Prosodic Linearization in Khoekhoe and beyond

Leland Paul Kusmer

1 Introduction: Prosodic linearization

Understanding the relationship between syntactic structure and linear strings is a challenge for modern syntactic theory. The most complete and widely accepted models — namely, the Headedness Parameter and the Linear Correspondence Axiom (LCA) (Kayne, 1994) — each capture aspects of the relationship, but are either too permissive (a Headedness Parameter relativized to individual categories permits nearly any linear order which keeps phrases contiguous) or too restrictive (the LCA is well-known for ruling out head-final configurations generally). Subsequent linearization models (e.g. Dobashi 2004, 2009; Fox & Pesetsky 2005) have typically relied on one of these two main models in order to capture other, non-order-related effects (e.g. island constraints on movement) — a valuable contribution, to be sure, but not one that significantly deepens our understanding of the linearization process itself.

In recent years an interesting new model has begun to emerge. Elfner (2012) discusses an anomalous displacement in Irish in which prosodically-light object pronouns are displaced to the right with no change in meaning; this displacement is capable of placing the pronoun in many syntactically improbable locations (e.g. inside conjunctions), which makes analyzing it as syntactic movement highly undesirable. Elfner instead proposes a model of the syntax-prosody interface in which the prosody is capable of altering the linearization to suit its own needs, displacing pronouns to the right in order to achieve a more well-formed prosodic structure. A different but related model has been proposed by Richards (2016), in which prosodic structure is built simultaneous to the syntax and can condition syntactic operations (including, for Richards, a change to head-finality) for its own reasons. These models share in common the hypothesis that linearization and prosody are inextricably related, with the pressures on one capable of conditioning the output of the other. I will refer to this hypothesis as prosodic linearization.

Prosodic linearization is a promising new avenue, but further research is required to develop a specific model. The model proposed by Elfner (2012) is a constraint-based model in keeping with modern prosodic and phonological theory (Selkirk & Lee, 2015), which has the advantage of making concrete typological predictions about possible prosodies and linearizations. However, as will be shown in my dissertation, Elfner's model both under- and overpredicts, generating a large number of unattested linearization patterns while also linking some attested patterns to only one possible prosodic structure.

Whereas Elfner begins with a specific linearization phenomenon which is apparently syntactically-anomalous, Richards (2016) takes as his goal the explanation of certain syntactic facts, including head-finality and the EPP. In order to do this, he proposes a model which dramatically reworks our current understanding of the syntax-phonology interface, allowing prosodic factors to influence operations in the narrow syntax. While this model does have a number of desirable typological consequences, there are other cases where the empirical validity of the relevant generalizations are not clear. Richards links various syntactic processes with 'prosodic activity' (which edge of a prosodic unit receives phonological marking), predicting a strong link between prosodic marking and (for instance) movement. It is not clear whether such a close link between the surface prosody and the underlying syntax can be borne out; furthermore, there are many languages in which prosodic activity occurs at both edges of a phrase, and in these cases it isn't at all clear what predictions Richards' system will make.

The primary aim of this dissertation is to propose a new model of prosodic linearization. In doing so, I will draw on two sources of data. The first is word order typology, including both headedness patterns and movement facts. I take word-order generalizations, the core of what any linearization scheme must capture, to represent the 'unmarked' cases in which prosodic constraints and linearization constraints do not interact. The second source of data will be prosodic displacement of the kind noted by Elfner (2012), which I take to represent those corner cases in which prosodic constraints oppose linearization ones; these give us an invaluable opportunity to see what aspects of linearization can be overridden and which cannot. My investigation will represent a significant contribution to the literature in that the literatures on linearization and on prosody have typically proceeded separately; I will bring together recent empirical and theoretical results from both domains. Insofar as the prosodic linearization hypothesis proposes that linearization is a PF phenomenon, it seems desirable to model it using the theoretical tools phonology; I take this to mean that it should be modeled in a constraint-based framework like Optimality Theory (Prince & Smolensky, 1994/2008), which has the additional advantage of making clear and testable typological predictions. As a specific case study, I will consider the linearization of tense particles in Khoekhoe, which seem to be linearized primarily by prosodic factors; I aim to build a model that will help us understand the specific role that prosody plays in Khoekhoe linearization and also the broader typology of word order and prosody cross-linguistically.

The remainder of this prospectus is structured as followed. In section 2 I'll offer some theoretical background on the problem of linearization and the prosodic linearization hypothesis, with a eye towards the specific problems I propose to solve in this dissertation. In 2.1 I'll argue that linearization is an unsolved problem with broad consequences for our understanding of the architecture of grammar; 2.2 will present the reasons to consider the prosodic linearization hypothesis, while 2.4 will discuss the technical problems facing any specific implementation of prosodic linearization. In 3, I'll present some background on the Khoekhoe language and make the argument that this language presents a unique case-study for prosodic linearization models. Finally, in section 4 I'll outline the general approach I will take to modeling prosodic linearization, along with a preliminary proposal for some specific constraints to be used.

2 Background

2.1 *Linearization as an unsolved problem*

I propose the following desiderata for a linearization scheme: First, it should predict which items will be linearized before their head as specifiers. Second, it should predict a general consistency of headedness while allowing for specific categories to be linearized exceptionally, to account for mixed-headedness. Finally, it give some insight into the linearization of moved items, allowing them to appear either 'high' (in their landing site) or 'low' (in their origin), but not both. None of the linearization schemes in the current literature adequately meet all three requirements.

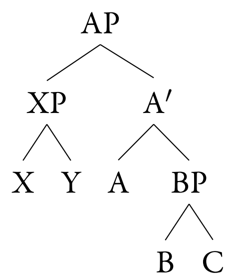
The earliest and best-known solution to the problem of linearization is the Headedness Parameter. The strongest form of this hypothesis proposes a language-wide parameter of the form in (1):

(1) When linearizing a phrase XP with immediate daughters X and Y, with X the head of XP, the following precedence relations hold:
   a. If Y is a specifier: All terminal strings dominated by Y precede all terminal strings dominated by X
   b. Otherwise:
      (i) If the Headedness Parameter is set to Initial: All terminal strings dominated by X precede all terminal strings dominated by Y.
      (ii) If the Headedness Parameter is set to Final: All terminal strings dominated by Y precede all terminal strings dominated by X.

Note first of all the exemption clause for specifiers, allowing the specifier position to be always left. Early versions of the Headedness Parameter allowed for specifiers to be linearized by a second parameter, but modern analyses have converged on left-specifiers. This requires the stipulation above that specifiers simply be exempt from the parameter. In fact, the situation is worse than this — without any commitment to a special role for specifiers in general (e.g. Spec-Head agreement), the Minimalist Program is left without a coherent way to identify what items are specifiers or not, and the exemption becomes somewhat circular: We know that something is a specifier because it is on the left, and it's on the left because it's a specifier.

More concretely, the strong Headedness Parameter hypothesis is descriptively inadequate: One needn't look any further than German to find an example of apparently mixed-headedness. There must be some means of relativizing the Headedness Parameter to particular categories, but once such relativity is introduced the parameter becomes nearly unrestricted. Could a language set the parameter differently for each category? Consider the tree in (2); under a relativized Headedness Parameter, it permits any of the six orders listed in (2-b). If the parameter can be independently relativized to each category, we lose exactly the generalizations that the parameter was intended to capture — namely that headedness is largely consistent within a language.

(2) a.



b. Orders:
(i) XYABC
(ii) YXABC
(iii) XYBCA
(iv) YXBCA
(v) XYACB
(vi) XYCBA

By contrast, Kayne's (1994) LCA provides some insight into the nature of specifiers, but at the cost of being overly restrictive. The LCA may be succinctly stated as in (3), and in essence requires that all specifiers be initial while also requiring all heads to precede their complements.

(3) If any two nodes α and β are such that α c-commands β and not the reverse, then all terminal strings dominated by α must precede all terminal strings dominated by β.

Of the six orders predicted by the Headedness Parameter for the tree in (2), the LCA allows only the first. Under such a system, head-final languages simply cannot exist; rather, apparently head-final structures must in fact be arrived at by movement. Without any independent evidence, it seems undesirable to say that the (slim) majority of the world's languages that are head-final involve extensive movement even in relatively unmarked root clauses.

The problem grows worse when we consider the linearization of structures involving movement. Minimalist theories of movement are quite diverse, but probably most contemporary analyses assume some form of copy-theory, in which the moved item in some sense exists in both its origin and landing site (see, e.g. Corver & Nunes 2007). The linearization of such a structure is significantly problematic: If the item is in some sense in both places, which place 'counts' for linearization? How do we avoid spelling out the item twice? Typical solutions involve deleting the lower copy (as assessed by c-command), but in some cases we apparently want the higher one deleted as well (as, perhaps, is the case in WH-*in situ*). These same theoretical issues plague the linearization of multidominant trees, as well.

In sum, both the Headedness Parameter and the LCA are only partially successful on the three desiderata laid out above, with both in particular requiring significant extra stipulations when a copy-theory of movement (or multidominance) is introduced. In this sense, linearization is still an open theoretical problem.

2.2 *Prosodic linearization*

Within the framework of the Minimalist Program (Chomsky, 1995) and the Y-model architecture of grammar, linearization seems like a good candidate for a PF phenomenon. Insofar as linearization concerns only how sentences are actually pronounced, it seems sensible to place linearization well after the narrow syntax is finished. This is the fundamental rationale behind the prosodic linearization hypothesis: If linear order is only determined at PF, then it stands to reason that it might interact with other PF phenomena. Put another way, the only thing that causes syntactic structures to be linearized is the need to fit them into a phonological structure; as such, constraints on the construction of that structure may influence the specific linear order assigned to a given syntactic tree. Phonology in general, and prosody in specific, is full of constraints on what makes a better or worse structure to utter, constraints which are utterly independent of the syntax. Under the prosodic linearization hypothesis, we expect to find that these constraints may influence the linearization of given tree structures.

There is preliminary evidence that such an approach should be considered. In perhaps the clearest example to date, Elfner (Elfner, 2012; Bennett, Elfner, & McCloskey, 2016) identifies a kind of anomalous displacement in Irish which seems motivated by phonological properties of particular utterances. The phenomenon in question concerns light pronouns. In certain contexts, these pronouns may appear arbitrarily far to the right of their canonical base position; in general, they are displaced to the right edge of a prosodic constituent, which may or may not align with a syntactic constituent. The example in (4) illustrates a weak pronoun (in this case, an expletive subject) postposing into the middle of a disjoint, for instance — surely not a kind of movement we wish to introduce into the narrow syntax![1]

(4)

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."

It's particularly noteworthy that the postposing is conditioned by the phonological status of the object: Only the weak, unstressed object pronouns may postpose. When the strong, stressed pronoun is used, it obligatorily stays in its base position. This, along with the fact that the landing site seems prosodically conditioned, strongly hints that this is a prosodic, not syntactic, process. Elfner proposes that this movement is driven by an independently-attested preference to not have weak items at the start of prosodic phrases; for independent reasons, the object position is one that must start a phrase. This is a constraint that has nothing to do with the syntax — it is a purely surface-oriented constraint. This is data then works in favor of the prosodic linearization hypothesis: Constraints on what makes a good prosody are forcing the sentence to be linearized in an alternative way.

In the case discussed so far, the prosodic linearization hypothesis is supported by the fact that the prosodic status of particular words is relevant to determining their position in the string. We might term cases like these *prosodically-driven displacement*, insofar as they apparently represent cases in which prosodically-light items are linearized differently than prosodically-heavy ones. But if the prosodic linearization hypothesis is right, then the effects of prosody on linearization should not be confined to displacement — prosodic structure and linear order should always (have the potential to) interact, even when the prosody determines the same order regardless of the weight of the individual words. In this way, prosodic linearization has the potential to give us new insight into all aspects of the linearization problem.

For example, Richards (2016) has developed a theory of prosodic linearization which follows very different lines than what has been outlined here so far, but shares the core notion that independent con-

[1] Example taken from Bennett, Elfner, & McCloskey (2016); the underlines indicate the 'normal', non-displaced position of the pronoun é.

straints on the prosody can determine which words are pronounced where. For Richards, the prosodic constraints in question involve a pressure to have certain syntactic relationships be mirrored by prosodic ones. For instance, he argues that when two heads are in an agreement relationship, there is pressure to group them together and to mark the goal in the prosody; given an independent choice about which edges of prosodic phrases receive phonological marking, this allows Richards to force heads in a probe-goal relationship to follow certain orders. Based on this and other constraints, his system attempts to provide prosodic explanations for common problems in linearization, including the nature of specifiers and the ordering of moved items. While his system assumes the grammar to have a fairly non-standard architecture (for Richards, prosody is built simultaneous to the narrow syntax and thus can feed syntactic processes like movement), it serves to illustrate the promise of the prosodic linearization hypothesis: That we might bring the extensive tools of prosodic theory to bear on certain broad issues in syntax and find an answer with some explanatory power.

2.3 *Other possible examples*

Clear empirical cases of prosodically-driven displacement, as in Elfner (2012), are few and far between. Some other cases in which a prosodic-displacement analysis has been proposed include:

- Halpern (1992); Bošković (2001), & many others: 2$^{\text{nd}}$ position clitics in Serbo-Croatian sometimes interrupt syntactic constituents, but never interrupt prosodic ones, making it seem as though they've been positioned prosodically rather than syntactically.
- Edmiston & Potsdam (2016): Malagasy complement clauses are displaced to the right, possibly for prosodic reasons.
- Clemens (2016): Pseudo-noun incorporation in Niuean results from pressure to keep the verb and noun prosodically grouped; Clemens & Coon (2016) extends this to various Mayan languages.
- Anttila et al. (2010): Prosodic factors can influence the choice of syntactic frame for English ditransitive. However, no specific model is proposed.
- Hahn (2013) proposes that some tense / aspect makers in Khoekhoe are prosodically displaced; see section 3 for more details.

All of the above models share in common that the prosody is capable of adjusting linear order. The Clemens (2016) proposal (and the work with Coon following it) are unique in that they propose a specific constraint on the syntax-prosody mapping which is responsible for relinearizing some NPs; this contrasts with e.g. Elfner (2012), in which it is prosodic markedness constraints that are responsible for prosodic displacement. The Edmiston & Potsdam (2016) proposal is also unique in that it propose that very heavy elements (i.e. clauses) are displaced for prosodic reasons, rather than light ones.

2.4 *Challenges for prosodic linearization*

Constructing a specific model of prosodic linearization is not without its challenges. On the one hand, there is the danger of overgeneration. It is notable that all of the cases of what I have termed prosodically-driving displacement involve the displacement of prosodically-light items, generally smaller than the minimal prosodic word. The literature on prosodic well-formedness has often relied on constraints like StrongStart or EqualSisters (Selkirk, 2011; Myrberg, 2013) as a means of enforcing parity between sister nodes in a prosodic structure — all things being equal, many languages prefer not to have a light unit as the sister to a heavy unit. But these constraints are just as likely to penalize a (prosodic) phrase being a sister to a clause (/ intonational phrase), predicting that even these heavier units may undergo prosodically-driven displacement in some cases. While there are some displacement processes that seem to target heavy units (like Heavy NP Shift), these often seem to have a different character than the examples of prosodically-driven displacement discussed above, insofar as they may sometimes have semantic or syntactic effects; it isn't clear that we want displacement of this kind to be derived at the PF stage, rather than in the narrow syntax. As such, one of the challenges of prosodic linearization is to ensure that only the lightest of elements can undergo displacement.

On the other hand, there is also a danger of undergeneration. If the prosody and the linearization are calculated simultaneously, there is danger of having some linear orders be associated with one and only one prosody. Empirically, both head-initial and head-final languages seem to display a range of possible prosodies; we want to make sure that our model of prosodic linearization can derive all of these combinations. Stated another way, while the existence of prosodically-driven displacement strongly indicates that prosodic structure building and linearization can interact, they still seem to be somewhat independent processes, and we want our model to capture this.

As an example of these pitfalls, consider the model from Elfner (2012). Elfner couches her model in Optimality Theory[2], with the Match constraints (Selkirk, 2011) as the primary mechanism of prosodic structure building, with Binarity enforcing binary branching in the prosody. She motivates the displacement of the weak pronoun with the constraint StrongStart, which penalizes prosodic phrases whose leftmost element is prosodically lighter than their second element. In order to allow the prosody to affect the linearization, she modifies Gen to generate not just all possible prosodifications of a particular string, but in fact all prosodifications of all permutations of that set of words; that is, the candidates under consideration consist of every possible order the words could be produced in along with every possible prosodic tree. She thus needs some 'faithfulness' constraint penalizing those orders that differ from what we intuitively expect to find given the syntactic structure; she uses Kayne's LCA for this purpose, simply treating it as a violable constraint.[3]

While Elfner's model works admirably for the Irish data, it makes typological predictions which show both under- and overgeneration. Optimality Theory is an inherently typological theory (Prince & Smolensky, 1994/2008), in that every ranking of constraints is hypothesized to correspond to some real language; by testing all orderings of constraints we can see the predicted language types. I have shown that the constraint-set used by Elfner produces a set of 50 possible languages[4]. Perhaps surprisingly, many of these languages are head-final. Given that the linearization constraint used (namely the LCA) predicts only head-initial languages, we might at first think this a success — head-finality emerges as an epiphenomenon based on the interaction of StrongStart and Match. However, I think that this close linking of prosody and headedness is not desirable: My typology predicts that all head-final languages will have a strictly uniform prosody, which does not seem to be empirically true. This is a case of undergeneration as discussed above — prosody and linearization in this scheme are too closely linked.

On the other hand, Elfner's constraint set also overgenerates in that it allows prosodically-driven displacement in a larger range of situations. For example, there are some languages in the typology in

[2] Actually, Elfner uses Harmonic Grammar, which differs from Optimality Theory in that it weights, rather than ranks, its constraints; this enables her to derive certain other phenomena in Irish prosody. For the specific issues we'll consider here, OT and HG are equivalent.

[3] In the later paper, Bennett, Elfner, & McCloskey (2016), the authors replace the LCA with a cover-constraint which simply penalizes deviations from some assumed order without telling us how to calculate that order. In both cases, the linearization is effectively done outside of the prosody, with the relevant constraint penalizing deviations from the string order calculated earlier.

[4] These 50 languages are only across a relatively limited set of possible inputs, which don't involve movement. If we were to consider more inputs, we would find a much larger typology.

which single-word complements follow their head but multi-word complements precede their head. Put another way, in these languages heads are prosodically-displaced rightward after their complement exactly when their complement is heavy enough to form a prosodic phrase; if the complement is too small to form a (binary) prosodic phrase, normal head-initial order is followed. This kind of weight-driven mixed headedness with respect to complements is not something we have observed empirically.

This highlights a broader challenge facing prosodic linearization, namely the lack of empirical data. Well-established cases of prosodically-driven displacement are thin on the ground, so we cannot really say with any confidence what this typology should look like. Worse, most of the currently-established cases are, like Elfner's example, quite marginal within the language in question — only a small number of words undergo displacement, and then only in a small range of syntactic structures. This makes it difficult to test what prosodic and syntactic factors might influence the availability of displacement, or the landing sites for such displacement. Without more extensive cases, it becomes hard to evaluate the merits of particular models of prosodic linearization.

3 Case study: Khoekhoe

Khoekhoegowab, also known as Khoekhoe, is one of the official languages of Namibia. With about 200,000 speakers, it is the largest language in the Khoesan family; despite this, the existing literature on Khoekhoe is not extensive, consisting primarily of a small handful of grammatical sketches. Syntactically, the language is strongly head-final both in the DP and CP domain[5]. Most root clauses have a second-position 'clause type' particle; as with most language showing second-position effects, the prefield is typically occupied by the subject, but other phrasal items may be raised there, displacing the subject down into the middlefield. This general clause structure is shown in (5).[6]

(5) a. taras ge ǁari ǂkhanis -a go mā -te
woman DECL yesterday book -A TAM give =me
"The woman gave me a book yesterday."

b. ǁari -s ge taras -a ǂkhanis -a go mā -te
yesterday -3SF DECL woman -A book -A TAM give =me
"Yesterday the woman gave me a book."

Note that, in (5), tense, aspect, polarity, and mood (hereafter TAM) are indicated by the particle *go*, which appears immediately before the verb. These TAM particles, along with the second-position clause type particle, are the only exceptions to the strong head-finality in Khoekhoe. Following an observation made by Hahn (2013), I will argue based on original fieldwork that the placement of these particles is an example of prosodically-driven displacement.

My fieldwork was conducted in the summer of 2017 in Windhoek, Namibia, and consisted primarily of elicitation with 5 native speakers of the Nama variety of Khoekhoe. My speakers consisted of four women and one man, with ages ranging from 20 to 50 years; all but one were native to the south of Namibia (where the Nama dialect is dominant); one was born in the north speaking the Damara dialect, but had spent most of her childhood in Windhoek. All speakers were recorded using a Zoom H5 recorder and a Shure SM10A head-mounted microphone. In addition to elicitation, I conducted one production experiment on the prosody of displaced preverbal particles; the design of this experiment will be detailed in section 3.3.

[5] Surprisingly, one class of serial verb construction in Khoekhoe seems to be head-initial; this is the only clear exception to head-finality that isn't likely a case of prosodically-driven displacement.

[6] In the Khoekhoe examples, the following abbreviations are used: (i) DECL — declarative clause marker; (2) -A — oblique case marker; (3) TAM — tense, aspect, & mood particle; 1/2/3 S/G F/M/N — person, number, & gender marking. All Khoekhoe examples are presented in the standard orthography.

3.1 *TAM particles*

Tense, aspect, (some) mood, and polarity in Khoekhoe are all expressed by a set of particles which are (to varying degrees) independent of the verb. The majority of these TAM particles are phonologically enclitic. For example, consider the alternation of the imperfect aspect particle *ra / ta* based the final segment of the preceding word, illustrated in (6). In (b), the lack of a clause-type marker in questions requires the TAM marker to encliticize to the subject pronoun -TS[7] and surface in its *ta* allomorph.

(6) a. stors ǁkha -ts ge ra !gû
store to -2SF DECL TAM go
"You are going to the store."

b. stors ǁkha -ts ta !gû?
store to -2SF TAM go
"Are you going to the store?"

The vast majority of Khoekhoe TAM particles are preverbal. In all of the preceding examples, these particles are in their default position immediately before the verb. Hahn (2013) first observed that in fact these particles are not fixed in place, and can 'float' left through the middlefield, encliticizing to any constituent therein. This is illustrated in (7), taken from Hahn's paper.

(7) taras ge (go) ǁari (go) ǂkhanis -a (go) mā -te
woman DECL (TAM) yesterday (TAM) book -A (TAM) give -me
"The woman gave me the book yesterday."

In my own fieldwork, consultants consistently volunteered sentences with the TAM particle in immediately-preverbal (hereafter, IPV) position, however, they would readily accept sentences with particles in other positions, and were not able to discern any semantic or pragmatic difference between them.[8] Notably, in both questions and embedded clauses where there is no 2nd position particle, speakers were more likely to volunteer TAM particles in 2nd position, but would generally offer IPV as an alternative. Other, intermediate positions were also accepted. Examples of TAM particles in 2nd position are given in (8).

(8) a. Dandagoba go tarasa ǂkhanisa mā?
D. TAM woman book give
"Did D. give the book to the woman?"

b. ... Dandagoba go tarasa ǂkhanisa mā -sa
D. TAM woman book give that
"...that D. gave the woman book".

In addition to these preverbal particles, however, there is a small set of TAM particles which obligatorily follow the verb. For example, the negative future *tide* is always postverbal:

(9) a. Namas ge taras -a mā tide
Nama DECL woman -A give TAM
"The Nama will not give (anything) to the woman."

b. *Namas ge (tide) taras -a (tide) mā
Nama DECL (TAM) woman -A (TAM) give

[7] This subject 'pronoun' is in fact itself a second-position clitic rather than a full pronoun.

[8] In fact, when presented (either orthographically or orally) with two sentences differing only in TAM placement, speakers were frequently unable to detect any difference until it was pointed out.

Table 1: Preverbal TAM particles

	IPA	Gloss
a	[ra]	present copula
ra / ta	[ra] / [tʰa]	imperfect
ge	[ke]	remote past
go	[ko]	recent past
ni	[ni]	future
ta	[ta]	negative non-finite
ga	[ka]	irrealis[9]
Compound particles:		
gere	keɾe	remote past imperfect
goro	koɾo	recent past imperfect
nira	niɾa	future imperfect
gara	kara	irrealis imperfect

Table 2: Postverbal TAMparticles

	IPA	Gloss
tama	[tama]	non-future negative
tide	[tite]	future negative
i	[iː]	non-present copula
hâ	[hãː]	perfective

A full list of the TAM particles of Khoekhoe is presented in Tables 1 & 2.

Hahn argues that the distribution of TAM particles in Khoekhoe is best analyzed as prosodically-driven displacement. Specifically, he argues that post- and pre-verbal particles do not form morphosyntactic natural classes, but prosodic ones: Light, monomoraic particles like *go* 'past' or *ra /ta* 'imperfect' are preverbal, while bimoraic particles like *tide* 'negative future' are postverbal. In this conception, TAM particles belong to a single, unified class (i.e. all TAM particles are exponents of the T° head or similar); some prosodic pressure (left unidentified by Hahn) causes the light ones to be displaced left. This is a particularly attractive analysis given that all Khoekhoe root words are minimally bimoraic (Brugman, 2009); the preverbal particles are exactly those which are too small to be prosodic words on their own.[10]

Further support for this analysis comes from the fact that the preverbal particles can and do occur after the verb, but only in specific prosodic conditions. In her extensive dissertation on tone and prosody in Khoekhoe, Brugman (2009) notes that verbs (in their base position) always behave as though the were non-initial in their prosodic phrase.[11] In verum focus contexts, however, the verb will be 'promoted' to its own prosodic phrase, thus showing phrase-initial prosody. In exactly this context, prever-

[10] The 'compound' particles (tense + *ra* imperfect) pose a potential problem for this analysis — they are indisputably bimoraic, yet behave exactly like the other preverbal particles. An answer perhaps lies in the fact that they are the only multimorphemic TAM particles; perhaps they really are two different syntactic heads, and the only reason that the two morphemes are not separable is that whatever prosodic pressure drives displacement of the tense particle to a particular position also drives the displacement of the aspect marker. This analysis will need to be worked out in more detail.

[11] For more detail on Khoekhoe prosody, see section 3.2

bal TAM markers may occur post-verbally:

(10) ǂkhanisa -b ge ǁama go!
book -3SM DECL buy TAM
"He DID buy the book!"

By contrast, there are no such pragmatic or prosodic restrictions on postverbal TAM particles occurring in their base position. These facts are compatible with an analysis in which there is some ban on light elements following the verb, except when the verb is 'promoted' to its own prosodic phrase.

There are three syntactic conditions in which normally-preverbal TAM particles are required to appear after the verb. All involve fronting of some verbal projection; I will term them **verb-fronting, VP-fronting,** and **VP-inversion**. In the verb- and predicate-fronting constructions, some constituent containing the verb is moved into the prefield, stranding the subject in the middlefield; the difference between them is whether this constituent includes the TAM particle or not. In the verb-fronting construction, only the verb is moved into the prefield, stranding the TAM particle in the middlefield; interestingly, in this construction the TAM particle is constrained to 2^{nd} position, i.e. as close to the verb as possible without leaving the middlefield. This is illustrated in (11).

(11) ǂû -b ge ra Dandagoba (*ra) ǁgan-e (*ra).
eat -3SM DECL TAM D. (*TAM) meat (*TAM)
"As for eating, D. is eating meat."

In the predicate-fronting construction, the TAM particle moves with the verb into the prefield. This is presumably remnant movement of e.g. TP after scrambling the arguments of the verb elsewhere. In this case, the TAM particle is required to follow the verb. This is not surprising, as the light TAM particles are enclitics and thus need a host to their left. This is illustrated in (12).

(12) ǂû ra -b ge Dandagoba ǁgan-e.
eat TAM 3SM DECL D. meat
"Dandago is eating meat."

It's worth noting that in both of the above cases, the verb will be prosodic-phrase-initial (by virtue of being sentence-initial). This parallels the constraint on verum focus light TAM particles following the verb.

The final construction in which light TAM particles must follow the verb is the inversion construction. This construction appears to be a variation on a fronting construction in which the entire TP is moved left of the prefield; I'll term this the **subject-final** construction (in contrast to the normally verb-final order Khoekhoe employs). It is illustrated in (13).

(13) ǁgan-e ra ǂû Dandagob ge
meat TAM eat D. DECL
"As for eating meat, D. is doing it."

In the inversion construction, TP is fronted just as in the subject-final construction, but also the verb is moved to the front (by some as-yet-unknown mechanism). This is illustrated in (14). Interestingly, in the inversion construction, all TAM particles (including light ones) are constrained to be immediately post-verbal. (In contrast, the subject-fronting construction without inversion allows the TAM to move left, as long as there is still a prosodic host further to the left.)

(14) ǂû ra ǁgan-e (*ra) Dandagob ge
eat TAM meat (*TAM) D. DECL

Table 3: Tone melodies

Citation		Sandhi	
SL-L	Low-Rising	SL-L	Low-Rising
SL H	Superlow High	L-SL	Low-Falling
L H-SH	Low High-Rising	L	Low
SH	Superhigh	H	High

"As for eating meat, D. is doing it."

It is perhaps notable that in the inversion construction the verb is also initial in its prosodic phrase, by virtue of being at the left edge of the sentence. This again supports the idea that the placement of light TAM particles is dependent on the prosody of the rest of the sentence.

3.2 Tone & Prosody

The relationship between tone and prosodic structure in Khoekhoe is extensively discussed in Brugman (2009). Tone is contrastive in Khoekhoe, and all lexical items are sorted into one of six categories.[12] Each category has two tonal melodies associated with it; tonal melodies are derived from a four-level tone system and include level tones, rises, and falls, but no complex contours. The two tone melodies associated with each category are termed the 'citation' and 'sandhi' melodies; citation melodies occur in 'prosodically strong' positions (meaning, for the most part, at the left edge of prosodic phrases), while sandhi melodies occur elsewhere. The relation between citation and sandhi melodies is apparently arbitrary: While Brugman notes that the inventory of sandhi melodies contains fewer contrasts and no rising tones (which are typologically more marked than falls), there is no apparent way to predict from the form of a citation melody what the associated sandhi melody will be. The citation and sandhi melodies are illustrated in Table 3.

Brugman finds that the citation form is used for the left-most item in a phrase, while the sandhi form is used for all others. For example, within the DP domain the left-most prenominal modifier receives citation while everything else (including the head noun) receives sandhi. This is illustrated in (15) (taken from Brugman 2009) with a set of words taken from the High-Rising citation melody category; in each case, the first word in the example is High-Rising, while all following words are Low.

(15) a. súű̋gu 'pots'
b. |ápa̋ sùugu 'red pots'
c. !ánı̋ |àpa sùugu 'six red pots'
d. ǁnáa̋ !àni |àpa sùugu 'those six red pots'

From a prosodic perspective, it seems that each DP is mapped onto some prosodic unit, and that the leftmost prosodic word in this unit is assigned the citation melody; Brugman identifies this unit as the Prosodic Phrase. DPs robustly map to prosodic phrases, but interesting VPs do not: Brugman notes

[12] Both Brugman and Haacke claim that there are more than six categories, with the additional tonal categories being rare / antiquated. I have not found any examples of such categories, and it's not clear to me how productive sandhi alternations are supposed to be in those categories.

that verbs[13] receive sandhi form whenever two conditions are met: (1) a second position clause-type particle is present, and (2) no light, preverbal TAM marking follows the verb. This means that in all matrix indicative clauses, verbs behave as though they are non-initial in a prosodic phrase; in all other configurations, they behave as though they are initial. Examples of this are given below; word receiving citation melody are in boldface.

(16) a.

khoeb	ge	**ǂkhanisa**	go	ǁama.
man	DECL	book	TAM	buy

"The man bought the book."

b.

mû	ta	ge	go	[**khoeb**	**ǂkhanisa**	go	**ǁama**	-sa]
see	1sg	DECL	TAM	man	book	TAM	buy	COMP

"I saw that the man bought the book."

The issue of why verbs should differ in form based on the presence of a second-position particle is puzzling; setting this aside, however, it can be seen from this discussion that the citation melody can be robustly used to diagnose the left edges of prosodic phrase in the pre- and middle-fields. This makes tone the primary target of analysis when diagnosing the prosodic structure of leftward-displaced light TAM particles.

3.3 Production experiment

In order to better understand what prosodic factors influence the placement of TAM particles, I designed and conducted a production experiment on the prosody of preverbal particles. The goal of the study is to observe preverbal particles across a range of possible positions in a variety of syntactic structures and to observe what, if any, prosodic features correlate with each position. The primary prosodic feature in question is of course tonal sandhi, but analysis will also look for boundary tones, lengthening, and pauses as potential correlates of prosodic structure.

The study was conducted with six native speakers of Khoekhoe, all literate in the language. Test items were presented in the form of a slide-show, with each slide displaying either a question or a question / answer pair; where there was a pair displayed, the answer was the target sentence, while the question provided a supportive pragmatic context for the syntactic structure being targeted. There were 68 test items in total of 10 syntactic conditions. Each speaker saw each item 3 times[14] in a different random order blocked by repetition (so each item was seen once before any was seen twice). Speakers were recorded with a Zoom H5 recorder and a head-mounted Shure SM-10A microphone, and were offered a short break every 34 items (six breaks evenly spaced through the experiment).

10 different syntactic conditions were tested; within each condition, previous elicitation established which positions were possible TAM positions. Where relevant, two constituents were placed in the middlefield in order to ensure at least three possible hosts for the TAM enclitic. The test items were not evenly distributed across the syntactic conditions: Those conditions which were the primary focus of investigation received 4 sentences, while other conditions received 2 or 1 in order to maintain a reasonable study length. Each sentence was presented in several forms, once with the TAM particle in each possible position.

In order to make it possible to observe the prosodic structure, lexical items were chosen only from those tonal classes which undergo sandhi. In particular, all but two items were chosen from the High (H) and High-Rising (H-HL) classes, which undergo a dramatic register shift to Low (L) and Low-

[13] This applies to verbs in their base-position, clause-finally. When the verb has been fronted to be clause-initial, it receives citation form.

[14] Except one speaker, who ran out of time to complete the 3rd repetition.

Table 4: Syntactic conditions

Condition	Description	Sentences	Positions	Context Question[15]
Verb-final	Matrix declaratives, unmarked word order	4	3	'What happened'?
Subject-final	Predicate-fronted matrix clauses with given subject	4	2	'What did SUBJECT do?'
Inversion	Predicate-fronted matrix clauses with VO word order	4	1	'What did SUBJECT do?'
Embedded	Embedded clauses under 'see' or 'know'	4	3	'What do you (see/know)'?
Questions	Polar questions without clause-type marker	4	3	n/a
Infinitive	Embedded infinitives under 'want'	2	3	'What happened?'
Irrealis	Matrix irrealis (optative interpretation)	1	3	'What should SUBJECT do?'
Conjoined object	Matrix SOV clause with 2 conjoined objects	1	3	'What happened?'
Appositive	Matrix SOV with an appositive modifier on the object	1	3	'What happened?'
Relative	Matrix SOV with an object relative on the object DP	1	3	'What happened?'

Falling (L-SL), respectively; the remaining lexical items were chosen from the Superlow (SL) class, which become Low-Rising (L-SL) under sandhi. This constraint should make it easy to diagnose the prosodic structure — with the exception of the two Superlow items, all and only the words at the left edge of prosodic phrases should have a high tone. Each particular combination of lexical items was used in at most two sentences, to reduce repetition. All sentences were checked for grammaticality with two speakers before the experiment was conducted.

The results of the experiment remain to be analyzed. Analysis will proceed as follows: First, each token utterance will be extracted to an individual sound file. Second, a forced aligner trained on data extracted from earlier elicitation sessions will be used to align word and syllable boundaries. Third, a naive transcriber will sort the lexical items from a subset of the tokens into the categories High, Low, and (Low-)Rising —– the only three distinctions necessary to fully parse the prosodic structure given the choice of lexical items discussed above. These transcriptions will be checked against my own transcriptions, then potentially used to train an automatic classifier. The final transcriptions will be manually inspected for any consistency with regard to the placement of the TAM particle. In addition, measurements of syllable duration and pause placement will be automatically extracted and again inspected for anything that correlates with TAM placement.

4 Theoretical approach

The core theoretical component of this dissertation will be a proposal for a model of prosodic linearization. My aim is for this model to provide some insight into the problem of linearization in general while also concretely modeling the prosodically-driven displacement found in Khoekhoe. In this section I will sketch the direction I intend to take in the construction of this model and will report on some preliminary results.

4.1 On the architecture of the grammar

Minimalist syntax generally subscribes to the Y-model of grammar (Chomsky, 1995), in which the narrow syntax feeds both PF and LF, but not the reverse. Given that this model has found significant support in the literature, it seems desirable to maintain this strict compartmentalization. The model I will propose here is in line with this goal and with the goals of the Minimalist Program more generally insofar as it treats linearization as a strictly PF phenomenon. Linearization of syntactic structures is something only necessary for the pronunciation of those structures, not for further syntactic operations or for interpretation, and hence should be confined to the PF branch exclusively.

If linearization is a PF phenomenon, then it should be analyzed in a fashion parallel to other well-studied phonological phenomena. The field of phonology has largely converged on constraint-based models generally, and on Optimality Theory in specific, as the right kind of architecture for characterizing phonological processes; it seems desirable, then, to give linearization a constraint-based analysis. In addition to analytical uniformity, such an analysis has the benefit of making typological predictions arise out of the interaction of independent constraints. The analysis here, rather than treating linearization as a single constraint *a la* Elfner (2012) will instead derive various orderings from different rankings of a few linearization-related constraints. These constraints will be responsible for deriving the facts of word order typology, and more generally for enforcing consistent mappings between syntactic structures and a set of typologically-possible linear orders.

These constraints will, however, be ranked in the same hierarchy as various prosodic markedness constraints. Prosody, and phonology in general, is a phenomenon which operates on strings — prosodic phenomena such as boundary tone placement or pause insertion pick out the edges of (sub)strings, while more long-distance phenomena such as tone spread seem to operate continuously across a substring; other prosodically conditioned phenomena, such as French liaison, seem to need both precedence and adjacency. And yet, prosodic phenomena also seem to require some reference to syntactic constituency to determine which substrings are relevant. From this standpoint, prosodification is a process relating syntactic structures to a particular set of substrings which are accessible to further phonological processes; this seems closely related to the process of linearization, and given the empirical evidence that phonological aspects of sentences sometimes seem to play a role in their linearization it seems desirable to allow the constraints on linearization and prosody to interact. Once these constraints are ranked together in one hierarchy, the prosodic constraints are capable of picking out winning candidates which violate the ordinary mapping from syntax to string enforced by the linearization constraints.

There is an alternative model to consider, one in which what I have so-far termed the linearization constraints and the prosodic constraints are in fact one in the same — that is, in which the constraints responsible for building prosodic structure are also the constraints responsible for generating linearization. While this seems theoretically more elegant, it is probably a bad idea. While it is true that prosody seems to be able to influence linearization, even to the point of overriding such basic principles such as the contiguity of syntactic constituents, the scope of such influence is actually quite limited. We can see this from two facts. First, word order typology seems to be independent of prosodic typology, with both head-initial and head-final languages displaying a range of possible prosodic structures; if there was only one set of constraints responsible for both linearization and prosodification, we might predict that at least some word-orders would correspond one-to-one with certain prosodic structures. Second, prosodically-driven displacement seems to be very limited in scope: All clearly-attested cases involve only very light items being displaced. While some analyses[16] have proposed that very heavy elements like intonational phrases may sometimes postpose, the fact remains that we do not see, for example, different weights of prosodic words or phrases being linearized differently. Given these facts, it seems desirable that the total constraint set for prosodic linearization should contain of a set of 'faithfulness' constraints enforcing a relation between syntactic order and linear order, without reference to prosody. Separate prosodic constraints — both 'faithfulness' constraints enforcing a relation between syntactic and prosodic constituency and 'markedness' constraints enforcing well-formed prosodic structures — will then interact with these linearization faithfulness constraints to compel prosodically-driven displacement in only those cases where it is observed.

4.2 *On syntactic structures and the desired typology*

The model proposed here aims to capture some facts about the typology of word order, and so it is worth stating from the outset what my assumptions about the empirical typology are. The goal is *not* to capture all possible word orders, or even all possible base word orders for clauses (SVO, VSO, etc.): There is strong support in the literature for the idea that some of these orders (for instance, verb-initial orders) are derived by syntactic movement. Instead, the main goal is to capture the same facts as the Headedness Parameter — namely, that those syntactic positions we call 'specifiers' should uniformly precede their head, while the ordering of head and complement should vary on a per-language basis.

I would like to add another dimension to the typology as well, namely that of movement. Movement configurations in the syntax are a problem for linearization in that some syntactic item seems to be associated with multiple positions and yet appears in the string only once. Many contemporary analyses of movement assume something like a copy theory, in which 'copies' of the moved item are merge into multiple locations in the syntactic structure; this model then requires some operation to ensure that

[16] For example, Edmiston & Potsdam (2016) proposes that embedded clauses in Malagasy postpose for prosodic reasons.

only one copy of the moved item is pronounced. But this 'copy deletion' operation looks like a 'post-syntactic' operation in that it only effects the linearization; as such, it seems desirable to model this process in a constraint-based framework, rather than as an operation on its own. That is, the constraints on linearization should be formulated such that no candidate with two copies of the moved item should win under any ranking of constraints. Furthermore, there are some arguments that moved items may, in some languages, appear in their lower positions — consider, for example, WH-*in situ* languages, where we certainly want the WH item to have some syntactic relation to the WH-complementizer without it appearing in spec,CP. This, then, is a second dimension for the typology: Whether moved items are linearized 'high' (in their landing site) or 'low' (in their base position).

For a concrete example, consider the tree in (16); this is intended to be schematic of a clausal structure, with X, Y, and Z representing the clausal spine while SP and OP represent a subject and an object, respectively. OP, here, has moved into the specifier of XP. If all of the syntactic heads in this structure are taken to represent independent words, then we have 7 words and 5040 possible orders those words could be placed in. However, our desired typology includes only the four orders in table 5; these orders are exactly those which show uniform headedness, place specifiers on the left, and linearize all elements of OP together in one of its two positions.

(17)

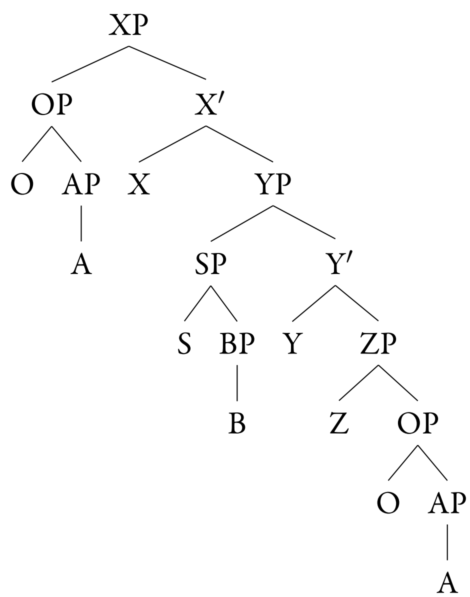


Table 5: Desired typology

	HEAD INITIAL	HEAD FINAL
LOW	XSBYZOA	SBOAZYX
HIGH	OAXSBYZ	OASBZYX

It's worth mentioning at least one alternative to copy theory, namely multidominance. In a multidominant structure, nodes in the syntactic tree can have multiple mothers, and moved items are taken to be Merged into multiple locations (rather than copied). There are at least two ways to build multidominant structures that are discussed in the literature. Engdahl (1980); Gärtner (2002) and others propose what I will term Remerge multidominance, in which moved items are simply targeted by the Merge operation twice — once to Merge them with their base-position sister, and then later in the derivation to Merge them with their landing-site sister. By contrast, Johnson (2017) argues for a Parallel Merge multidominance, in which moved items are targeted by the Merge operation only once, but that operation is a more-than-2-place function which merges the targeted item with all of its sisters at once. In Parallel Merge multidominance, a moved item is not placed directly into a higher specifier; instead, it becomes sister to a head which isn't in the main derivational spine, which itself is then later Merged into a specifier position.

While Remerge multidominance and copy theory have potentially different implications for the narrow syntax and the LF branch, from the point of view of PF it is not clear that they have any substantive difference. A set of constraints which viewed the two copies of a moved item as indistinguishable, for instance, should always score the same violations on a copy-theory tree and the equivalent Remerge tree. By contrast, Parallel Merge multidominance introduces at least one extra head, which potentially significantly changes the geometric relationships between the moved head and the heads it has moved past. My goal in designing this linearization system is that it should yield the desired typological results for copy-theory and Remerge trees; it will also be interesting to consider what modifications would need to be made to the theory to adequately linearize Parallel Merge multidominant trees.

There is one more aspect to the typological predictions that must be considered, and that is the typology of prosodically-driven displacement. The empirical facts here are significantly less clear. The danger of overpredicting the prevalence of prosodic displacement is clear — most or all displacement involves very light (sub-word) items, so the relevant constraints should in some way more strongly restrict the linearization of full words. But the exact nature of the linearization faithfulness constraints will also determine what displacements are available as 'repairs' for sub-optimal prosodic structures, and the typology of these displacements is significantly understudied. It is exactly this part of the typology that I hope Khoekhoe may shed some light on; my goal will be to design a constraint set that accounts well for Khoekhoe and then to consider what predictions are made by that system.

4.3 *On the necessity of computational modelling*

In Optimality Theory, the core assumption is that all possible candidate outputs compete in parallel. In the case of prosody, this means that all possible prosodic structures are in competition. This is a very large candidate set indeed — even ignoring intonational-phrase level structure, with 5 words in the terminal string the number of prosodic structures is already 1,440; at 7 words, it is over 155 thousand. The problem gets worse when we add in linearization: If Gen must generate both all permutations of the words and all prosodifications of each of those permutations, then at 5 words we must already consider nearly 173 thousand candidates. Obviously, most of those candidates are harmonically bounded (meaning that they cannot win under any ranking of constraints), but determining which can win is still a non-trivial task. Evaluating the typological predictions of a set of prosodic linearization constraints, then, is well beyond what can be done by hand.

In order to solve this problem, I have designed and built a computational model of prosodic linearization capable of generating all candidates for a given syntactic structure and evaluating them across a range of constraints; this system can relatively quickly determine the predicted typology of a particular constraint set.[17] This mo I will be using this throughout my research to test the success of my models. This code, and a thorough description of its architecture, will thus also form a core part of my dissertation.

[17] My model is not the first computational method for doing prosodic typology — SPOT (J. Bellik, 2015) calculates prosodic typologies across a range of prosodic constraint sets using JavaScript. To the best of my knowledge, however, my model is the first to incorporate linearization.

4.4 *A preliminary proposal*

In this section, I will present a preliminary proposal for the set of linearization faithfulness constraints. These constraints are capable of correctly linearizing the tree in (17), generating all and only the linear orders in table 5. These constraints purely enforce linearization, with no relation to prosodic structure; integrating them into the broader set of prosodic constraints remains to be done. As noted above, these constraints work equally well for Remerge multidominance and copy-theory; they don't currently work for the equivalent Parallel Merge structures.

The constraint set I propose here relies on defining a few structural relations in the syntax; some of these are familiar syntactic concepts, while some are new to this proposal. First, I define c-command in the familiar way, but with special care paid to making sure that moved items c-command everything that they have moved past:

(18) X **c-commands** Y if some sister of X recursively dominates Y.

In order to evaluate my constraints, it will be useful to define $A(\alpha)$, the set of **ancestors** of α. Intuitively, $A(\alpha)$ is the set of all syntactic nodes which (recursively) dominate α.

(19) The **ancestors** of a node α, $A(\alpha) := \lambda X.$ X recursively dominates α.

My constraints will also rely on the notion of **proper ancestors**, $pA(\alpha)$. Intuitively, the proper ancestors of a node α are those nodes which dominate α in every position in which it appears. If α has not been moved, then $pA(\alpha) = A(\alpha)$; if α has been moved, then instead $pA(\alpha) \subset A(\alpha)$. A formal definition is as follows:

(20) The **proper ancestors** of a node α, $pA(\alpha) := \lambda X.$ X recursively dominates α and α does not c-command X.

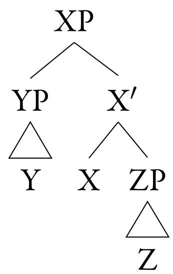
With these definitions in hand, I propose three linearization faithfulness constraints — the minimum number of constraints necessary to derive the four possible outputs in table 5. The constraints are as follows:

(21)
a. SubsetLeft: For each pair of nodes X and Y, where Y is a terminal node and $A(X) \subset A(Y)$: Score one violation if Y precedes some word dominated by X.
b. ProperSubsetLeft: For each pair of nodes X and Y, where Y is a terminal node and $pA(X) \subset pA(Y)$: Score one violation if Y precedes some word dominated by X.
c. HeadsRight: Score one violation for each pair of words X and Y where some projection of X dominates Y and X precedes Y.

Intuitively, SubsetLeft and ProperSubsetLeft both enforce head-initiality; in fact, in practice both enforce something like Kayne's LCA . The fundamental intuition of the LCA seems to be that nodes closer to the root of the tree (i.e. which were Merged into the tree later in the derivation) should precede those which are further from the root of the tree. In the case of the LCA 'closeness to the root' is measured via asymmetric c-command. In the constraint set proposed here, I instead measure this property directly, via a subset relationship between ancestors — nodes which are closer to the root of the tree definitionally have fewer ancestors than those which are further away. The distinction between *ancestors* (in the case of SubsetLeft) and *proper ancestors* (in the case of ProperSubsetLeft) tracks the location of moved items: ProperSubsetLeft will count only those nodes which are ancestors of the moved item's highest position, thus compelling the moved item to precede everything it has moved past, while SubsetLeft will count all the ancestors from the lowest position.

It's worth noting at this point that both SubsetLeft and ProperSubsetLeft, insofar as they mimic the action of the LCA , also compel specifiers to be on the left and complements to be on the right. Consider the tree in (22), for instance; I've enumerated the (proper) ancestors of all nodes and the relevant subset relations. Because YP has only the root of the tree in its ancestry, it is compelled to precede both X and Z; similarly, the presence of ZP in the ancestry of Z compels it to follow X.

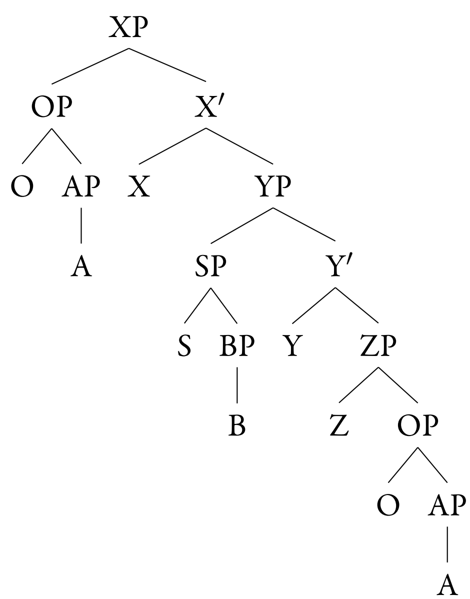
(22)



a. $A(X) = pA(X) = \{XP, X'\}$
b. $A(YP) = pA(YP) = \{XP\}$
c. $A(Y) = pA(Y) = \{XP, YP\}$
d. $A(ZP) = pA(ZP) = \{XP, X'\}$
e. $A(Z) = pA(Z) = \{XP, X', ZP\}$
f. **Subset relations**:
   (i) $A(YP) \subset A(X)$
   (ii) $A(YP) \subset A(Z)$
   (iii) $A(X) \subset A(Z)$

If both of these constraints are head-initial, then it just remains to add a constraint forcing head-finality. This is exactly the job of HeadsRight, which quite simply compels all heads to follow everything dominated by their projections. In the factorial typology, whenever SubsetLeft or ProperSubsetLeft are undominated, a head-initial order will obtain; when HeadsRight is undominated, the relative ranking of ProperSubsetLeft and SubsetLeft will determine whether moved items are placed 'low' (following any filled specifiers they have moved over) or 'high' (preceding everything they have moved over); crucially, in either case the moved item precedes the head that selected it, and so no violations of HeadsRight accrue. This typology is illustrated by the tableau below, using the tree in (23) (repeated from (17)) as the input. I have confirmed by exhaustive computational search that there are no other winners — only these for linearizations are possible given this constraint set.

(23)



4.5 Remaining issues

This constraint set thus derives the desired typology of headedness and movement. There are remaining issues to be resolved, however. None of these constraints make any reference to prosody; worse, adding

Table 6: Word order typology

Input:	(23)	SubsetLeft	ProperSubsetLeft	HeadsRight
Initial, Low:	XSBYZOA	0	1	13
Initial, High:	XSBYZOA	4	0	7
Final, Low:	SBOAZYX	5	5	0
Final, High:	OASBZYX	6	4	0

Match, a commonly-assumed framework for the construction of prosodic structure, disrupts the desired typology. However, this fundamentally results from the fact that, when a moved item is linearized in its 'high' position, Match can never be fully satisfied. To provide a concrete example, in the tree in (23), ZP can never be fully matched unless OP is linearized in its 'low' position, adjacent to Z; if OP is high, then the elements of ZP do not form a contiguous string and so cannot be matched by a prosodic phrase. This is not the desired effect: We would like our prosodic structure constraints to provide some pressure to map ZP to a prosodic phrase even if something has moved out of it. Thus, Match itself is inadequate to the task of prosodifying structures involving movement unless copy-deletion or a similar process applies before the prosody. Thus, one of the immediate next goals of this project is to design an alternative version of Match Theory which will enable the prosodification of movement structures.

Another issue, pending the amendment of Match Theory, is the issue of appropriately constraining prosodically-driven displacement. As noted above, it seems that displaced items are generally subminimal words; put another way, it seems that there is more pressure to put full prosodic words in their 'normal' linear order than there is for small items which do not form prosodic words. It remains to be seen how the linearization faithfulness constraints above can incorporate this intuition; it may be that violations should be calculated over pairs of prosodic words rather than nodes in the syntax.

It is also not immediately clear how to account for mixed-headedness in this system. On the one hand, it would be easy to use targeted constraints that picked out specific categories — for instances, a HeadsRight(D) constraint that would force head-final DPs while remaining neutral on other categories. However, such a system quickly winds up with the same overprediction problem as the Headedness Parameter: Why not have targeted constraints for each individual category? Furthermore, there are typological patterns evident in mixed-headedness, for instance the Final-Over-Final Constraint (Biberauer et al., 2014) which notes that a head-final phrase typically may not select a head-initial one; the solution to incorporating mixed-headedness into this constraint set would optimally address this pattern. The resolution to this and the other open issues will form the core of the theoretical component of this dissertation.

References

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

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.

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

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

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.

Corver, Norbert, & Jairo Nunes. 2007. *The copy theory of movement*, volume 107. John Benjamins Publishing.

Dobashi, Yoshihito. 2004. Multiple spell-out, label-free syntax, and PF-interface. *Explorations in English Linguistics* 19:1–47.

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

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

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

Engdahl, Elisabet. 1980. The syntax and semantics of questions in Swedish. Doctoral Dissertation, University of Massachusetts, Amherst, Massachusetts.

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

Gärtner, Hans-Martin. 2002. *Generalized transformations and beyond*. Berlin: Akademie-Verlag.

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.

J. Bellik, N. Kalivoda, O. Bellik. 2015. Syntax-prosody for ot. JavaScript application.

Johnson, Kyle. 2017. Rethinking linearization. Ms.

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

Myrberg, Sara. 2013. Sisterhood in prosodic branching. *Phonology* 30:73.

Prince, Alan, & Paul Smolensky. 1994/2008. *Optimality theory: Constraint interaction in generative grammar*. John Wiley & Sons.

Richards, Norvin. 2016. *Contiguity Theory*, volume 73 of *Linguistic Inquiry Monographs*. MIT Press.

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.

Selkirk, Elisabeth, & Seunghun J Lee. 2015. Constituency in sentence phonology: an introduction. *Phonology* 32:1–18.