Optimal Linearization:
Prosodic displacement in Khoekhoegowab and beyond

Leland Paul Kusmer
Doctoral Dissertation
(excerpt)

Chapter 2

Linearization

Contemporary syntactic research has largely converged on the Minimalist (Chomsky, 1995, 2005) notion that our theory of syntax should include only those elements necessary for both the interpretive component (the Logical Form) and the articulatory component (the Phonological Form). One consequence of this is that syntactic structures are typically taken to be inherently unordered: The LF is not generally known to be sensitive to linear order, and so our model of grammar should put linearization on the PF branch, after syntactic structure is built. This is not a new idea — the notion that a single syntactic structure might be mapped to different linear orders by different languages lead to the formulation of the Headedness Parameter in early generative inquiry — but under Minimalism the centrality of the linearization problem has increased.

If syntactic trees themselves are inherently underordered, then our model needs to include a function which maps trees to strings. This function should be sufficiently limited to generate all and only the mappings we find in natural language. Put another way, research into linearization is a question of typology: How can we get from syntactic structures to a limited set of possible word orders? This is not our only goal, however: we also seek a model which gives some insight into why we observe the typological patterns that we do.

In this chapter, I'll start by reviewing those typological patterns themselves: What word orders do and do not appear? I'll then turn to reviewing prior approaches to the problem of linearization. These fall broadly into two groups. The first group starts with the classical Headedness Parameter; I'll show, however, that this starting point is not adequate to our typological needs. The second group starts with the Linear Corresondence Axiom (Kayne, 1994a), which aims to be a maximally-restrictive model deriving word order from asymmetric c-command; again, however, I'll show that this model is either much too restrictive or not restrictive at all, depending on one's ancillary assumptions. Finally, I'll briefly anticipate the next chapter by turning to a small class of recent models which have used violable constraints to enforce linearization.

2.1 Empirical word-order typology

Models of word-order typology suffer from too many degrees of freedom. If we want to ensure that our model includes some particular word order, we have at least two options:

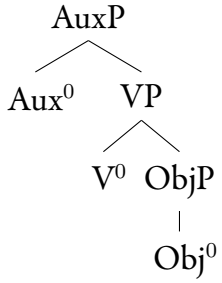
We could hold constant our syntax and tweak our linearization model to produce the desired order; or we could hold constant our linearization model and propose that the word order in question is derived via syntactic movement. In some cases, syntactic research has clearly converged on a movement solution; for example, VSO word orders as in Irish (Chung & McCloskey, 1987; McCloskey, 2011) are almost universally derived via movement of the verb, rather than a linearization scheme which somehow separates the *in situ* verb and object. When setting out to build a model of linearization, then, it behoves us to be conservative in choosing what phenomena we hope to explain: Our goal should be to model only those properties of word order for which syntactic theory does not currently offer any explanation.

One such property is this: recent work has made it clear that the specifier position, insofar as it can be coherently defined on purely syntactic grounds, is always linearized to the left of its head (e.g. Kayne, 1994b; Abels & Neeleman, 2012, a.o). Evidence for this claim includes the universally-leftward direction of WH-movement (e.g Bach, 1971, and many others)[1]; the general paucity of any rightward movement (see e.g. Overfelt, 2015, and references therein); and the rare and often controversial status of OSV base word order. This universal doesn't follow from anything inherent to the syntax — it would change nothing substantial about our theory if specifiers were universally on the right, or alternated based on headedness direction. We should thus hope to find an explanation for this in the linearization function.

The other word order universal which will concern us in this chapter is the Final-Over-Final Constraint (Sheehan et al., 2017): If a phrase is head-final, then its complement will be as well; but if a phrase is head-initial, its complement may have either headedness. This is illustrated with a schematic tree in (1); any part of a tree with the same geometry will have the same word order prediction. If we are allowed to set the Headedness Parameter individually for each phrase, we predict 4 possible orderings for this tree; empirically, though, the Final-Over-Final Constraint (FOFC) rules out the order in which VP is head-initial but AuxP is head-final:

(1) **The Final-Over-Final Constraint:** The complement of a head-final phrase is also head-final.

(2) a.



b.

	Aux Initial	**Aux Final**
V Initial	Aux V Obj (e.g. English)	* V Obj Aux
V Final	Aux Obj V (e.g. Bambara)	Obj V Aux (e.g. Hindi)

[1] The one purported exception to this universal is American Sign Language; however, the data there is highly unclear and the analysis controversial. See e.g. Petronio & Lillo-Martin (1997)

The FOFC has been extensively discussed in the literature, notably in a recent book by Sheehan, Biberauer, Roberts, & Holmberg (2017); evidence for the constraint is presented there and in the references contained therein. I will present a small sample of the evidence here, however, coming from WALS (Dryer & Haspelmath, 2013). WALS does not code directly for the kind of disharmonic orders that interest us here, but it does include a proxy: Feature 94A covers the placement of "adverbial subordinators", a subset of complementizers, with respect to their embedded clause; we can take this as tracking the order of C and its complement S. We can then look at the relationship between these embedding complementizers and the headedness of the language overall (as measured by Feature 95A, "Relationship between the order of Object and Verb and the order of Adposition and Noun Phrase"). The results are tabulated in (3).

(3) **The FOFC in WALS:**

	C S	S C
Head-Initial	258 (87%)	1 (0.001%)
Head-Final	37 (13%)	85 (99.99%)

As can be seen, languages in which a head-final C embeds an otherwise head-initial clause are vanishingly rare[2], with only one such language listed in WALS.[3] This provides evidence for only a small subset of the range of cases covered by the FOFC, and the reader is directed to the existing literature for exemplification of the other cases. Nonetheless, it can be seen that the FOFC is at least a very strong trend and likely a universal.[4]

These, then, are the typological facts we should target when designing a linearization function: Specifiers always precede heads and their complements; and the complements of head-final phrases must also be head-final.

2.2 The Headedness Parameter

The classical approach to linearization is the Headedness Parameter, which hypotheses a parameter controlling whether heads occur on the left or the right of their phrase. This hypothesis fails to account for either of the empirical generalizations above. First, the Headedness Parameter offers no explanation for why specifiers always precede their heads. Put another way, head-final languages are fully head-final in that the head does occur at the right edge of its phrase; but head-initial languages the head is preceded by the specifier. The Headedness Parameter offers no explanation for this striking asymmetry; we are left to simply stipulate that the parameter applies only to heads and complements, but not specifiers.

The Headedness Parameter model also fails to capture the typological facts; depending on one's assumptions, it either wildly undergenerates or wildly overgenerates. The undergeneration case is commonly known: If we assume that the headedness parameter

[2] Difference of proportions: $\chi^2 = 227.8$, $\mathrm{df} = 1$, $p < 0.0001$.

[3] The one language listed is Buduma (Lukas & Nachtigal, 1939), a Chadic language.

[4] The low percentage of C S languages which are head final in this data (13%) is a sampling artifact — head-final languages are under-represented in Feature 94A generally. Note that the disharmonic case does comprise 30% of the head-final languages in this sample.

can't be set for individual heads (but rather is global to the entire language), we predict that all phrases in a language will have identical headedness. As we've already seen above, this is easily falsifiable: German is a frequently-studied example of a language with mixed-headedness; casting our net a bit more broadly, WALS (Dryer & Haspelmath, 2013) lists 66 languages in which the relative ordering of the verb and its object differs from the ordering of adposition and noun. This is a small percentage of the sample, to be sure, but it represents only one of the ways that a language might display mixed-headedness; whatever model we use, it clearly must rule in these mixed cases.

On the other hand, if we allow languages to set the Headedness Parameter differently for each individual phrase type, we wildly overgenerate. In particular, we will fail to capture the FOFC: If parameter settings are independent for heads, a final-over-initial configuration is just as likely as an initial-over-final one. The Headedness Parameter is thus a wholly-inadequate model for linearization.

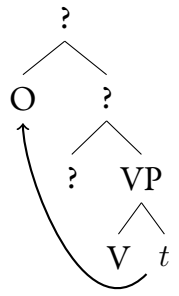
While this model has been largely abandoned in recent work, there are still a few models that follow similar lines. One example is Wouter Zwart (2011), which proposes that the Merge structure-building operation is asymmetric, generating ordered pairs; while he does address the FOFC, it is still unclear what would prevent his system from switching order in a non-FOFC-respecting way.

2.3 The Linear Correspondence Axiom

Kayne (1994b) proposes the Linear Correspondence Axiom, which states that asymmetric c-command in the syntax is directly mapped onto precedence in the linearized string. This has the immediate benefit of explaining why specifiers are always on the left: The specifier always asymmetrically c-commands the head, and so everything in the specifier must precede the head. The cost, of course, is that the LCA rules out head-finality entirely: Heads always asymmetrically c-command the contents of their complements, and so under this model will always precede them. On first glance, then, the LCA vastly undergenerates: Of the three FOFC-compliant word orders, it seems to predict only one.

In order to escape this prediction, Kayne himself proposes that apparently head-final orders are in fact generated by movement. For example, Object-Verb word order might be generated by some kind of object raising, as in (4):

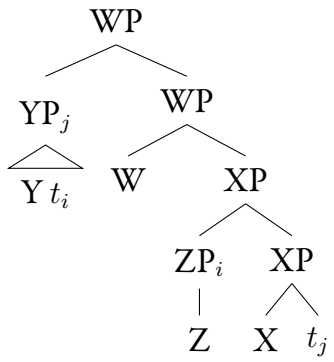
(4)



Of course, getting the entire clausal spine to be head-final then requires a sequence of roll-up movements: The object above VP, then the VP above TP, etc. These movements generally have no independent support, and it's not clear what could motivate them. Worse,

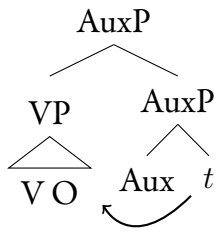
though, is that allowing this kind of movement in fact renders the LCA completely non-restrictive. For example: A large part of the original motivation for the LCA was to derive an apparent ban on rightward movement: If movement is always to a c-commanding position, by the LCA it must always be leftward. But, as Abels & Neeleman (2012) point out that, given the option of remnant movement (for which we generally have independent motivation), it is perfectly possible to generate an LCA-compliant structure giving the appearance of rightward movement:

(5)



Worse, for our purposes, is that it is quite easily possible to generate the missing final-over-initial disharmonic word order, without even requiring remnant movement:

(6)



The LCA, then, despite its apparent success in providing an explanation for the asymmetry of specifiers, ultimately suffers the same fate as the Headedness Parameter: Depending on how much poorly-motivated syntactic movement one is willing to countenance, it either vastly undergenerates or vastly overgenerates, with no obvious way to arrive at a happy medium.

Most contemporary approaches to linearization use the LCA as a starting point and thus inherit its flaws. For example, Fox & Shoichi (2005) propose that linearization proceeds cyclically by phase, with each new phase adding asymmetric c-command relations (and thus precedence) to the order; crucially, they propose that this process is monotonic — once an order has been established between two words, it cannot be changed. This allows them to derive successive cyclicity and other restrictions on syntactic movement. This gives us considerable new insight into these restrictions, but fails to address the underlying typological issues with the LCA. Dobashi (2009) similarly shows why phase-based linearization requires the phase edge to remain accessible for later syntactic processes, but still accepts the one-to-one correspondence between asymmetric c-command and precedence. Collins & Stabler (2016) ignores asymmetric c-command, but still posits a universal specifier-head-complement order. With all its flaws, the LCA

thus continues to reign as the state-of-the-art approach to linearizing syntactic structure.

2.4 Violable Linearization

To anticipate the next chapter a little, it's worth taking a look at a few linearization schemes which make use of violable constraints to model cases where non-syntactic factors seem to adjust the linearization. Morphophonology has used a variety of violable constraints to order morphemes since the introduction of Optimality Theory (Prince & Smolensky, 1993/2004); various morphosyntax analyses have adopted this for clitic ordering (Legendre, 1998, e.g.). These analyses generally share in common that the portion of the input of interest — generally the heads or features which are spelled out as clitics — are unordered in the input and are subject to ALIGN constraints (McCarthy & Prince, 1994a) which try to position them relative to some edge. I'm not aware of any attempt to extend this style of analysis to cover the full range of linearization, however.

More directly relevant here are approaches which assume an order-enforcing constraint is in conflict with other constraints not related to word order. An early example of this is López (2009), who proposes that the LCA itself is a violable constraint in competition with various prosodic constraints. He uses this approach to explain Clitic Right Dislocation in Romance, arguing that the apparent rightward movement is in fact leftward movement to an intermediary position, but that a prosodic constraint requiring the verb to phrase together with its extended projection overrides the LCA and causes the moved item to be linearized on the right. Similarly, Elfner (2012), in analyzing Irish pronoun postposing[5], uses an LCA constraint penalizing deviation from spec-head-comp order; in Bennett et al. (2016) this constraint is softened into NOSHIFT, which penalizes deviation from some order, determined from the syntax by a deliberately unspecified algorithm.

In fact, all three of these proposals effectively assume that the linear order is part of the input. All three analyses define their respective constraints as penalizing deviations from some linear order, rather than from a mapping between syntactic structure and linear order. That is, the linearization constraint in question is effectively a "string edit distance" function, which measures the difference between two strings by how many individual swaps you would have to make to get from one to the other. These approaches, in other words, give us no added insight into how we get the preferred word order in the first place; they merely give us a mechanism for allowing linearization and prosody to interact. In chapter 5 I'll propose an alternative that makes use of violable constraints more extensively in order to give us some additional insight into the linearization function.

[5] See section 3.2.2 for a more detailed discussion of the phenomenon.

Chapter 3

Prosodic Displacement

In the last chapter, I surveyed the prior approaches to linearizing syntactic structure. While there is a great deal of variety in these approaches, they share in common a restriction on what information is available to the linearization function. In particular, they restrict the linearization function to seeing *syntactic* information, i.e. constituency and labelling (as opposed to e.g. phonological form). While this restriction is generally left implicit, it follows from a view of grammar in which linearization takes place in the narrow syntax, or at best at the interface in which syntactic form becomes phonological form.

There is a growing body of evidence that this restriction may not be tenable. For example, consider the case of Irish pronoun postposing as discussed by Elfner (2012) and Bennett et al. (2016). In Irish, some unstressed pronouns may be postposed arbitrarily late in the clause, shown in (1). If these pronouns are stressed, however, they must be pronounced in their base position.

(1) Fuair sé ___ óna dhearthair an lá cheana é
get.PAST he from.his brother the-other-day it
"He got it from his brother the other day."

This is a case in which the phonology of the pronoun seems to affect its linearization: The linearization function treats pronouns with a particular phonological property differently from those without that property. In a model where the linearization function only has access to the syntactic structure, this would be impossible to account for. If we're going to capture cases like Irish, we need to extend our model.

The Irish pronoun postposing phenomenon is a case of what I will term *prosodic displacement*. I'll use the word 'displacement' generally to refer to all those linguistic phenomena in which some constituent seems to have more than one position — for example, being pronounced in a different position than it is interpreted, or being interpreted differently in multipile positions. We can immediately distinguish at least three classes of displacement: overt syntactic movement, covert LF movement, and PF displacement. PF displacement (or PF movement) has a long history in the literature; see, for example, Chomsky (1995); Aoun & Benmamoun (1998); Sauerland & Elbourne (2002); Embick & Noyer (2001). I'll use the term prosodic displacement to refer to a subset of PF displacement which is apparently conditioned by the phonological or prosodic properties of the displaced item and its context, rather than some condition on the syntax-phonology

interface. Since syntactic & prosodic theory both have heretofore assumed that the linearization function only saw syntactic structure, both have ignored the possibility of prosodic displacement as a systematic phenomenon.

In this chapter, I'll argue that phenomena like Irish pronoun postposing, which show a word-order alternation deependent on prosody, force us to consider a prosodic displacement analysis. I'll start by proposing a set of criteria we can use to diagnose PF displacement generally and prosodic displacement in specific. With these criteria in hand, I'll review the previously-proposed cases of prosodic displacement, building evidence that whatever linearization function we choose it must have access to prosodic information; I'll also consider a number of proposed cases of PF displacement which are excluded by these criteria. In the next chapter, I'll introduce a new and particularly-extensive case of prosodic linearization from Khoekhoegowab.

3.1 Diagnosing prosodic displacement

Before I can propose an analysis of prosodic displacement, we need clear criteria for identifying when it occurs. That is, say we have some word-order alternation: A particular word (or class of words) is pronounced in one position in one context, but a different position in another context. We already have one clear mechanism for deriving such an alternation, namely syntactic movement; what could motivate us to provide a prosodic displacement analysis for a given alternation instead of a syntactic movement analysis?

In what follows, I will strive to be conservative in what I analyze as prosodic displacement. It's entirely possible that some phenomena which have previously been understood as syntactic movement would be better analyzed as prosodic displacement, but for the time being it seems wise to only include those phenomena which have no reasonable syntactic analysis. Most of the criteria proposed here then are concerned not so much with ruling *in* prosodic displacement phenomena but with ruling *out* phenomena which the narrow syntax could easily explain. That is, the first three criteria are concerned with selecting those word-order alternations for which only a PF displacement analysis is available; the fourth and final criterion selects for a prosodic displacement analysis specifically.

With that in mind, the criteria I will use for diagnosing prosodic displacement are as follows:

1. **Syntactic implausibility**: We should prefer a PF analysis if the displacement involved violates commonly-accepted generalizations about syntactic movement.

2. **Semantic inactivity**: We should prefer a PF analysis if the displacement involved fails to affect the compositional semantics of the utterance.

3. **Morphosyntactic heterogeneity**: We should prefer a PF analysis if the contexts in which displacement occurs show no uniform morphosyntactic property.

4. **Prosodic homogeneity**: We should prefer a PF analysis if the contexts in which displacement occur show some uniform prosodic property.

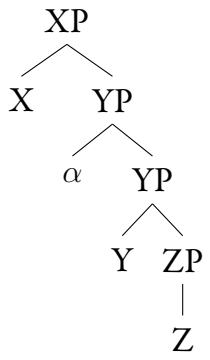
The rest of this section will discuss these criteria in greater depth; the rest of the chapter will be devoted to seeing how these criteria apply to specific examples of prosodic displacement from the literature.

3.1.1 Syntactic implausibility

The first criterion for identifying PF displacement is a basic one: Can syntactic movement easily generate the proposed structures? 'Implausibility' is fairly subjective, so it's worth our while to specify at least some of the ways a given alternation might be implausible; to do this, we need to enumerate some of the typically-assumed properties of syntactic movement.

For one, syntactic movement is typically assumed to be **monotonic**. In fact, in most contemporary syntactic theory movement is assumed to always be 'upwards', i.e. towards less embedded positions, as in (2). There are a limited class of cases which have sometimes been analyzed as lowering (for example, English affix hopping) though it is not clear that we should analyze these as a syntactic phenomenon at all. Even allowing for syntactic lowering, however, it would be extremely surprising to see a single phenomenon which moved some constituent either up or down. For example, imagine that the structure in (2) was sometimes pronounced as (3-a) and sometimes as (3-b), as though α had sometimes raised to spec,XP and sometimes lowered to spec,ZP; this would be a highly implausible candidate for a syntactic movement analysis, as the movement would need to be non-monotonic.

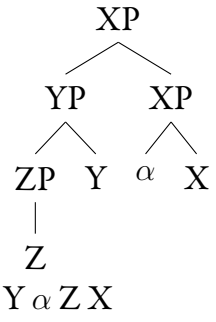
(2)



(3) a. α X Y Z
b. X Y α Z

Another kind of syntactic implausibility involves **locality**. Movement is known to be subject to various locality conditions, both inter- and intra-linguistically defined. An easy example of a locality condition is islandhood: If a particular structure is known to be an island for otherwise-uncontroversial syntactic movement, we should regard a particular displacement as implausible if it apparently does not respect this island. Islandhood is a locality condition on the origin of movement, but we can also point to locality conditions on the landing site of movement. For example, consider the syntactic structure in (4-a) when pronounced as (4-b); the object α has apparently been displaced into the constituent YP. If a particular syntactic movement seems to land inside a higher constituent, we should regard it as non-local.

(4) a.



b. Y α Z X

3.1.2 Semantic inactivity

A displacement is a good candidate for a PF analysis if it has no semantic effect. While syntactic movement does not always create changes in the compositional semantics of the sentence, it at least always has the option to. PF displacement, by virtue of occurring derivationally after the hand-off from the narrow syntax to the interpretation and pronunciation portions of the grammar, should not have this option.

However, it should be noted that this criterion pertains only to to *compositional* semantics. That is, PF displacement should not change how the denotation of the displaced item is combined with the denotation of the rest of the clause; for example, it shouldn't allow DPs to gain new theta-roles, or quantifiers to change scope. But this still allows for the possibility that the different word orders will be associated with different pragmatic meanings. We'll see in considering the Irish pronoun postposing case that this is an important caveat.

3.1.3 Morphosyntactic heterogeneity

The final criterion for ruling out syntactic movement as an analytic option is morphosyntactic heterogeneity: the contexts in which a displacement is observed should not form a morphosyntactic natural class to the exclusion of those contexts where the displacement is not observed. For example, WH-movement is morphosyntactically homogeneous in that it targets only and all items with WH-features; it would be very unusual if certain WH-items were immune to movement, or if certain non-WH-DPs also underwent the same movement.

3.1.4 Prosodic homogeneity

The final criterion is also the only one which specifically picks out prosodic displacement, rather than PF displacement generally: the contexts in which the displacement is found should form a prosodic natural class excluding the contexts in which the displacement is not found.

3.2 Prior examples of prosodic displacement

Cases of prosodic displacement which meet all four of the criteria laid out here are rare. To some degree, this is likely because these criteria were deliberately constructed to be very conservative in what phenomena would count. To my knowledge, only three clear cases have been described so far:

- Halpern (1992); Schütze (1994); Bošković (2001), & many others: second-position clitics in Serbo-Croatian (and related languages) sometimes interrupt syntactic constituents, but never interrupt prosodic ones.
- Elfner (2012); Bennett et al. (2016): Irish pronouns sometimes postpose when unstressed, but never when stressed.
- Edmiston & Potsdam (2016): Malagasy complement clauses are displaced to the right, unless they consist only of a single phonological phrase.

In this section, I'll lay out the evidence for each of these phenomena, evaluating them against the criteria proposed above. Afterwards, I'll review a few more cases where PF displacement has been proposed which fail to meet these criteria.

3.2.1 Second-position clitics

The Balkan languages, particularly Serbo-Croatian, have perhaps the best-studied case of prosodic displacement, namely its second-position clitics (2PC s), which have a long literature going back to at least (Halpern, 1992; Legendre, 1998, and others). These are clitics in the phonological sense of prosodically-dependent light items which nonetheless don't show the behavior of affixes. The class of 2PC s includes items from an array of morphosyntactic categories, including auxiliaries, a question particle, the reflexive markers, and various non-nominative pronouns. All of the clitics in a given sentence cluster in the second position of the clause, regardless of the other word order (which is fairly free):[1]

(5) Taj pesnik mi je napisao knijgu.
That poet me Aux written book
"That poet wrote me a book."

What has typically attracted researchers to a prosodic displacement analysis for 2PC s is the first criterion: Syntactic implausibility. In particular, the clitics do not always follow the first XP as in (5); they can alternatively follow the first word as in (6). This frequently results in clitics apparently interrupting other constituents. Schütze (1994) and others call these the '1P' (1st phrase) and '1W' (1st word) positions; this alternation is illustrated further below.

(6) Taj mi je pesnik napisao knijjgu.
That me Aux poet written book
"That poet wrote me a book."

[1] Unless otherwise noted, all Serbo-Croatian examples are taken from (Schütze, 1994).

(7) a. Prošle godine su otvorili ugostiteljsku školu
last year AUX open hotel-and-catering school

b. Prošle su godine otvorili ugostiteljsku školu
last AUX year open hotel-and-catering school
"Last year they opened a hotel-and-catering school."

Generating any of these word-orders by syntactic movement would involve either moving into an already-built XP or would involve a variety of unusual extractions from inside those XPs. Serbo-Croatian 2PC s thus meet the first criterion — there does not appear to be a plausible syntactic movement analysis. The second criterion, semantic inactivity, is also easily met: None of the descriptions of the phenomena find any difference in compositional meaning between the 1P and 1W positions.[2] Likewise, the third criterion is easily assessed: The 2PC s themselves, comprising everything from a question particle to pronouns, do not form any morphosyntactic natural class that would exclude all those morphemes which do not obligatorily appear in second position. Neither do the sentences showing the 1W position form a natural class — the first word may be from (nearly) any morphosyntactic category.

Evidence that Serbo-Croatian 2PC s meet the fourth criterion, prosodic homogeneity, comes from the few cases in which the 1W order is not grammatical. For example, the clitic cluster may not come between (most) prepositions and their arguments:

(8) a. Na sto ga ostavi.
on table it leave
"Leave it on the table."

b. *Na ga sto ostavi.
on it table leave

Schütze (1994) nodes that the relevant generalization seems to be that the host item to the left of the clitic cluster must be a prosodic word, not just any syntactic terminal. Prepositions like *na* 'on' seem to be proclitics themselves insofar as they do not receive an independent accent and thus do not constitute their own prosodic words. Percus (1993) notes that there are some prosodically-heavier prepositions that do have the accentual properties of prosodic words and can, at least marginally, host clitics:

(9) ?Okolo je sobe trčao Marko.
around AUX room run M.
"Marko runs around the room."

This, then, is prosodic homogeneity: The 1W position always has a prosodic word to the left of the clitic cluster. Serbo-Croatian thus meets all four criteria for prosodic displacement. This fact has been well-recognized in the literature, if not in precisely the terms presented here; for example, Halpern (1992) proposes a PF operation of "prosodic inversion" which reorders a clitic and a potential host in order to satisfy the prosodic needs of the clitic. This operation works well enough for the Serbo-Croatian case, but we will

[2] Though see discussion in Schütze (1994) on the factors which condition the selection of 1P or 1W position.

see that it has little to say for the other cases of prosodic displacement discussed in this chapter. In chapter 7 I will return to this issue and propose an analysis which allows us to unify the Serbo-Croatian case with the other examples.

3.2.2 Irish pronoun postposing

Elfner (2012), expanded by Bennett et al. (2016), show that Irish light object pronouns often appear far to the right of where object DPs would generally be expected, with no detectable difference in semantic or pragmatic import. For example, in (10) the pronominal object appears after the clause-final adjunct:[3]

(10) Fuair sé ___ óna dheartháir an lá cheana é
get.PAST he from.his brother the-other-day it
"He got it from his brother the other day."

Bennett, Elfner, & McCloskey (2016) present convincing evidence that this displacement lacks the signature of a syntactic movement process. First, the displacement is highly syntactically implausible. For example, compare the example in (10) with the example in (11). In (10), the object pronoun seems to be raising in that it is displaced past a variety of adjuncts, including the temporal adjunct *an lá chearna* 'the other day'. In (11), by contrast, a light expletive subject pronoun has seemingly been lowered into the middle of the conjoined predicates. This is an example of non-monotonicity of movement — the same displacement phenomenon apparently moves an item either up or down. Additionally, the example in (11) involves displacement into a coordinate structure; if this were syntactic movement, it would seemingly violate the Coordinate Structure Constraint Ross (1967).

(11) is cuma ___ 'na shamhradh é nó 'na gheimhreadh
COP.PRES no.matter PRED summer it or PRED winter
"It doesn't matter whether it's summer or winter." (Bennett et al., 2016, 183)

On the criterion of semantic inactivity, Bennett, Elfner, & McCloskey (2016) show quite convincingly that even within the same syntactic structure pronouns may freely displace to a variety of syntactic positions with no difference in meaning. It should also be clear from the previous two examples that conditions under which postposing occurs are morphosyntactically heterogeneous: Pronouns displace from both object and (some) subject positions, and can land in a variety of locations.

The last criterion, prosodic homogeneity, is satisfied by the fact that pronoun postposing affects only light, stressless pronouns; stressed pronouns obligatorily appear in their base position. The relevant generalization governing postposing, then, is a prosodic one; this is a clear case of prosodic displacement.

3.2.3 Malagasy clausal extraposition

Edmiston & Potsdam (2017) argue that clausal extraposition in Malagasy takes place at PF. This is particularly interesting in that it is the only clear case of prosodic displacement

[3] All Irish examples are drawn from Bennett et al. (2016).

I am aware of which affects items heavier than a phonological clitic. Clausal extraposition is cross-linguistically quite common and is typically given a syntactic movement analysis; however, Edmiston & Potsdam (2017), expanding on Law (2007), give quite compelling evidence that the Malagasy case must be post-syntactic.

Malagasy shows VOS default word order (Keenan, 1976). However, most embedded clauses obligatorily extrapose to the right edge of the clause:[4]

(12) Nividy (fiara vaovao) Rabe (*fiara vaovao)
PST.buy car new Rabe car new
"Rabe bought a new car."

(13) Manantena (*fa hividy fiara aho) Rabe (fa hividy fiara aho)
hope that FUT.buy car I Rabe that FUT.buy car I
"Rabe hopes that I will buy a car."

Initially, this right-extraposition seems syntactically plausible. However, there is a language-specific test which applies here. Malagasy typically only allows matrix subjects and some adjuncts to be extracted. Objects and constituents inside objects cannot be extracted. Descriptively, objects are islands for extraction (Keenan, 1976, 1995).

(14) a. Iza no hividy boky?
who FOC FUT.buy book
"Who will buy a book?"
b. *Iona no hividy Rabe?
what FOC FUT.buy Rabe
Intended: "What will Rabe buy?"
c. *Momba iona no hividy boky Rabe?
about what FOC FUT.buy book about Rabe
Intended: "What will Rabe buy a book about?"

Nonetheless, CPs can and must move from within a complex object, as illustrated in (15). Clausal extraposition thus violates an otherwise-unviolated generalization about Malagasy syntax, making a syntactic movement analysis implausible.

(15) Nanambara ny faniran-dRabe (*fa hanambady ny faravaviny aho) Rasoa
PST.reveal DET desire-Rabe that FUT.marry DET daughter I Rasoa
(fa hanambady ny faravaviny aho)
that FUT.marry DET daughter I
"Rasoa revealed Rabe's desire that I marry his daughter."

Turning to the second criterion, Edmiston & Potsdam (2017) present a wide variety of arguments that extraposed CPs are always interpreted in their base position within the VP[5]. That is: VOS word order in Malagasy is achieved by VP fronting; CP complements to the verb (or the object) are universally interpreted as though they are still within the VP. I'll present only their argument from NPI licensing here: Negation in Malagasy

[4] All Malagasy examples are taken from Edmiston & Potsdam (2017).

[5] Edmiston & Potsdam (2017) talk about a "Predicate Phrase" for reasons not germane to this discussion; I'll use VP for clarity.

is expressed with a preverbal particle *tsy*; this particle is unable to license NPIs in subject position (16), implying that these NPIs are strong in the sense of requiring strict c-command by negation (Zwarts, 1998). However, NPIs within embedded clauses continue to be licensed (17), showing that they are being interpreted within the VP. This, along with other arguments from binding and Principle C, establish that clausal extraposition is semantically vacuous.

(16) *Tsy nanongo an' i Koto n'iza n'iza
NEG PST.pinch ACC Koto anyone
Intended: "No one pinched Koto."

(17) Tsy mino aho [fa marary velively izy]
NEG believe I that sick at.all he
"I don't believe that he's sick at all."

For the final two criteria, Edmiston & Potsdam (2017) direct us to those CPs for which extraposition is optional, which they term *degenerate* CPs. For example, in (18) the embedded subject has been elided due to Topic Drop; this CP can optionally remain *in situ*.

(18) Milaza [fa nahita gidro tany an-tsena Ø] Rabe
say that PST.saw lemur LOC PREP-market Rabe
"Rabe says that he (Rabe) saw a lemur at the market."

Other degenerate clause types include controlled clauses, existential clauses, and subject relative clauses. There is no morphosyntactic natural class which would contain only those clauses that obligatorily extrapose and exclude the degenerate ones. There is, however, a clear prosodic natural class: Edmiston & Potsdam (2017) show that degenerate clauses are exactly those in which the subject is null, which allows the entire clause to form a single phonological phrase. Malagasy phonological phrases robustly show a distinctive final rise, and most clauses show one phonological phrase for the VP and another for the subject; degenerate clauses show only one at the end of the VP. The authors argue that intonational phrases in Malagasy are preferentially binary; without a subject in the degenerate clause, there is a preference to downgrade it from an intonational phrase to a phonological phrase. The result is prosodic homogeneity: Clauses that obligatorily extrapose are exactly the ones which constitute intonational phrases.

3.2.4 Other proposals

The four criteria discussed here are quite restrictive in what phenomena will count as prosodic displacement. There are several other proposals for prosodic (or PF) displacement analyses where the phenomena in question do not meet the criteria for inclusion here. This is not to say that a PF displacement analysis is not correct for those cases — merely that such an analysis is not strictly necessary. These proposals fall broadly into three classes. First, there are 'PF movement' analyses aiming to account for (mostly) syntactically plausible movement which have no apparent semantic effect. Second, there are proposals which aim to provide an alternative account for displacement effects which are syntactically plausible and semantically active, but nonetheless seem to lack an obvious

syntactic motivation. Finally, there is a growing literature suggesting ways that prosodic or otherwise phonological effects might mediate between the choice of different syntactic structures. I'll briefly enumerate a few of these attempts below, in order to explain why they will not be taken up in this dissertation.

PF Movement

Commonly-cited PF movement phenomena include for instance Aoun & Benmamoun (1998); Sauerland & Elbourne (2002) on total reconstruction, or Chomsky (1995, 2005); Göbbel (2007) on clausal extraposition in English. Both of these examples start with a syntactically plausible and homogeneous movement that nonetheless seems to have no semantic effect; a PF displacement analysis is thus appealing, but not necessary. A distinct but related class comes from the Distributed Morphology literature (e.g. Embick & Noyer, 2001), which proposes a set of movement-like operations that occur after the narrow syntax; while the phenomena accounted for using these techniques are sometimes syntactically implausible and often semantically inactive, the are often not prosodically homogeneous, instead apparently being driven by morphological features. Again, this doesn't rule out prosodic displacement analyses, but still allows other options.

Alternative analyses to syntactic movement

The second category comprises PF displacement accounts of phenomena for which there are already syntactic analyses; in particular, these displacement phenomena are syntactically plausible and semantically active, so adopting a PF displacement analysis would rely on showing that it gives some general benefit over a syntactic one.

First, López (2009) proposes that Clitic Right Dislocation in Romance is the result of prosodic pressure overriding syntactic pressure when determining word order. In particular, he argues that there is pressure to phrase the verb together with its extended projection; this forces certain adjuncts, which would otherwise disrupt that phrasing, to be displaced out of the way. This analysis is highly interesting, but the phenomenon is still amenable to a purely-syntactic analysis. The implausibility of analyzing right dislocation as syntactic movement relies entirely on the assumption that rightward movement is never possible, which may not be warranted (see e.g. Overfelt, 2015, and references therein); furthermore, Clitic Right Dislocation does in fact change the binding possibilities of the moved item, showing that this movement is not semantically inactive. Lopez's analysis relies on syntactic movement to a middlefield position being prosodically marked in that it would separate the verb from the rest of its extended projection, but that markedness is not a property of the prosody itself: Rather, he argues that this structure is marked because it fails to maintain a certain syntax-prosody relationship. That is, the relevant prosodic homogeneity here is not phonological in nature — it requires us to know something about the syntax in order to evaluate whether it is, in fact, homogeneous. This is a highly interesting proposal, but fails to meet the criteria for inclusion here.

A second such example comes from Clemens (2016), followed by Clemens & Coon (2016). Here, the target phenomenon, observed Niuean and some Mayan languages, is the VSO / VOS word order alternation termed 'pseudo noun incorporation': While

VSO is the default word order, syntactic & semantic properties of certain objects trigger VOS surface order. Contra the standard syntactic analyses of this phenomenon, Clemens argues that this is due to a constraint ARGUMENT-φ, which requires that heads and their arguments occupy the same phonological phrase. Under this analysis, those objects which escape incorporation are exactly those that are headed by phasal D (rather than being bare NPs), causing them to be spelled out before their argument relation to the verb can be established. Clemens also shows convincingly that the VOS order does have a distinctive prosody, meeting the criterion of prosodic homogeneity. However, this phenomenon fails to meet any of the other criteria. The syntactic movement needed to generate the observed orders is VP movement either preceded or not by extraction of the object, which is a well-established and supported analysis (Coon, 2010, see e.g.). Pseudo noun incorporation does have a distinctive semantic effect, which is exactly the change Clemens & Coon are trying to capture by proposing that the incorporated objects are NPs rather than DPs. And again, like López (2009), this proposal relies on a marked syntax-prosody relation, rather than simply a marked prosodic structure, to motivate displacement. Once again, the PF displacement analysis is insightful, but not strictly necessary to capture the target phenomenon.

Phonology mediating choice of structure

Finally, there is another class of proposals in which prosodic factors seem to mediate the choice of syntactic structure. For example, Anttila et al. (2010) show that phonological markedness factors into the choice between the double-object and prepositional frames for English ditransitives; Shih & Zuraw (2017) show from a corpus study that phonological markedness plays a role in the selection of Noun-Adjective or Adjective-Noun order in Tagalog; Breiss & Hayes (2019) finds that bigrams which produce phonologically-marked clusters at the word boundary are systematically under-represented in English. Another interesting case comes from Weir (2015), who shows that English fragment answers seem to involve A′-movement that isn't possible in the absence of ellipsis; he argues that the requirement that focused items be stressed allows actual syntactic movement of the focused item in order to escape ellipsis.[6] All of these phenomena share in common that phonology seems to mediate between sentences that have different underlying syntactic structures. This is not PF displacement I mean it here, as it cannot be accounted for by proposing a different linearization scheme for a single syntactic structure. These facts require a significantly more complicated model in which the phonology is able to give feedback to the syntax during the process of sentence-construction. Such an idea is intriguing, but well beyond the scope of this dissertation.

[6] Weir argues that the fragments move at PF, not in the narrow syntax. However, the movement in question has all the properties of syntactic movement, meaning that it cannot be an effect of the linearization function, all that's relevant to the current discussion.

Chapter 4

Prosodic Displacement in Khoekhoe

Khoekhoegowab, commonly called Khoekhoe, is a Central Khoisan language spoken in Namibia by around 200,000 speakers, making it the largest language in the Khoisan group. It is a language of considerable syntactic interest, but the particular phenomenon of relevance here is the unusual positioning of tense, aspect, and polarity particles. Some but not all of these particles, which otherwise behave like the heads of their respective phrases, show up before the verb, despite Khoekhoe being an otherwise head-final language; what's more, these preverbal particles can show up in a wide range of positions, apparently without semantic or pragmatic effect:[1]

(1) a. Nesi =b ge ǁna xamma ne ǂnū gomasa ni nâ.
now =3MS DECL that lion this black cow FUT bite

b. Nesi =b ge ǁna xamma ni ne ǂnū gomasa nâ.
now =3MS DECL that lion FUT this black cow bite
"Now that lion will bite this black cow."

I will show in this chapter that preverbal TAP particles meet all the criteria for prosodic displacement proposed in chapter 3:

- TAP particles are frequently displaced to syntactically-implausible landing sites.
- The position of the TAP particle has no semantic consequence.
- The class of preverbal TAP particles is morhosyntactically heterogeneous.
- But the preverbal particles are prosodically homogenous both in the form of the particle and in their effects on sentential tone.

[1] Unless otherwise noted, all Khoekhoe examples are drawn from original fieldwork. Examples are presented in the standard orthography except where tone is relevant to the discussion, in which case the four tone levels are marked as <a̋, á, à, a̗> from superhigh to superlow. The following abbreviations are used: DECL — declarative clause type marker; -OBL — oblique case marker; TAP — tense, aspect, & polarity particle; IMP— imperfect aspect; PERF — perfect aspect; NEG.NF — negative non-future; NEG.FUT — negative future. ϕ-feature markes will be glossed as PERSON — GENDER — NUMBER; for example, 3CD indicates 3rd person, common gender, dual.

4.1 Basics of Khoekhoe syntax

Khoekhoe is, in general, a strongly head-final language. Looking first at the clausal domain, verbs are final; light verbs follow their main verbs; and embedding complementizers robustly follow the clause they introduce (Haacke, 2006).

(2) Arib ge |hôasa ra saru.
dog DECL cat IMP chase
"The dog is chasing the cat."

(3) ǂKhanisa =ta ge ra khomai ǂgau.
book =1S DECL IMP read want
"I want to read the book."

(4) a. Mî =ta ge ra [arib ge |hôasa go mû ti.]
say =1S DECL IMP dog DECL cat PST see C.QUOT
"I'm saying that the dog saw the cat."
b. Axab ge [!garise ra ā se] ra !khoe.
boy DECL loudly IMP cry C.ADJ IMP run
"The boy is running while crying loudly."

Turning to the nominal domain, we find that all DPs end with a ϕ-feature-bearing enclitic, which also encodes specificity; I take this to be the D^0 head. All nominal modifiers precede the noun, including demonstratives, which behave generally identically to adjectives. Completing the picture, adpositions follow their complement.

(5) a. khoe =b
person =3MS
"the man"
b. khoe =di
person =3FP
"the women"
c. khoe =i
person =3CS
"some person"
d. khoe =khom
person =1MD
"we two men"

(6) a. ne ǂnū goma =s
this black cow =3FS
"this black cow"
b. ǁna ti !nona |ho =n
those my three friend =3CP
"those three friends of mine"

(7) ǁîb om =s |kha
his house =3FS to
"to his house"

Khoekhoe shows a range of second-position clitics which divide the clause into a prefield and a middlefield. Most prominently, root clauses typically have a second-position 'clause type' particle indicating the speech-act; only the declarative marker *ge* is obligatory, but 'emphatic' or echo questions are marked with *kha*, while 'emphatic' declaratives are marked with *kom*.

(8) a. Netse =b ge Dandagoba ni ǂna.
today =3MS DECL D. FUT dance
"Today Dandago will dance."
b. Netse =b kha Dandagoba go ǂna?
today =3MS ECHO D. PST dance
"Dandago danced today? (echo / surprisal)"
c. Dandagob kom ǁkhawa ra ǂna o.
D. EMPH again IMP dance C.EMPH
"Dandago really is dancing again."

The prefield, which I take to correspond to a specifier position in the CP layer of the clause, is typically occupied by the subject. However, topicalized constituents may be raised there, leaving the subject stranded in the middlefield. When this happens, a second-position clitic tracking the ϕ-features of the subject obligatorily precedes the any clause-type marker. This is shown for sentences with second-position clause type markers in (8-a-b); (9) shows that the subject clitic appears even when there is no (overt) clause type marker.

(9) Netse =b Dandagoba go ǂna?
today =3MS D. PST dance
"Did Dandago dance today?"

4.1.1 TAP particles

Given the otherwise-head-final word order, it is striking that tense is frequently marked by a particle in preverbal position.

(10) Khoeb ge oms |kha go oa.
man DECL home to PST return
"The man went back home."

All tense, aspect, and polarity (TAP) information in Khoekhoe is expressed with a set of particles,[2] which are often fusional across those three domains of meaning. Most of these particles appear preverbally as in (10); some, however, appear after the verb:

(11) Khoeb ge oms |kha oa tama.
man DECL home to return NEG.NF
"The man didn't go back home."

Notably, there are contexts where the preverbal particles like *go* 'past' may occur after the verb. First, in some cases it is possible or even preferable to front the verb and its tense marker into the prefield. When this happens, the TAP particle obligatorily follows the verb no matter which class it belongs to:

[2] These TAP particles appear to be phonological enclitics, as evidenced by the fact that the imperfect marker shows allomorphy based on the final consonant of the word it encliticizes to: *ta* after consonants, *ra* otherwise.

(12) Khomai go =b ge Dandagoba ǂkhanisa.
read PST =3MS DECL D. book
"Dandago read the book."

Additionally, in certain embedded clauses it is possible to scramble the verb to the left of the TP; in these cases, the particle again obligatorily (immediately) follows it:

(13) ǁAmaxu ra netsē ǀapa ǂkhanisa ti ǀhōs ge.
sell IMP today red book my friend DECL
"It's my friend who's selling the red book today."

By contrast, there are no circumstances under which a postverbal TAP particle like *tama* 'negative non-future' can precede the verb:

(14) *Ne taras ge !haise tama !gû.
this woman DECL quickly NEG.NF walk
"This woman doesn't walk quickly."

I take the TAP particles to be the (sometimes fused) heads of TP, AspectP, and PolarityP, analogous to auxiliaries; I'll continue to refer to them with the neutral term 'particle'. These facts make attractive an analysis in which all heads in the clausal spine are head-final, including T^0 and the other heads expressed by the TAP particles, but where some process causes certain particles to be displaced to a preverbal position. I'll argue that this displacement is postsyntactic and in fact has all the hallmarks of prosodic displacement.

4.2 First criterion: Syntactic implausibility

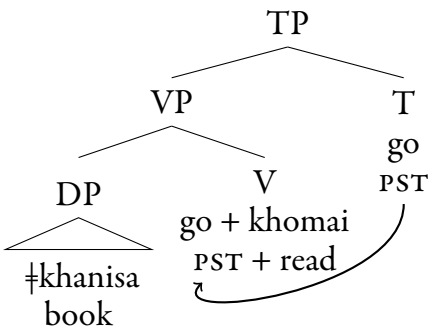
Above, I gave evidence that, other than the case of preverbal TAP particles, Khoekhoe is uniformly head-final, motivating an analysis in which the preverbal particles achieve their position by some kind of displacement. It's worth taking a bit to consider whether this displacement could possibly be syntactic movement. I will argue in this section that preverbal TAP placement does meet the first criterion for identifying prosodic displacement: If we were to understand it as syntactic movement, it would be movement with a highly unusual signature.

4.2.1 First possibility: Lowering

The first possible syntactic movement analysis we must consider is the simplest one: Perhaps the preverbal TAPs themselves move into a preverbal position. Under the standard assumption that T^0 is higher in the clausal spine than V^0, this would be the result of lowering:[3]

(15)

[3] For ease of exposition, I'm going to proceed as though all TAPs originate in T^0, ignoring aspect and polarity heads unless they are specifically relevant. The same arguments given in this section would apply to TAPs originating in separate Asp^0 or Pol^0 heads, both of which are generally assumed to be higher than V^0.



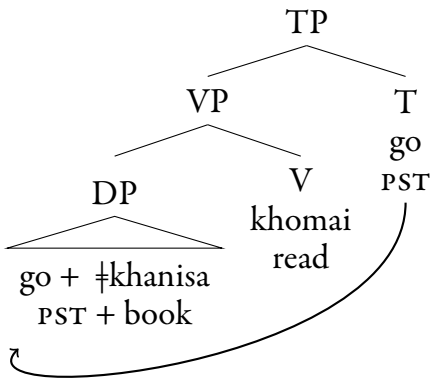
Lowering has a controversial status in syntax. In modern syntactic theory it is commonly assumed that phrasal movement only goes upward. Head movement is generally also treated as proceeding monotonically upwards, with the possible exception of certain kinds of post-syntactic operations (e.g. affix hopping in English). The case of lowering in (15) could plausibly be of this second sort, i.e. postsyntactic lowering of T^0 onto V^0. It would be the only case of prefixing affixation in Khoekhoe, but perhaps the TAP 's status as a clitic rather than an affix can explain this difference.

This analysis becomes impossible to maintain, however, in light of additional data: preverbal TAPs are not always immediately preverbal. It is possible, though rare, for the particles to appear earlier in the middlefield, separated from the verb by at least one other XP:

(16) Dandagob ge go ǂkhanisa khomai.
D. DECL PST book read
"Dandago read the book."

This no longer has the signature of post-syntactic lowering: The TAP particle would be lowering to attach to an arbitrary phrase.

(17)

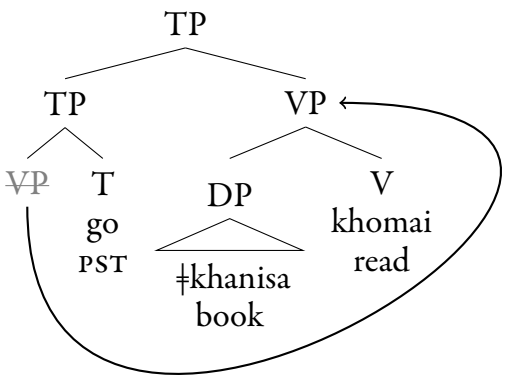


Given the controversial status of lowering in contemporary syntactic theory, it would seem unwarranted to extend it to cover the sort of movement depicted in (17). As such, we'll dismiss the lowering analysis.

4.2.2 Second possibility: Raising

The inverse of the lowering analysis, in which the verb raises to T^0, is also made implausible by the fact that preverbal TAP particles can be separated by the verb by arbitrary XPs in the middlefield, as shown in (16). A better raising analysis involves not head-movement but phrase-movement: Under this analysis, the VP (or some arbitrarily-large phrase containing the verb) would raise and right-adjoin to TP:

(18)

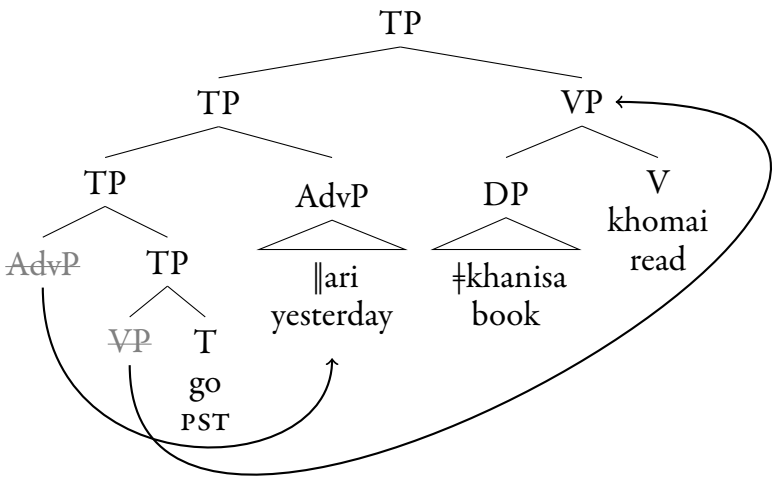


It's not clear what could motivate such a movement, nor why it should only be obligatory with particular TAPs. This problem compounds when we consider that TAP particles can precede temporal adverbs, which are commonly assumed to be adjoined to TP:

(19) Dandagob ge go ǁari ǂkhanisa khomai.
D. DECL PST yesterday book read
"Dandago read the book yesterday."

Consider the movements that would be necessary to produce this word order: First, the temporal adverb would need to raise to a right-adjoined position (20-a); then VP would raise to a higher right-adjoined position (20-b). These movements would need to happen in this precise order, otherwise the ungrammatical (21) would result. Given the stipulative nature of this analysis, it seems worth dismissing the raising analysis.

(20)



(21) *Dandagob ge go ǂkhanisa khomai ǁari.
D. DECL PST book read yesterday

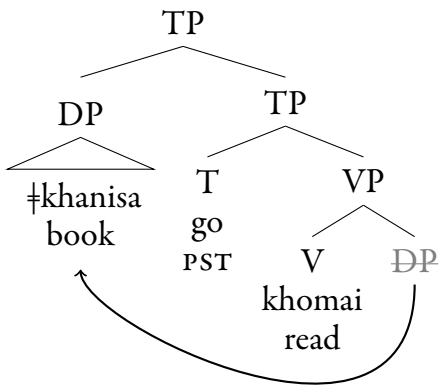
4.2.3 Third possibility: Fronting

Washburn (2001) argues that Khoekhoe clauses are underlyingly head-initial. Under this analysis, preverbal tense particles are in their base position; instead, it is everything else in the VP that has moved. That is, to derive the word order in (22), the object DP *ǂkhanisa* 'book' is forced to evacuate the VP and move to a specifier of TP, as shown in (23).[4]

(22) Dandagob ge ǂkhanisa$_i$ go khomai t_i.
D. DECL book PST read
"Dandago read the book."

[4]Washburn proposes that only TP assigns case in Khoekhoe and that it can only do so to items in its specifiers; however, he assumes that it can assign case to multiple specifiers. On this analysis, VP-internal material is forced to move to Spec,TP to receive case. There are some difficulties with this analysis; for one, it isn't clear why that VP-internal adverbs would be forced to move for case. Second, it isn't clear that the *-a* marker that he takes to be case in fact represents anything of the sort. See Kusmer & Devlin (2018) for a more thorough summary and analysis of the distribution of the *-a* marker.

(23)



Other than the position of preverbal tense markers, Washburn's only evidence for this analysis is the fact that weak object pronouns appear postverbally. He argues that these are the only objects allowed to retain their base position.

(24) Taras ge ǂkhanisa ǁari go mā -te.
woman DECL book yesterday PST give -1S.OBJ
"The woman gave the book to me yesterday."

Washburn's analysis does not account for the postverbal TAPs. If T^0 is underlyingly head-initial, why should some TAPs follow the verb? We might propose that head-movement raises V to T in these cases, but recall that Washburn's primary evidence that VP was head-initial came from the position of light object pronouns. When there is a postverbal TAP, these object pronouns precede it:

(25) Taras ge ǂkhanisa ǁari mā -te tama.
woman DECL book yesterday give -1S.OBJ NEG.NF
"The woman didn't give me the book yesterday."

If the light object pronoun is in fact a DP in its base position, then (25) cannot be (only) V-to-T movement. If instead the light object pronoun is simply an agreement clitic on the verb (perhaps on v^0), then we lose our motivation for having VP (and TP) be head-initial in the first place.

4.2.4 VP coordination

The nail in the coffin for a syntactic analysis of preverbal TAP particles comes from VP coordination. When two VPs are coordinated under a single T^0, the postverbal tense markers obligatorily occur clause-finally (26). By contrast, the preverbal TAPs may freely occur in either conjunct (27):

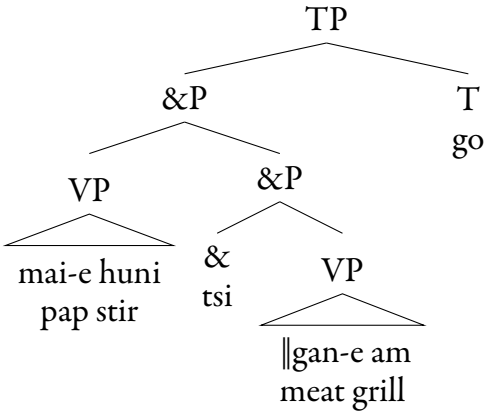
(26) Aob ge mai-e huni tsi ǁgan-e am tama.
man DECL pap stir and meat grill tama
"The man didn't stir the pap and grill the meat."

(27) a. Aob ge mai-e huni tsi ǁgan-e go am.
man DECL pap stir and meat PST grill

b. Aob ge mai-e go huni tsi ǁgan-e am.
man DECL pap PST stir and meat grill
"The man stirred the pap and grilled the meat."

The two sentences in (27) show no difference in meaning; the TAP evidently scopes over both verbs. Syntactically, then, (27) should have a structure like (28):

(28)



None of the syntactic analyses considered above will plausibly allow us to derive the correct word orders from the tree in (28). In all cases, the relevant syntactic movement would involve extracting part of the coordinate structure (or lowering into the coordinate structure), in violation of the Coordinate Structure Constraint Ross (1967). Insofar as this constraint is believed to be universal, we should disprefer any possible syntactic analysis of preverbal TAPparticles in Khoekhoe.[5]

4.2.5 Summary

In this section, I've demonstrated that the placement of preverbal TAP particles in Khoekhoe meets the first criterion for identifying prosodic displacement: All syntactic movements that could account for this word order have little to no independent motivation in the language and would need to violate the Coordinate Structure Constraint.

4.3 Criterion 2: Semantic inactivity

The final criterion for prosodic displacement is that the displacement is semantically vacuous. In at least most cases, this is trivially true in the case of Khoekhoe preverbal TAP particles: these particles can appear before any XP in the middlefield with no change

[5] In fact, the evidence for the Coordinate Structure Constraint in Khoekhoe is complex and mixed: The language broadly allows extraction from the first conjunct, as discussed in Kusmer (2018). However, it universally disallows extraction from the second conjunct. Deriving the correct word order for the VP coordination case via syntactic movement would certainly involve extracting from the second conjunct, and therefore we are justified in excluding this analysis based on the Coordinate Structure Constraint.

in meaning, as shown in (29). In elicitation, speakers uniformly commented that these sentences were identical in meaning and usage, and in fact frequently had trouble distinguishing them from one another even when primed to look for differences in word order.

(29) a. Ti |hōs ge go -ro ǂkhani-e ǁkhawa xoa.
my friend DECL PST IMP book again write

b. Ti |hōs ge ǂkhani-e go -ro ǁkhawa xoa.
my friend DECL book PST IMP again write

c. Ti |hōs ge ǂkhani-e ǁkhawa go -ro xoa.
my friend DECL book again PST IMP write
"My friend was writing a book again."

4.4 Criteria 3 & 4: Morphosyntactic heterogeneity, prosodic homogeneity

The last two criteria discussed in chapter 3 for identifying prosodic displacement are syntactic heterogeneity and prosodic homogeneity: The candidate displacement structures should show some regularity in prosodic form and a lack of any such regularity in morphosyntactic features. Khoekhoe TAP particles show prosodic uniformity in two ways. First, as initially noted by Hahn (2013), whether a given TAP particle will appear in pre- or post-verbal position is determined only by its prosodic weight, not by any morphosyntactic features. Second, the position of the TAP particle predicts whether the verb will undergo sandhi or not, in ways not easily explained by reference to syntactic structure.

4.4.1 Particle weight

As noted above, Khoekhoe TAP particles come in two flavors, either pre- or post-verbal. Hahn (2013) was the first to notice that the only predictor of which class a given particle will fall into is its prosodic weight: Particles with at least two moras appear postverbally, while particles with exactly one mora appear preverbally. A complete list of Khoekhoe TAP particles is presented in (30) & (31).

(30) Preverbal TAP particles

	IPA	Gloss
a	[ra]	present stative
ra / ta	[ra] / [ta]	imperfect
ge	[ke]	remote past
go	[ko]	recent past
ni	[ni]	future
ta	[ta]	negative non-finite
ga	[ka]	irrealis[6]
Compound particles:		
gere	[kɛrɛ]	remote past imperfect
goro	[kɔrɔ]	recent past imperfect
nira	[nɪra]	future imperfect
gara	[kara]	irrealis imperfect

(31) Postverbal TAP particles

	IPA	Gloss
tama	[tama]	non-future negative
tide	[tɪte]	future negative
i	[iː]	non-present stative
hâ	[hãː]	perfect

Note that there is no good morphosyntactic predictor of which class a given particle will take. Negative markers appear both preverbally (*ta* 'negative infinitive') and postverbally (*tama* 'negative non-future'). Stative aspect markers (which are arguably copular) appear both preverbally (*a* 'stative present') and postverbally (*i* 'stative non-present'). Aspect markers include both preverbal *ra* 'imperfect' and postverbal *hâ* 'perfect'.

By contrast, though, prosodic weight is a perfect predictor of which class a particle will take. The only bimoraic particles which appear preverbally are the compound particles, transparently composed of two monomoraic particles. All monomorphemic but bimoraic particles appear postverbally.

Recall from example (12) that preverbal particles do appear postverbally under certain conditions, while postverbal ones never appear preverbally. These facts, taken together with the observation that the rest of the language is overwhelmingly head-final, support an analysis in which T^0 (and Asp^0 and Pol^0) is in some sense head-final, but under certain prosodic conditions undergoes displacement into pre-verbal position.

4.4.2 Tone sandhi

Lexical tone is contrastive in Khoekhoe (Brugman, 2009). All lexical items are associated with one of six tonal classes[7]; each tonal class is, in turn, associated with a particular tonal

[7] Functional vocabulary also has contrastive tone, but it works differently from the tone on lexical vocabulary; see Brugman (2009) for details.

melody made up of a sequence of at most two levels out of the four contrastive tone levels. The word will be produced with this melody, called the 'citation melody', in isolation or in certain prosodically strong positions (defined in more detail below).

(32) Citation melodies: Brugman (2009)

Melody	Description	Example	Gloss
SL	Superlow	[!ǹàas]	'servant'
L	Low	[‖àas]	'tie'
SL-L	Low-rising	[!nǹàs]	'story'
H	High	[ǂáas]	'plain'
SH	Superhigh	[!nā́as]	'tortoise'
H-SH	High-rising	[ǂā́ás]	'spittle'

As noted, the citation meldoy only surfaces in certain prosodic contexts. There is a proces of tonal sandhi that applies in most contexts. Sandhi is an opaque tonal substitution process mapping each of the six citation melodies onto another, apparently arbitrary melody; it can broadly be characterized as a weakening process in the sense that it reduces the number of cross-linguistically marked tonal melodies: The inventory of sandhi melodies is overall lower than the inventory of citation melodies and contains no rising countours (which are generally marked cross-linguistically, see e.g. Yip (2002)). The six citation melodies and their sandhi counterparts are given in table (33).

(33) Sandhi forms:

Citation	Sandhi
SL-L	SL-L
SL H	L-SL
L H-SH	L
SH	H

I have said that the citation melodies appear in prosodically strong positions, while sandhi applies everywhere else. It's time to make that more precise. Within the nominal domain, the generalization is clear: The leftmost item in a DP (or PP) receives citation form, while all other items undergo sandhi.

(34) Sandhi in DPs (citation forms in red): Brugman (2009)

a. súúku
pots

b. |ápā̋ sùùku
red pots

c. !nánī̋ |àpa sùùku
six red pots

d. ǁnáã !nàni ǀàpa sùùku
those six red pots

The association between the left edge of phrases and citation melody is preserved when the verb is moved to the left perhiphery (and thus winds up at the left edge of the clause): In this context, the verb takes citation melody regardless of what occurs later in the clause.

(35) Khőmai go =b ge Dandagoba ǂkhanisa.
read PST =3MS DECL D. book
"Dandago read the book."

4.4.3 Sandhi on verbs

The situation becomes more complex when we consider *in situ* verbs, however. Previous works on verbal sandhi give contradictory generalizations: Haacke (1999) states that verbs take sandhi whenever there is a preverbal TAP; Brugman (2009) states that all root-clause verbs undergo sandhi, while all embedded verbs retain their citation form. In order to resolve this conflict, I conducted a prosodic production experiment. The final generalization resulting from this experiment is as follows: Root clause verbs undergo sandhi whenever they are preceded by a TAP; embedded clause verbs do not undergo sandhi except in quotative clauses, where they behave like root verbs.

Methodology

Four native speakers of Khoekhoegowab were asked to read sentences from a set of slides. There were 54 items total, of which 14 were fillers and the remaining 40 were divided into 6 classes testing different sentence structures. Within each class, half the items had a preverbal TAP and half the items had a postverbal one. All of the test items used the verbs listed in (36); these verbs were selected to be entirely sonorant (to aid in F0 tracking) and to have either High or High-Rising citation melodies, which are the two melodies showing the most dramatic register shift under sandhi.

(36) List of verbs used:

Verb	Gloss	Citation	Sandhi
oa	'return'		
ā	'cry'	H	L-SL
om	'build'		
mû	'see'		
huni	'stir'	H-SH	L
am	'grill'		

Each speaker saw all 54 sentences twice. The sentences were all recorded on a Zoom H5 recorder using a Shure SN10A-CN head-mounted microphone. After recording, individual items were segmented and then force-aligned using the Montreal Forced Aligner (McAuliffe et al., 2017), which was trained on a dataset including roughly 4.5 hours of

transcribed Khoekhoe speech from 8 speakers. After alignment, the verb was identified and hand-transcribed.

In order to confirm the accuracy of the transcriptions, quantitative analysis was performed. Parselmouth (Jadoul et al., 2018) was used to extract the mean F0 of the verb. F0 range varies widely from between speakers and utterances; in order to normalize verb F0, I used a moving window of one-half the utterance before the verb; the mean of this window was extracted, and then the normalized value was taken by dividing the mean F0 of the verb by the mean F0 of the window preceding it. In R (Team, 2013), I constructed a logistic model to predict the transcription, with a fixed effect of normalized verb mean F0 and random effects of speaker and lexical item. This model was trained on 25% of the data; when tested on the remaining data it achieved 91% accuracy.

Results

Let's first consider those test items with root clauses. There were 12 in total — 8 declaratives and 4 questions, giving a total of 80 tokens. The results are tabulated in (37); only one verb with no preverbal TAP was observed to undergo sandhi, and no verbs with preverbal TAPs failed to undergo sandhi.

(37) Verb sandhi in root clauses:

	Citation	Sandhi
Preverbal	0	40
Postverbal	38	2

Next we'll turn to embedded clauses. There were three classes of embeddings tested: Nominalized complement clauses (38-a); quotative complement clauses (38-b); and clefted relative clauses (38-c).[8] Note that the quotative clauses are unique in having the declarative clause type marker *ge* present; both of the other clause types have no (overt) marker. The results are tabulated in (38).

[8] The analysis of these OVS clauses as involving relative clauses hinges on three facts. First, the subject obligatorily undergoes sandhi in this context, implying that the preceding material is contained within the DP, making the noun not leftmost within its own phrase. Second, OVS word order is ungrammatical when the subject is immediately preceded by a determiner.

(i) [ǁAri tsaurase go -ro !gû] (*ne) kai xammi ge.
yesterday softly PST IMP walk (*this) big lion DECL
"Yesterday the big lion was walking softly."

This is unexpected if OVS is derived by e.g. TP fronting, but expected if the OV is a subject relative clause within the DP. Finally, this word order has a unique pragmatic meaning: it is used to convey that the subject is new information while the rest of the clause is given, parallel to cleft structures in other languages.

(38) Verb sandhi in embedded clauses:

		Citation	Sandhi
Nominalized	**Preverbal**	16	0
	Postverbal	16	0
Quotative	**Preverbal**	8	0
	Postverbal	3	5
Relative	**Preverbal**	16	0
	Postverbal	16	0

In the nominalized and relative clause conditions, verbs failed to undergo sandhi regardless of the position of the TAP. In the quotative condition, the results are more complicated. All the verbs preceded by a TAP underwent sandhi, exactly as would be the case in a root clause. Three of the tokens where the verb was not preceded by a TAP retained their citation melody, as expected in root clauses; the remainder were transcribed as undergoing sandhi. There are at least two possible explanations for this. The first is that verbs in quotative clauses really do unexpectedly undergo sandhi with some variability. Perhaps a simpler explanation is simply that these cases were mistranscribed: These sentences are quite long, and the effect of downdrift means that the total F0 range used by the speaker that late in the sentence is low, making the difference between the citation and sandhi forms ambiguous. I'll assume going forward that these transcriptions are in error due to this ambiguity.

The final condition tested in this study involves VP coordination. Recall from the discussion of syntactic implausibility that in VP coordination, a preverbal tense-marker can appear in either conjunct. This gives us the means to disambiguate two possible hypotheses about where verbs undergo sandhi. The first hypothesis is the one given by Haacke (1999): Verbs undergo sandhi whenever there is a preverbal TAP— that is, sandhi on the verb depends on the morphosyntactic features of the TAP. The alternative hypothesis is that verbs undergo sandhi whenever they are *preceded* by a TAP — that is, sandhi on the verb depends on the linear position of the TAP. In order to resolve this, items in the coordination condition came in three varieties: They could have a postverbal TAP (39-a); a preverbal one preceding only the second verb (39-b); or a preverbal one preceding both verbs (39-c). The results are tabulated in (40).

(39) a. Aob ge mai-e huni tsi ǁgan-e am tama.
man DECL pap stir and meat grill NEG.NF
"The man didn't sir the pap and grill the meat."

b. Aob ge mai-e huni tsi ǁgan-e go am.
man DECL pap stir and meat PST grill.

c. Aob ge mai-e go huni tsi ǁgan-e am.
man DECL pap PST stir and meat grill
"The man stirrred the pap and grilled the meat."

(40) Sandhi in VP coordination:

	Citation	Sandhi
Preceded by TAP:	2	30
Not preceded by TAP:	40	0

The results support the linear position hypothesis: root clause verbs undergo sandhi when they are preceded by a TAP; they fail to undergo sandhi if they are followed by a TAP, even if that TAP is in fact from the preverbal class.

4.4.4 Summary: Prosodic uniformity

In this section, I've shown that sentences containing preverbal TAP particles show prosodic uniformity (and morphosyntactic non-uniformity) on two different levels. First, while both pre- and postverbal TAPs are morphosyntactically heterogeneous, the preverbal ones are uniformly prosodically light and the postverbal ones are uniformly prosodically heavy. Second, sandhi on verbs in root clauses is entirely predicted by the linear position of the TAP particle, not by its morphosyntactic identity. Thus, on both counts, Khoekhoe preverbal tense markers meet the second and third criteria for identifying prosodic displacement.

4.5 Conclusions

The placement of Khoekhoe TAP particles has all the hallmarks of prosodic displacement. The particles fall into two distributional classes based on whether they precede or follow the verb. The preverbal particles can appear in a range of positions that would be implausible landing sites for syntactic movement. Furthermore, both classes of particle are morphosyntactically heterogeneous. By contrast, both classes of particle are prosodically uniform: preverbal particles are at most one mora, while postverbal particles are at least two, the minimum number of moras the language requires of a prosodic word. Additionally, root clauses in which a TAP particle have displaced past a verb uniformly show sandhi on the verb. Finally, this displacement has no discernible semantic or pragmatic effect. All together, the preverbal position of some Khoekhoe TAP particles seems to be derived by prosodic displacement. In the next chapter, we'll briefly leave Khoekhoe behind in order to consider how prosodic displacement, and linearization more generally, might be accounted for; in chapter 6 we'll return to Khoekhoe to see how to derive the linear order and prosodic structure of preverbal TAPs.

Chapter 5

Optimal Linearization

In chapter 2, I argued that extant models of linearization don't provide good explanations for typological effects. In chapters 3 & 4, I showed that whatever linearization model we choose to adopt must be capable of accounting for PF displacement. In this chapter, I will propose a model of linearization which begins to provide some explanation for these two problems. I start from the perspective that linearization is a PF phenomenon (Kayne, 1994a; Chomsky, 1995) and should be modelled the same way we model other phonological processes, namely with violable constraints. This allows us to modeel PF displacement by having constraints on linearization come into competition with prosodic markedness constraints. In contrast to the violable linearization models mentioned in chapter 2, however, I propose that the mapping from syntactic structures to linear strings occurs fully post-syntactically: Rather than proposing a single "word order faithfulness" constraint penalizing deviance from a pre-specified order, I propose a family of constraints which enforce certain relationships between syntactic structure and word order, working together to derive the correct output. Modelling linearization in this way has the benefit of making clear, well-defined typological predictions, in the form of factorial typology: Different rankings of constraints should predict all and only the classes of word order actually observed.

I will call this general approach Optimal Linearization, and will demonstrate that, given the right constraint set, we can predict the typological gap described as the FOFC while still offering a coherent explanation for why specifiers are always left.[1] My proposed constraint set models word order typology as arising from the competition of two core constraints: One, HeadFinality, encodes a general preference for heads and their non-maximal projections to follow their sisters. The other, Antisymmetry, encodes a competing preference for syntactic objects higher in the tree to be linearized earlier in the string; it closely mimics the effect of the familiar LCA (Kayne, 1994a). These are both violable constraints; in some cases satisfaction of one constraint will entail violation of the other. Competition of these two constraints will derive the two harmonic word orders (head-initial and -final). Within this framework, the leftward position of specifiers occurs not because the specifier c-commands the head, but rather because the terminals within the specifier *fail* to c-command it; specifiers are therefore placed on the left as

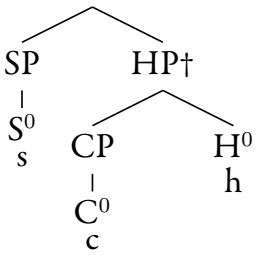
[1] In particular, I aim to capture the ordering of specifiers and complements; I will not take up the positioning of adjuncts here. See section 5.4 for thoughts on how this system might be extended to address the ordering of adjuncts.

the grammar tries to achieve the "most head-final" ordering possible with heads still preceding their complement. Finally, a third constraint HeadFinality-α is identical to HeadFinality except that it considers only the order of those heads dominated by some node α. The addition of this constraint allows us to derive exactly those disharmonic orders compatible with the FOFC. In chapters 6 & 7 I'll show that these same constraints allow us to account for PF displacement phenomena, and in fact fair better than the previous violable-linearization models.

5.1 Harmonic word orders

I'll introduce Optimal Linearization by illustrating how it models a subset of the complete typology. In particular, I will start by considering only the "harmonic" word orders — those word orders that are consistently head-initial or head-final in all phrases. Intuitively, we want the Optimal Linearization procedure to take a syntactic structure like (1-a) and produce one of the two orders in (1-b) (and no others). (The nodes have been named corresponding to their structural position — so the specifier is SP, the head is HP, and the complement is CP.)

(1) a.



b. **Head-Initial:** shc
c. **Head-Final:** sch

In the Headedness Parameter model, these two orders are controlled by a single parameter. In a violable-constraint framework, it's more natural to have them controlled by two constraints: When one constraint (call it HeadFinality) is dominant, the output will be the head-final order *sch*; when the other constraint (call it Antisymmetry) is dominant, the output will be the head-initial order *shc*. Further, we want this to extend to all phrases — that is, if there is more material in SP or CP, we want those phrases to be linearized the same way as HP. The goal of this section will be to define the constraints HeadFinality and Antisymmetry to achieve exactly this result.

Before getting to the constraints themselves, however, I first need to introduce the rest of the Optimal Linearization model.

5.1.1 Some housekeeping

Before getting into the constraints themselves, it's worth taking a second to formalize what exactly the complete model looks like.[2] The general architecture of OT involves two core components: Gen takes an *input* and generates from it a number of *candidates*

[2] While I endeavor to introduce the formal mechanisms of OT in this text, readers unfamiliar with the framework are referred to McCarthy 2002 for a more complete introduction.

(i.e. potential outputs); Eval takes the input and candidate set and, using a set of ranked violable constraints, selects a winner, which is the output of the model overall. Any given language is taken to have a fixed ranking of constraints. Taken together, Gen, Eval, and the ranked constraints are a function from the possible inputs in the language to the possible outputs.

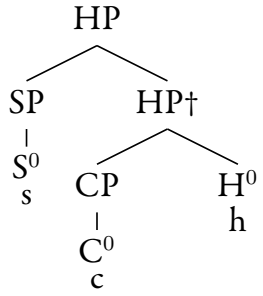
In Optimal Linearization, the input to Gen is the output of the narrow syntax, i.e. a phrase marker produced by some particular theory of syntax. While Optimal Linearization is compatible with a variety of syntactic theories, I will use structures compatible with Merge-based derivations and the Minimalist Program generally (Chomsky, 1995). I will assume that the candidates created by Gen are strings composed of whatever phonologically-contentful Vocabulary Items are produced by the Spell-Out of the set of syntactic terminals in the input. I'll refer to these vocabulary items generically as "words". The set of candidates produced by Gen will be the full set of possible orders of words, so if there are n syntactic terminals mapped to phonologically-contentful words, there are $n! = n(n-1)(n-2)...$ candidates from which a single unique winner will be selected. Phonologically null syntactic terminals remain part of the input to the linearization component, but are never present in any of the candidates.

As a matter of notational convention, I will use capital letters to denote syntactic terminals (A, B) and lower case letters to refer to the words corresponding to them (a, b). In addition, I will reserve the letters {X, Y, Z} for variables ranging over syntactic labels; letters from the beginning of the alphabet denote specific syntactic objects. The symbol $<$ denotes string precedence, so $x < y$ means some word x precedes some word y. As a last notational convention, I will draw all syntactic trees in a head-final fashion; remember, however, that syntactic trees have no order!

5.1.2 HeadFinality

Having dispensed with the preliminaries, let's now turn to the derivation of head-final orders. This will be accomplished by a constraint HeadFinality which, given the input (2-a), prefers the order in (2-b) to all other possible orders (2-c).[3]

(2)



a. → sch
b. *shc, *csh, *chs, ...

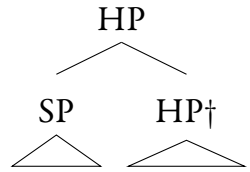
Let's think about what properties the winning order sch has that the other possible orders don't. First, it orders the specifier s before everything that isn't the specifier; any order that doesn't have s initial will be dispreferred. Put another way, the correct output

[3] Optimal Linearization requires that we be able to distinguish phrasal nodes from non phrasal nodes, but "bar-levels" have no special status in this theory; as such, I've labelled all non-terminal nodes as "XP", here and in all other trees. However, for expositional reasons it will be convenient to have unique labels for each node; accordingly, I've marked the phrasal, non-maximal nodes with †.

has HP† following its sister. Second, the correct output orders the complement c before the head h; any order that has $h < c$ will be dispreferred. Put another way, H^0 follows its sister.

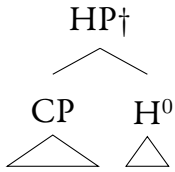
By visualizing each branching node separately, it can be seen that these two ordering conditions share a structural description. One ordering relation relates the daughters of HP to each other; the other relates the daughters of HP† to each other. In each case, the daughter that shares a label with the node in question (HP† for HP; H for HP†) is set to follow the daughter that doesn't (SP for HP; CP for HP†).

(3) a.



→ s < {c, h}

b.



→ c < h

It's going to be useful to have a pair of terms that distinguish these two structural relations. I'm going to call the daughter that shares a label with its parent the **descendant** or **endogenous daughter**; the one that doesn't share a label with it I'll call the **in-law** or **exogenous daughter**. Put another way, when two nodes undergo Merge, the one which projects becomes the descendant and the one that doesn't becomes the in-law. Specifiers and complements will always be in-laws of the nodes immediately dominating them; heads and their non-maximal projections will always be descendants.

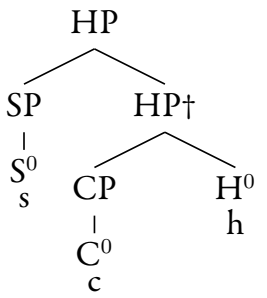
Intuitively, then, HeadFinality is a constraint that prefers orders in which, for every branching node, the material dominated by its in-law precedes all material dominated by its descendant. Optimality Theory constraints are generally stated in terms of the output configurations they disprefer, i.e. the configurations which incur violations of the constraint. Putting HeadFinality into that form:

(4) HeadFinality : Assign one violation for each branching node XP dominating a pair of terminal nodes X^0 & Y^0 such that:
   a. Y^0 is dominated by the in-law of XP;
   b. X^0 is not dominated by the in-law of XP; and[4]
   c. $x < y$.

I'll illustrate the action of this constraint in an OT tableau. The candidate orders are listed in the leftmost column; the next column lists which branching nodes incur violations of HeadFinality. In this input, there are only two branching nodes and so the constraint scores a maximum of two violations. The arrow indicates the winning candidate sch, the only candidate which scores no violations.

[4] If X^0 is dominated by XP but not dominated by the in-law of XP, then it is by definition dominated by the descendant of XP. Once we turn to linearizing movement structures in section 5.2, we will encounter cases in which a particular head is dominated by both the in-law and the descendent of XP; defining the constraint as shown here will prevent it from giving contradictory orders in these cases.

(5) a.



b.

	(a)	HeadFinality
	shc	*H†
→	sch	
	csh	*HP
	chs	*HP
	hcs	*HP *HP†
	hsc	*HP *HP†

While this is a simple example, it serves to illustrate the action of HeadFinality generally. The constraint will linearize any XP in the same fashion as HP in this example — with everything contained in the specifier foremost, and X^0 final.

5.1.3 Antisymmetry

The constraint HeadFinality suffices for deriving harmonically head-final word orders. In order to derive the head-initial orders we need a constraint that opposes HeadFinality. That is, we want some constraint Antisymmetry such that the same tree in (5) is mapped to the order shc whenever Antisymmetry$\gg$ HeadFinality. It may at first seem tempting to make Antisymmetry the inverse of HeadFinality — that is, have it require the descendant to proceed the in-law. However, this won't work, as head-initial orders and head-final ones are not symmetric: In both orders, the specifier must precede everything that follows it. We need to look for something else that will create head-initial orders than just the reverse of HeadFinality.

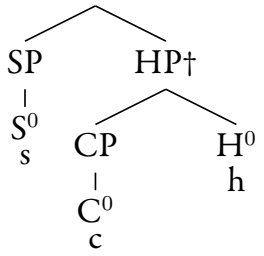
I propose that we take a (metaphorical) page from Kayne's (literal) book (1994) and make Antisymmetry a constraint that enforces correspondence between asymmetric c-command and precedence. Unlike Kayne, however, I will only consider relationships between terminal nodes. This frees us from making stipulations about segments & categories, and will also have some other benefits that I will make clear momentarily. Intuitively, then, the constraint that we're looking for is one that penalizes words that occur in the opposite order as the asymmetric c-command relation between their terminals. More formally:

(6) Antisymmetry: Assign one violation for each pair of terminal nodes X^0 & Y^0, where:
   a. X^0 asymmetrically c-commands Y^0; and
   b. $y < x$.

This constraint ranges over pairs of nodes that stand in an asymmetric c-command relation. In the basic spec-head-comp structure we've been investigating so far, there is only one such pair: The head H^0 asymmetrically c-commands everything in CP (namely C^0). As such, Antisymmetry will score a maximum of one violation whenever $c < h$. However, Antisymmetry will not order the specifier S^0 with respect to either of the other heads — while the phrase SP asymmetrically c-commands both h and c, S^0 itself does not. How, then, will the system order the specifier? Conveniently, we already have a constraint which accomplishes this: HeadFinality requires that HP be linearized such that everything in SP precedes everything in HP†. In a violable constraint system

like OT, low-ranked constraints remain active even when dominated by a higher ranked constraint; even when ANTISYMMETRY≫ HEADFINALITY, then, HEADFINALITY is still active and can enforce the leftward position of the specifier. I've presented this in tableau form below. ANTISYMMETRY eliminates the three candidates in which $c < h$; of the three that remain, only one fails to incur a violation of HEADFINALITY for HP, namely the one that orders the specifier on the left.

(7) a.



b.

	(a)	ANTISYMMETRY	HEADFINALITY
→	shc		*HP†
	sch	*$h < c$	
	csh	*$h < c$	*HP
	chs	*$h < c$	*HP
	hcs		*HP *HP†
	hsc		*HP *HP†

This is a case of "the emergence of the unmarked" (McCarthy & Prince, 1994b): The lower-ranked constraint acts to select the winner exactly when the higher-ranked one fails to choose. In this case, the higher-ranked ANTISYMMETRY doesn't select between the different placements of the specifier s within the string — it only requires that the head precede its complement. The fact that the specifier is on the left in the winning candidate is a reflection of the system choosing the "most head-final" order among those compatible with the order $h < c$. Optimal Linearization thus gives us new insight into a previously-mysterious fact about word order typology, namely that specifiers are always left-most even in otherwise "head-initial" languages. Put another way, it has always been somewhat problematic that so-called head-initial languages are never fully head-initial, but rather always require specifiers to precede the head. Optimal Linearization lets us understand this fact as a preference for head-finality emerging even in otherwise head-initial languages.

So far we've considered only a single, abstract tree where the specifier and the complement contain only a single word. Hopefully it is clear that adding more words to either specifier or complement will behave in the expected way: HEADFINALITY will provide pressure to linearize all the specifier material before head & complement and also all the complement material before the head; ANTISYMMETRY, likewise, will provide pressure to linearize the head before all the complement material — the head, after all, does asymmetrically c-command all of its complement. The same general pattern of linearization will be replicated within each XP, just as we'd expect. There is one class of syntactic structure not yet accounted for, however, namely structures involving movement. This is what I'll turn to in the next section.

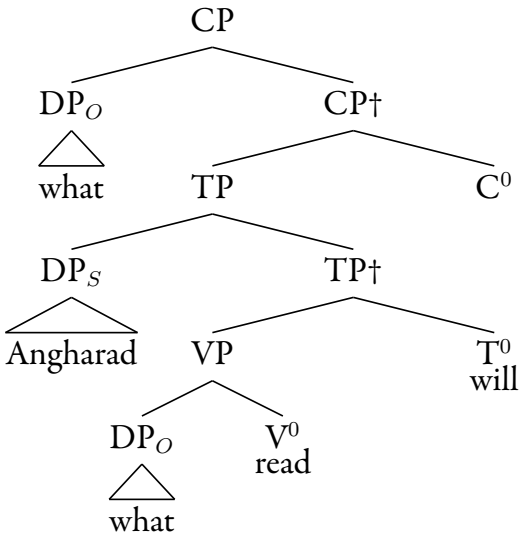
5.2 Linearizing movement

One of the goals of any linearization algorithm must be to explain why moved items appear in the location that they do (and only that one). That is: Once an XP has moved, what prevents it from being linearized according to its base position? And what prevents

it from being spelled out twice, once according to each position? In most traditional theories of linearization there is an operation of "copy-deletion" which applies before linearization and transforms the tree at PF such that moved items are only in one position. However, Johnson (2016) outlines some possible undesirable consequences of introducing this extra transformation between the syntax and the linearization. Instead, I propose to keep to the original intuition that it is linearization itself that forces moved items to be spelled out in a particular location. The input to Optimal Linearization, then, will still have moved items in all of their positions. I will assume for the moment that GEN only creates candidates that have exactly one word for each (phonologically-contentful) syntactic terminal, even if that terminal has multiple copies — that is, GEN can't distinguish individual copies of a moved item. This prevents moved items from being linearized in multiple positions (a.k.a. multiple spellout). This may or may not be a desirable assumption, as multiple spellout of movement chains has been proposed as an analysis of resumption (Sichel, 2014, e.g.) and verb-doubling predicate clefts (Koopman, 1984; Kandybowicz, 2006; Cable, 2004, e.g.). If we want to capture these phenomena using multiple spellout, we would need to relax this restriction on GEN but then add additional constraints to enforce single spellout in all but the relevant contexts. Such a project is beyond the scope of this paper, so for the moment I'll use the constrained version of GEN.

With that in mind, let's consider what we want the Optimal Linearization constraints to do in the case of movement structures. I'll use English *wh*-movement as an illustrative example; (8) presents a simplified structure for an object *wh*-question.[5]

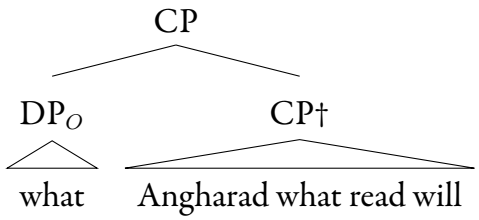
(8)



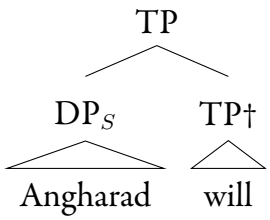
Let's first consider how we want HEADFINALITY to treat the moved item. Recall that HEADFINALITY scores violations based on branching nodes. There are 5 branching nodes in (8), but one of them (CP†) has a branch with no phonologically-contentful words (C^0) and so will never score a violation. The remaining 4 branching nodes are as follows:

[5] More specifically, this is an embedded question. For expository reasons, I've omitted subject-auxiliary inversion and have shown the subject being base-generated in spec,TP.

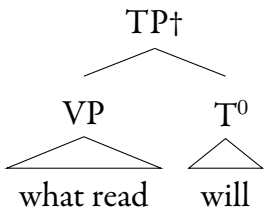
(9) a.



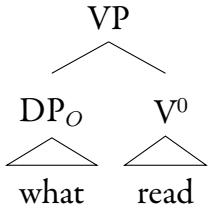
b.



c.



d.



At once we can see that there's a problem. HeadFinality will score a violation for any branching node for which material in its descendant precedes material in its in-law. (9-a) shows that the constraint will score a violation for CP if $Angharad$ (which is in the descendant CP†) precedes $what$ (which is in the in-law DP$_O$). (9-c), however, shows that the constraint will score a violation for TP whenever $what$ (which is in the descendant TP†) precedes $Angharad$ (which is in the in-law DP$_S$). This produces produces a contradictory ordering for this tree.

Of course, the problem is that the constraint as defined can't distinguish between the 'high' and 'low' positions of the moved item. We want $what$ to be linearized according to its higher position[6] , namely spec,CP. In other words, we want the constraint HeadFinality to consider DP$_O$ only when it is evaluating the node CP; the contents of DP$_O$ should not be relevant for the linearization of any lower branching node. In order to accomplish this, I will borrow from Abels (2003) the idea of *total domination*. Intuitively, some node X dominates a node Y only if it dominates all copies of Y. Formally:

(10) X **totally dominates** Y iff all copies of Y are dominated by a copy of X.

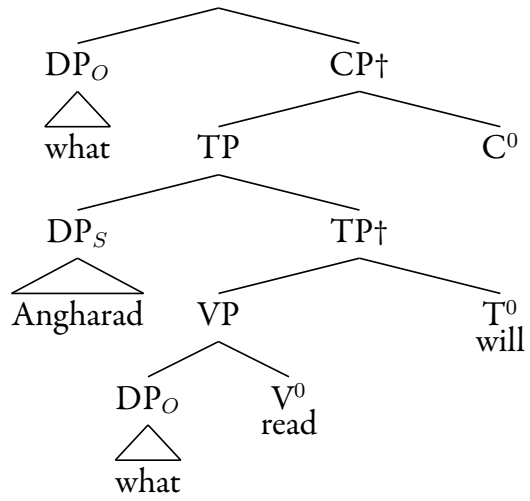
In (8), DP$_O$ is totally dominated by only two items: itself (total domination is reflexive) and CP. All of the other terminal nodes are totally dominated by everything which (non-totally) dominates them — in the absence of movement, domination and total domination are identical. This allows us to revise our definition of HeadFinality to linearize the moved item according to its highest position:

[6]This may not always be true if for instance *wh-in situ* languages covertly raise the *wh* item (Watanabe, 1992; Cole & Hermon, 1998, e.g.) — in covert movement in general it seems that the linearization scheme must pick the lower copy. Fully accounting for these facts is beyond the scope of this paper, but we might propose that for instance there are two versions of each of the Optimal Linearization constraints, one which sees the lower copy and one the higher; the ranking of these versions relative to each other would determine whether movement overt or covert. Further refinement would be needed to ensure that overt and covert movement could coexist in the same language.

(11) HeadFinality (revised): Assign one violation for each branching node XP **totally dominating** a pair of terminal nodes X^0 & Y^0 such that:
   a. Y^0 is dominated by the in-law of XP;
   b. X^0 is not dominated by the in-law of XP; and
   c. $x < y$.

Because CP is the only branching node which totally dominates *what* in (8), the only way for *what* to violate HeadFinality is for it to follow anything contained in CP but not in DP_O, i.e. any word in CP†other than itself. As such, *what* (and in fact all of DP_O, if it were larger) will be linearized leftmost, in accordance with its moved position. I've illustrated this in the tableau in (12); space does not permit me to include all 24 candidate orders, so I've chosen a representative set. The winning candidate is a fully head-final pseudo-English.[7]

(12) a.



b.

	(a)	HeadFinality
→	what Angharad read will	
	Angharad what read will	*CP
	what Angharad will read	*TP†
	Angharad will read what	*CP *TP†

Of course, to achieve the correct head-initial order for English we need to consider Antisymmetry. Here, we face a similar problem: V^0 still asymmetrically c-commands everything (non-reflexively) dominated by DP_O, and so Antisymmetry will exert pressure for $read < what$ as though *wh*-movement had never occurred. Again, what we want is a notion of *total c-command* parallel Abels (2003): V^0 fails to c-command *what* in all of its positions, and therefore won't be ordered before it. Total c-command is easy to formalize:

[7] Here we see the relevance of defining HeadFinality such that the material in the in-law must precede the material 'not in the in-law' (as opposed to 'in the descendent'), as mentioned in fn. 4: *what* is contained in both CP's in-law and descendent. If the constraint were defined in terms of the descendent, it produce the nonsensical ordering of *what* > *what*. The problem gets worse if the moved item has multiple words, for example if DP_O were *which book*: Here the constraint would both require *which* > *book* (since *which* is in the in-law and *book* is in the descendent) and *book* > *which* (since the reverse is also true).

(13) a. X **totally c-commands** Y iff:
(i) X does not dominate Y; and
(ii) everything that totally dominates X also totally dominates Y.
b. X **asymmetrically totally c-commands** Y iff X totally c-commands Y and Y does not totally c-command X.

In (8), V^0 does not totally c-command DP_O: for one, V^0's immediate mother VP does not totally dominate DP_O. In fact, there is nothing that totally c-commands the moved item. All that remains, then, is to update our definition of Antisymmetry to use total c-command:

(14) Antisymmetry(revised): Assign one violation for each pair of terminal nodes X^0 & Y^0, where:
a. X^0 asymmetrically **totally c-commands** Y^0; and
b. $y < x$.

Again, I've illustrated the action of this constraint in a tableau; as before, it fails to order any specifier, but HeadFinality emerges to accomplish that.

(15)

	(a)	Antisymmetry	HeadFinality
	what Angharad read will	*$will < read$	
	Angharad what read will	*$will < read$	*CP
→	what Angharad will read		*TP†
	Angharad will read what		*CP *TP†

With this last modification to the constraints, Optimal Linearization will now linearize all moved phrases according to their highest position.[8]

5.3 Disharmonic word orders

Up to this point, I've restricted my attention to only the two harmonic word orders. There is a third order compatible with the Final-Over-Final Constraint: A head-initial phrase can embed a head-final one (but not the reverse). For example, German embedded clauses have a head-initial complementizer but are otherwise head-final[9] (16-a); for an example lower in the clause, verbal auxiliaries in many of the Mande languages Kastenholz (2003) precede the VP, while the verb itself follows its complement (16-b).

[8] A reviewer asks to what extent the winning candidate is affected by details of the syntactic analysis, in particular by the addition or subtraction of functional material; for example, in (8) I have omitted vP; how would the linearization change if it were included? If the additional material is phonologically contentful, then the resulting candidates will be different and no direct comparison is possible; on the other hand, if the additional material is phonologically null, it will have no affect on the linearization whatsoever: Because only contentful words are present in the output candidates (by assumption), no violations will ever be scored involving a node dominating no contentful material. In essence, linearization operates on a "flattened" structure with null heads (and their immediate projections) are removed; this is reminiscent of the way the Match constraints as defined in Elfner (2012); Bennett et al. (2016) flatten syntactic structure to prosodic structure.

[9] Under the most common analyses of V2, matrix clauses are also an example of a mixed-headed order; I'll stick to embedded clauses here in order to avoid the complexities of head movement.

(16) a. German:

... dass Fritz mich gesehen hat.
that Fritz me seen has

"...that Fritz has seen me."

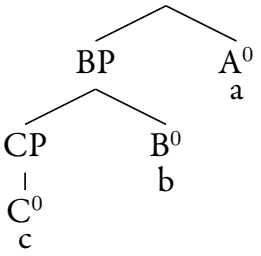
b. Bambara: (Dumestre 1984, via Kastenholz 2003)

fíɲɛ́ yé tásuma jíidi
wind PERFECT fire increase

"The wind made the fire flare up."

Abstractly, the FOFC-compliant disharmonic order follows the schema in (17): AP is linearized in a head-initial fashion, while BP is head-final.

(17) a.



b. **Disharmonic order:** acb

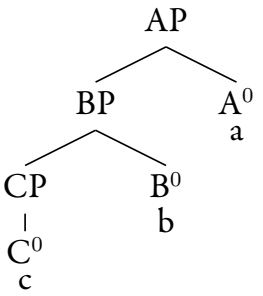
At present, the Optimal Linearization constraint set includes just two constraints, giving a maximum of two rankings / language classes. In order to allow for the disharmonic order, we'll need to add an additional constraint. I propose that this constraint is a relativized version of HeadFinality which only considers those nodes (reflexively) dominated by some node α. For example, in (17), α is BP; the constraint would score a violation for BP (which does reflexively dominate itself) if $b < c$, but would not consider the ordering of a at all. This leaves Antisymmetry free to order AP head-initially.

This constraint captures the core generalization of the FOFC: head-finality "propagates down" the tree such that any node dominated by a head-final node will also be head-final itself. Formally, HeadFinality-α is defined nearly identically to HeadFinality except for a clause specifying its domain of application:

(18) HeadFinality-α : Assign one violation for each branching node XP **dominated by** α and totally dominating a pair of terminal nodes X^0 & Y^0 such that:
   a. Y^0 is dominated by the in-law of XP;
   b. X^0 is not dominated by the in-law of XP; and
   c. $x < y$.

HeadFinality-α and HeadFinality are in a subset ("stringency") relationship: HeadFinality-α will always assign a strict subset of the violations assigned by HeadFinality. In practical terms, this means that whenever they are ranked "together" (i.e. both above or both below Antisymmetry), their effects will be indistinguishable. Only under the ranking HeadFinality-$\alpha \gg$ Antisymmetry $\gg$ HeadFinalitywill they give rise to the disharmonic order. This is illustrated in the tableau in (19):

(19) a.



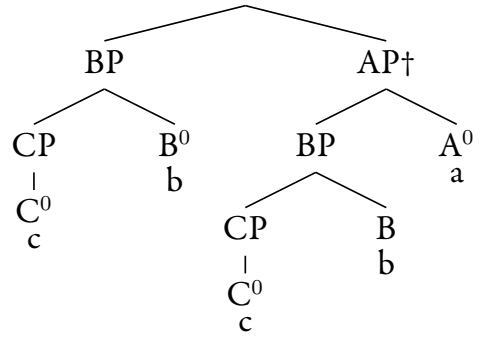
b.

	(a)	HF-BP	ANTISYM	HF
	abc	*BP		*AP *BP
	bac	*BP	$*a < b$	*AP *BP
	bca	*BP	$*a < b, *a < c$	*BP
	cba		$*a < b, *a < c, *b < c$	
	cab		$*a < c, *b < c$	*AP
→	acb		$*b < c$	*AP

Undominated HEADFINALITY-α effectively divides the syntactic structure into two domains: everything below α is linearized purely by HEADFINALITY-α , while everything above it is linearized by the combination of ANTISYMMETRY and HEADFINALITY, just as in the harmonic word order case. It's worth taking a moment to demonstrate that this applies even when movement is involved. There are two relevant cases: Movement of α itself, and movement of some phrase within α to a position outside of it. In both cases, we want the moved item to be head-final within itself, but positioned in a head-initial fashion with respect to the rest of the clause.

The case where α itself moves is illustrated in (20), where BP has moved to the specifier of AP. Both copies of BP (reflexively) dominate themselves, and so both are linearized head-finally; likewise, both copies of CP are dominated by a copy of BP, and so CP would also be linearized head-finally (if there were any other material in it). The only change is that A^0 no longer totally c-commands B^0 and C^0, so ANTISYMMETRY will fail to order it before them; instead, the general constraint HEADFINALITY will emerge to order the specifier on the left.

(20) a.

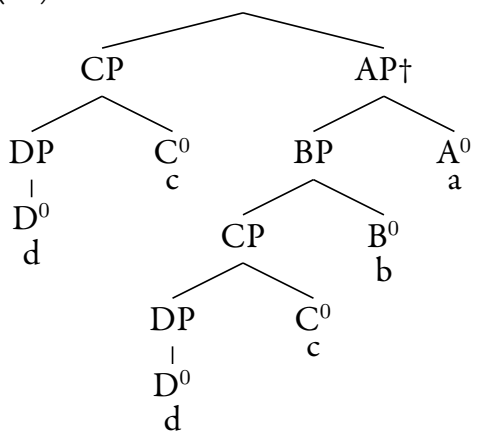


b.

	(a)	HF-BP	ANTISYM	HF
	abc	*BP		*AP *BP
	bac	*BP		*AP *BP
	bca	*BP		*BP
→	cba		$*b < c$	
	cab		$*b < c$	*AP
	acb		$*b < c$	*AP

Movement from within α requires a slightly larger tree to fully see. In (21), $\alpha = \text{BP}$ as before; this time, the complement of BP has moved up to the specifier of AP. Once again, HEADFINALITY-BP applies within CP, which is dominated (though not totally dominated) by BP; only the general HEADFINALITY orders the material in CP with respect to a and b, however, putting the moved item on the left.

(21) a.



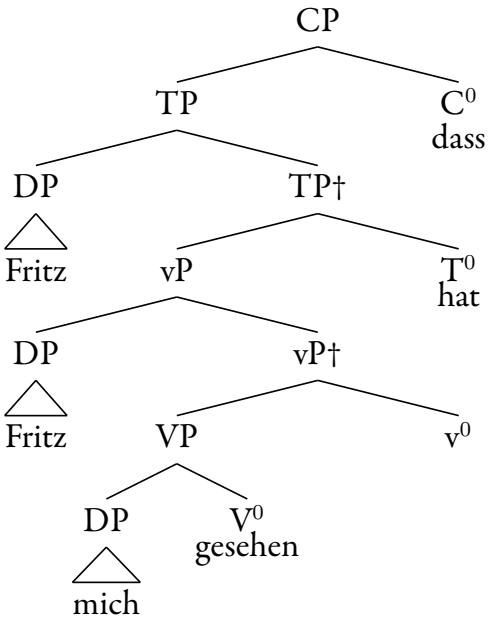
b.

	(a)	HF-αBP	ANTISYM	HF
→	dcab			*AP†
	cdab	*CP		*AP† *CP
	adcb	*CP		*AP *AP†
	abdc			*AP *AP† *BP
	dcba		$*a < b$	

I'll close this section by illustrating how the constraints described here linearize embedded clauses in German. German is a well-known case of a disharmonic word order: Complementizers are on the left, but the rest of the clausal spine is head-final. Thus, the domain of head-finality is TP; that is, HEADFINALITY-TP is undominated. I've given a simplified syntactic structure in (22). [10]

(22) a. ... dass Fritz mich gesehen hat.
that Fritz me seen has
"...that Fritz has seen me."

b.



(23)

	(22-b)	HF-TP	ANTISYM	HF
→	dass Fritz mich gesehen hat		3: *V < O, *Aux < V, *Aux<O	*CP
	Fritz mich gesehen hat dass		7: ... *C<S, *C<O, *C<V, *C<Aux	
	dass Fritz hat gesehen mich	*TP, *VP	0	*TP, *VP, *CP
	dass Fritz hat mich gesehen	*TP	1: *V<O	*TP, *CP

[10] For the purposes of this illustration, I'm ignoring the morphology of the verb itself.

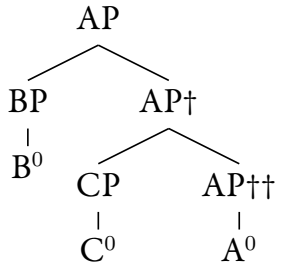
As shown in (23), the constraint HeadFinality-TP eliminates all candidates in which any head below TP is not final within its phrase. Antisymmetry further eliminates those candidates where C^0, the only head not in the domain of head-finality, is not initial. The interaction of these two constraints derives the correct disharmonic word order.

5.4 Conclusion

Optimal Linearization is the proposal that linearization is accomplished at PF by a set of violable constraints which make reference to the syntactic structure I've shown that this model is capable of making detailed predictions about word order typology; I've also shown that it gives new insight into the asymmetric positioning of specifiers, allowing us to understand it as an emergence of an unmarked preference for head-finality.

There is one notable aspect of linearization which has not been taken up here, namely the ordering of adjuncts. The constraints as presently defined will treat adjuncts identically to specifiers. For example, in (24), take CP to be some modifier phrase adjoined to AP. Similar to the specifier case, C^0 neither c-commands nor is c-commanded by any other head in this structure, and so Antisymmetry is silent on its ordering; HeadFinality will prefer to order AP† head-finally, i.e. with $c < a$. Similar logic results in $b < c$. From this we can generalize that adjuncts will universally be linearized before their head but after the specifier, regardless of constraint ranking.

(24) a.



b.

	(a)	Antisymmetry	HeadFinality
	abc	0	2
	bac	0	1
→	bca	0	0
	cba	0	1
	cab	0	1
	acb	0	2

This is not a desirable result, insofar as right-adjunction is quite common. Perhaps more interestingly, adjuncts are known to be extremely variable in their distribution (Ernst, 2001), both across and within specific kinds of adjuncts. Untangling this complex distribution will require other factors beyond the three constraints presented here. In some cases, the complex distribution of adjuncts has been taken to reflect more complex syntactic structure (as in e.g. Cinque, 1999). In other cases, it seems that the syntactic (or possibly prosodic) weight controls whether adjuncts are on the left or the right of their head, as in English examples like *a big dog* vs. *a dog bigger than me*. Roberts (2017) presents evidence that the positioning of adjuncts is, in fact, subject to the FOFC, so the constraints presented here still have a role to play in any analysis of their distribution, but considerably more refined tools will be needed.

Additionally, there is one aspect of the FOFC that these constraints do not capture: it only applies within certain domains. For example, German DPs appear to be head-initial, even though they are often contained inside the head-final TP; more generally, DP-internal ordering and the ordering of elements in the clausal spine seem to be inde-

pendent of one another as far as the FOFC is concerned. Biberauer et al. (2014) codify this by restricting the FOFC to looking at heads within one Extended Projection (Grimshaw, 1991). Optimal Linearization is certainly compatible with such a notion; one possible analysis would involve a stringent version of HeadFinality that is relativized not to some node but rather to an entire Extended Projection — for instance, in the case of German, one that examined only nodes in the verbal EP. There's also another possible explanation: Perhaps linearization precedes by phase (as in e.g. Fox & Pesetsky, 2005), with the possibility that the linearization constraints are ranked differently for the DP-phase and the CP-phase. Without committing to this particular analysis, I will leave further investigation of these options aside for now.

This is far from the first time that PF constraints have been proposed which make reference to the syntax. There is a large family of "prosodic faithfulness" constraints which enforce correspondence between syntactic and prosodic structures. For example, the Match constraints (Selkirk, 2011) ensure that syntactic constituents are matched by prosodic constituents that dominate the same set of terminal nodes. These constraints must have access to the syntactic structure, and in fact must even have access to the labelling of syntactic nodes in order to distinguish words, phrases, and clauses. Similarly, Clemens (2014) proposes the constraint Arg-φ, which penalizes prosodic structures in which heads and their arguments are not phrased together; this constraint needs access to selection relations.

The Optimal Linearization constraints fit this pattern: They use c-command, dominance, and labelling to choose between differently-ordered candidates. In so doing, they accomplish three main things. First, it captures the same empirical facts about linearization that are encoded in the classical Headedness Parameter, but does so using a constraint-based model consistent with how other PF-branch phenomena are treated. This frees us from having to stipulate properties like the placement of specifiers, instead allowing these properties to emerge from constraint interactions. Second, Optimal Linearization additionally allows for the disharmonic orders consistent with the FOFC without needing to stipulate any new syntactic principles — we can build syntactic trees exactly as before while still deriving the correct orders. And finally, as I'll show in the next few chapters, Optimal Linearization provides a model for interactions between linearization and phonological or prosodic markedness that allows us to capture PF displacement phenomena.

Bibliography

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut.

Abels, Klaus, & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15:25–74.

Anttila, Arto, Matthew Adams, & Michael Speriosu. 2010. The role of prosody in the english dative alternation. *Language and Cognitive Processes* 25:946–981.

Aoun, Joseph, & Elabbas Benmamoun. 1998. Minimality, reconstruction, and PF movement. *Linguistic Inquiry* 29:569–597.

Bach, Emmon. 1971. Questions. *Linguistic Inquiry* 2:153–166.

Bennett, Ryan, Emily Elfner, & James McCloskey. 2016. Lightest to the right: An apparently anomalous displacement in Irish. *Linguistic Inquiry* 47:169–234.

Biberauer, Theresa, Anders Holmberg, & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45:169–225.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier Science.

Breiss, Canaan, & Bruce Hayes. 2019. Phonological markedness effects in sentence formation. In submission.

Brugman, Johanna Christina. 2009. Segments, tones and distribution in Khoekhoe prosody. Doctoral Dissertation, Cornell University.

Cable, Seth. 2004. Predicate clefts and base-generation: Evidence from Yiddish and Brazilian Portuguese. Unpublished manuscript.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36:1–22.

Chung, Sandra, & James McCloskey. 1987. Government, barriers, and small clauses in Modern Irish. *Linguistic Inquiry* 18:173–237.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.

Clemens, Lauren Eby. 2014. Prosodic noun incorporation and verb-initial syntax. Doctoral Dissertation, Harvard University.

Clemens, Lauren Eby. 2016. Argument-ϕ: A prosodic account of pseudo noun incorporation. In *Proceedings of NELS 46*, ed. Brandon Prickett & Christopher Hammerly, 217–226.

Clemens, Lauren Eby, & Jessica Coon. 2016. Deriving verb initial word order in Mayan. Presentation at NELS 47.

Cole, Peter, & Gabriella Hermon. 1998. The typology of wh-movement. Wh-questions in Malay. *Syntax* 1:221–258.

Collins, Chris, & Edward Stabler. 2016. A formalization of Minimalist syntax. *Syntax* 19:1:43 – 78.

Coon, Jessica. 2010. VOS as predicate fronting in Mayan. *Lingua* 120:345 – 378.

Dobashi, Yoshihito. 2009. Multiple spell-out, assembly problem, and syntax-phonology mapping. *Phonological Domains: Universals and Deviations. Berlin: Mouton de Gruyter* 195–220.

Dryer, Matthew S., & Martin Haspelmath, ed. 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL `http://wals.info/`.

Dumestre, Gérard. 1984. *Dictionairre bambara-français*. Paris: Institut National des Langues et Civilsations Orientales.

Edmiston, Daniel, & Eric Potsdam. 2016. Linearization at pf: Evidence from malagasy extraposition. Presented at NELS 47, University of Massachusetts Amherst.

Edmiston, Daniel, & Eric Potsdam. 2017. Linearization at pf: Evidence from malagasy extraposition. In *Proceedings of NELS 47, University of Massachusetts Amherst*, ed. Andrew Lamont & Katerina Tetzloff.

Elfner, Emily. 2012. Syntax-prosody interactions in Irish. Doctoral Dissertation, University of Massachusetts Amherst.

Embick, David, & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Ernst, Thomas. 2001. *The syntax of adjuncts*, volume 96. Cambridge University Press.

Fox, Danny, & David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical linguistics* 31:1–45.

Fox, Danny, & Takahashi Shoichi. 2005. MaxElide and the re-binding problem. In *Proceedings of SALT XV*, ed. Effi Georgala & Jonathan Howell, 223–240. Cornell University, Ithaca, NY: CLC Publications.

Göbbel, Edward. 2007. Extraposition as PF movement. In *Proceedings of WECOL 2006*, 132 – 145. California State University, Department of Linguistics Fresno.

Grimshaw, Jane. 1991. Extended projection. Unpublished manuscript, Brandeis University.

Haacke, Wilfrid H. G. 1999. *The tonology of Khoekhoe (Nama/Damara)*. Kŏln: Rŭdiger Kŏppe Verlag. Quellen zur Khoisan-Forschung 16.

Haacke, Wilfrid HG. 2006. Syntactic focus marking in Khoekhoe ("Nama/Damara"). *ZAS Papers in Linguistics* 46:105–127.

Hahn, Michael. 2013. Word order variation in Khoekhoe. In *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar*, 48–68.

Halpern, Aaron Lars. 1992. Topics in the placement and morphology of clitics. Doctoral Dissertation, Stanford University Ph. D. dissertation.

Jadoul, Y., B. Thompson, & B. de Boer. 2018. Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics* 71:1–15.

Johnson, Kyle. 2016. Towards a multidominant theory of movement. Lectures presented at ACTL, University College.

Kandybowicz, Jason. 2006. On fusion and multiple copy spell-out: The case of verb repetition. In *The copy theory of movement on the PF side*, ed. Jairo Nunes & Norbert Corver. Oxford University Press.

Kastenholz, Raimund. 2003. Auxiliaries, grammaticalization and word order in Mande. *Journal of African Languages and Linguistics* 24:31 – 53.

Kayne, Richard S. 1994a. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.

Kayne, Richard S. 1994b. *The antisymmetry of syntax*, volume 25 of *Linguistic Inquiry Monographs*. MIT Press.

Keenan, Edward. 1976. Remarkable subjects in Malagasy. In *Subject and Topic*, ed. Charles Li, 247 – 301. New York: Academic Press.

Keenan, Edward. 1995. Predicate-argument structure in Malagasy. In *Grammatical Relations: Theoretical approaches to empirical questions*, ed. C. Burgess et al., 171 – 216. Stanford: CSLI.

Koopman, Hilda. 1984. *The syntax of verbs*. Dordrecht, The Netherlands: Foris Publications.

Kusmer, Leland Paul. 2018. Optimal Linearization: Word order typology with violable constraints. In submission.

Kusmer, Leland Paul, & Kerrianne Devlin. 2018. Default case in Khoekhoegowab. Presented at the Annual Conference on African Linguistics.

Law, Paul. 2007. The syntactic structure of the cleft construction in Malagasy. *Natural Language and Linguistic Theory* 25:765–823.

Legendre, Géraldine. 1998. Second position clitics in a verb-second language: Conflict resolution in Macedonian. In *Proceedings of ESCOL 1997*, 138 – 149.

López, Luis. 2009. Ranking the linear correspondence axiom. *Linguistic Inquiry* 40.2:239–276.

Lukas, J., & G. Nachtigal. 1939. *Die sprache der buduma im zentralen sudan: auf grund eigener studien und des nachlasses von g. nachtigal*. Abhandlungen für die Kunde des Morgenlandes. Deutsche morgenländische gesellschaft, Kommissionsverlag F.A.Brockhaus. URL `https://books.google.com/books?id=ZzQpOwAACAAJ`.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, & Morgan Sonderegger. 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*.

McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge University Press.

McCarthy, John J., & Alan Prince. 1994a. Generalized alignment. In *Yearbook of morphology 1993*, 79 – 153. Dordrecht: Springer.

McCarthy, John J, & Alan S Prince. 1994b. The emergence of the unmarked: Optimality in prosodic morphology. In *Proceedings of the North East Linguistic Society 24*. GLSA (Graduate Linguistic Student Association), Dept. of Linguistics, University of Massachusetts.

McCloskey, James. 2011. The shape of Irish clauses. In *Formal approaches to Celtic linguistics*, ed. Andrew Carnie, 143 – 178. Newcastle upon Tyne: Cambridge Scholars Publishing. URL `http://ohlone.ucsc.edu/jim/papers.html`.

Overfelt, Jason. 2015. Rightward movement: A study in locality. Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Percus, Orin. 1993. The captious clitic: Problems in Serbo-Croatian clitic placement. Phonology generals paper.

Petronio, Karen, & Diane Lillo-Martin. 1997. Wh-movement and the position of spec-CP: Evidence from American Sign Language. *Language* 73:18–57.

Prince, Alan, & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Blackwell Publishers.

Roberts, Ian. 2017. The Final-over-Final Condition and Adverbs. In *The Final-over-Final Condition*, ed. Michelle Sheehan, Theresa Biberauer, Ian Roberts, & Anders Holmberg, chapter 6, 97 – 120. MIT Press.

Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Sauerland, Uli, & Paul Elbourne. 2002. Total reconstruction, PF-movement, and the derivational order. *Linguistic Inquiry* 33:283–319.

Schütze, Carson T. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In *MIT Working Papers in Linguistics*, ed. Andrew Carnie, Heidi Harly, & Tony Bures, volume 21, 373–473.

Selkirk, Elisabeth. 2011. The Syntax-Phonology Interface. In *The Handbook of Phonological Theory*, ed. John Goldsmith, Jason Riggle, & Alan Yu, volume 2, chapter 14, 435–483. Oxford: Blackwell Publishing.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts, & Anders Holmberg. 2017. *The Final-over-Final Condition: A Syntactic Universal*, volume 76. MIT Press.

Shih, Stephanie S., & Kie Zuraw. 2017. Phonological conditions on variable adjjective and noun word order in Tagolog. *Language* 93:317 — 352.

Sichel, Ivy. 2014. Resumptive pronouns and competition. *Linguistic Inquiry* 45:655–693.

Team, R Core. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Washburn, Paul. 2001. A Minimalist approach to Khoekhoe declaratives. *Cornell Working Papers in Linguistics* 18:28–56.

Watanabe, Akira. 1992. Subjacency and S-structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1:255–291.

Weir, Andrew. 2015. Fragment answers and exceptional movement under ellipsis: a PF-movement account. In *Proceedings of NELS*, ed. Christopher Hammerly & Brandon Prickett, volume 45, 175 – 188.

Yip, Moira. 2002. *Tone*. New York: Cambridge University Press.

Wouter Zwart, Jan. 2011. Structure and order: asymmetric Merge. In *The Oxford handbook of linguistic minimalism*, 96 – 118.

Zwarts, Frans. 1998. Three types of polarity. In *Plurality and quantification*, ed. F. Hamm & E.W. Hinrichs, 177 – 238. Dordrecht: Kluwer.