Whither head movement. *Natural Language & Linguistic Theory* URL http://dx.doi.org/10.1007/s11049-018-9420-5. *Cited pages 26 and 31*

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press. *Cited pages 3, 5, 6, 7, 8, and 12*

Kayne, Richard S. 2013. Why Are There No Directionality Parameters. In *Theoretical Approaches to Disharmonic Word Order*, 219–244. Oxford University Press. URL http://dx.doi.org/10.1093/acprof:oso/9780199684359.003.0008. *Cited pages 2, 6, and 12*

Klima, Edward S. 1964. Negation in English. In *The Structure of Language*, ed. Jerry Alan Fodor and Jerrold Jacob Katz, 246–323. Englewood Cliffs: Prentice Hall. *Cited page 6*

Legate, Julie Anne. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34:506–515. URL http://dx.doi.org/10.1162/ling.2003.34.3.506. *Cited page 19*

Martin, Roger, and Juan Uriagereka. 2014. Chains in Minimalism. In *Language Faculty and Beyond: Minimalism and Beyond*, 169–194. Amsterdam: John Benjamins Publishing Company. URL http://dx.doi.org/10.1075/lfab.11.07mar. *Cited pages 18, 19, and 23*

Matushansky, Ora. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37:69–109. URL http://dx.doi.org/10.1162/002438906775321184. *Cited page 31*

May, Robert. 1985. *Logical Form*, volume 12. Cambridge, Mass: MIT Press. *Cited page 8*

Moro, Andrea. 1997. Dynamic Antisymmetry: Movement as a Symmetry-breaking Phenomenon. *Studia Linguistica* 51:50–76. URL http://dx.doi.org/10.1111/1467-9582.00017. *Cited page 3*

Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25:273–313. URL http://dx.doi.org/10.1007/s11049-006-9017-2. *Cited pages 2, 13, 14, 27, 28, 29, and 33*

Nevins, Andrew. 2010. *Locality in Vowel Harmony*. Cambridge, Mass.: The MIT Press. URL http://dx.doi.org/10.7551/mitpress/9780262140973.001.0001. *Cited pages 33 and 35*

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. The MIT Press. URL http://dx.doi.org/10.7551/mitpress/4241.001.0001. *Cited pages 15 and 16*

Osborne, Timothy, Michael Putnam, and Thomas M. Gross. 2011. Bare phrase structure, label-less trees, and specifier-less syntax. Is Minimalism becoming a dependency grammar? *The Linguistic Review* 28. URL http://dx.doi.org/10.1515/tlir.2011.009. *Cited page 12*

Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. Amsterdam: John Benjamins Publishing Company. URL http://dx.doi.org/10.1075/la.101.14pes. *Cited pages 15 and 24*

Preminger, Omer. 2014. *Agreement and Its Failures*, volume 68. Cambridge, Mass.: The MIT Press. URL http://dx.doi.org/10.7551/mitpress/9780262027403.001.0001. *Cited page 33*

Rizzi, Luigi. 2001. Relativized Minimality Effects. In *The Handbook of Contemporary Syntactic Theory*, ed. Mark Baltin and Chris Collins, 89–110. Oxford, UK: Blackwell Publishers Ltd. URL http://doi.wiley.com/10.1002/9780470756416. *Cited page 21*

Rizzi, Luigi. 2011. Minimality. In *The Oxford Handbook of Linguistic Minimalism*, ed. Cedric Boeckx, 220–238. Oxford, UK: Oxford University Press. URL http://dx.doi.org/10.1093/oxfordhb/9780199549368.013.0010. *Cited page 21*

Samuels, Bridget D., Marc Hauser, and Cedric Boeckx. 2017. Looking for UG in Animals. In *The Oxford Handbook of Universal Grammar*, ed. Ian Roberts, 527–546. Oxford, UK: Oxford University Press. URL http://dx.doi.org/10.1093/oxfordhb/9780199573776.013.22. *Cited page 12*

Sheehan, Michelle, Theresa Biberauer, Ian Roberts, and Anders Holmberg. 2017. *The Final-Over-Final Condition*, volume 76. Cambridge, Mass.: The MIT Press. URL http://dx.doi.org/10.7551/mitpress/8687.001.0001. *Cited page 33*

Starke, Michal. 2001. Move Dissolves Into Merge: A Theory of Locality. Doctoral Dissertation, University of Geneva, Geneva. *Cited pages 2, 13, and 14*

Uriagereka, Juan. 2011. *Spell-Out and the Minimalist Program*. Oxford, UK: Oxford University Press. URL http://dx.doi.org/10.1093/acprof:oso/9780199593521.001.0001. *Cited pages 5, 6, 7, 8, 10, 11, and 36*

Zwart, Jan-Wouter. 2011. Structure and Order: Asymmetric Merge. In *The Oxford Handbook of Linguistic Minimalism*, ed. Cedric Boeckx, 96–118. Oxford, UK: Oxford University Press. URL http://dx.doi.org/10.1093/oxfordhb/9780199549368.013.0005. *Cited page 12*