Head movement and linear edges

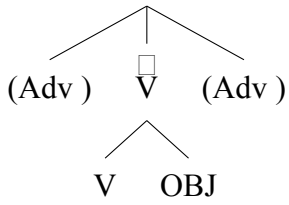
March 18, 2026

1 Introduction: Head Movement

1.1 Motivating Head movement

It is a well-known fact that verbal heads do not necessarily appear in the position where they are introduced in the tree. Given that verbs select objects in their complements, and that adverbs like *often* are located at the VP/vP-edge (1), verbs would be expected to appear in a position strictly adjacent to the objects. That is indeed the case for English lexical verbs, as shown in (2).

(1)



(2) a. Mary often reads books
b. *Mary reads often books
c. Mary reads books often

However, in many other languages, comparable adverbs appear between verbs and objects. This is the case in French, for instance, where the adverb may appear between the verb and object (Joseph Embley Emonds 1970; Pollock 1989). In fact, the adverb cannot appear to the immediate left of the verb:

(3) a. Jean embrasse souvent Marie
J. kisses often M.
b. *Jean souvent embrasse Marie
J. often kisses M.
c. Jean embrasse Marie souvent
J. kisses M. often
”John often kisses Mary”

Hence, the linear position of the verb doesn't reflect constituency.

The general solution to such apparent linear order - constituency mismatches is movement (Internal Merge), where in our cases, the head adjoins to the next higher head (Koopman 1984; Travis 1984; Baker 1985). This type of account is known as *head movement*. In the French examples, V adjoins to a higher head (say T) and, as a result the verb appears to the left of a VP-level adjunct.

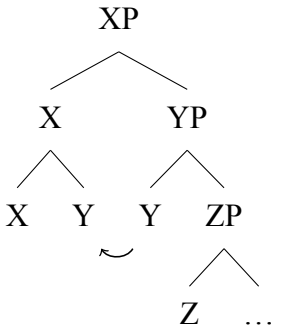
1.2 Challenges for Head Movement accounts

Traditionally, head movement has been perceived as a movement operation applying in narrow syntax, much akin to A- or Ā-movement. At the same time, there are at least four dimensions in which head movement deviates from other types of movement.

1.2.1 Cyclicity

The first one is that head movement violates common conceptions of cyclicity — we dub this the *Cyclicity Problem*. As can be seen in the tree below, head adjunction does not target the root of the tree and in that sense, head movement is not cyclic: it violates the Extension Condition.

(4)



1.2.2 Semantic vacuity

The second dimension is that unlike cases of A- or Ā movement, head movement does not appear to feed LF. Head movement appears semantically vacuous — we term this the *Vacuity Problem*. This can be seen in examples like (5)-(6):

(5) Mary may not eat any cookies

In (5), the modal *may*, an existential quantifier over possible worlds, underwent head movement across negation, but it does not allow a reading where it takes scope above negation. This contrasts with cases of phrasal movement of an existential quantifier over negation, which does lead to surface scope readings:

(6) A student has not passed the exam

The claim that head movement does not feed semantics is not uncontroversial, though. For instance, it has been argued in the literature that head movement can sometimes feed semantics.

Lechner (2004) and Iatridou and Zeijlstra (2013) for instance argue that scopal relations between negation and modals are best captured under the assumption that head movement feeds semantics.

For instance, Iatridou and Zeijlstra (2013) argue that *must* in (7a) moves from a position below negation to one above it. Being a Positive Polarity Item (PPI), it cannot reconstruct below negation and takes scope in situ, unlike polarity-insensitive modals like *may* and *can*, which must reconstruct below negation.

(7) a. Mary mustn't leave
b. Mary can't leave

However, those arguments are not too strong, as already pointed out by Iatridou and Zeijlstra (2013, fn. 22), since it can also be maintained that *must* undergoes postsyntactic QR irrespective of where it appears in its surface position, much in the same vein as *some girl* in *I didn't see some girl* can QR across negation (see Homer 2015). In fact, Jeretič and Thoms (2023) show that an analysis in terms of head movement feeding semantics actually suffers from various problems.

This shows that an argument that is based on head movement of quantifying elements can be subject to an alternative analysis where the relevant element undergoes overt but semantically vacuous head movement and covert QR. As we will show below, such an alternative actually makes better predictions in cases where moving a head to a higher position affects the way elements take scope.

Matyiku (2016) presents data from Western Texas English where subject-auxiliary inversion disambiguates particular readings, for instance in the case of universal subjects in negative clauses:

(8) a. Everybody didn't leave $\neg > \forall; \forall > \neg$
b. Didn't everybody leave $\neg > \forall; *\forall > \neg$

The reason why the inverse scope reading of (8b) is not available, must then also have to do with the fact that the (pragmatic) mechanism that determines which of the readings that can be grammatically derived end up being available in a language, must make reference to surface orders (see Fanselow and Cavar 2001, 2002; Bobaljik and Wurmbrand 2012 a.o.). Without such a mechanism any clausal combination of a quantifier and another scope-taking element in any language, where one moves across the other, should be ambiguous, contrary to fact. That means that even though indirectly, head movement may have semantic effects in that it blocks certain readings, head movement does not directly feed semantics. Hence, head movement cannot be said to feed semantics in these cases in the sense that it plays a role at LF. This also relates to cases where a Negative Polarity Item (NPI) subjects can appear under negated auxiliaries, as in *Which one of them doesn't anybody like?* (McCloskey 1996, see also Roberts 2010). It has been independently known since Ladusaw 1979 that NPI licensing is subject to surface orders (his so-called surface roof constraint), even though the mechanism responsible for the anomaly of unlicensed NPIs applies at LF (see Chierchia 2013 a.m.o.)[1]

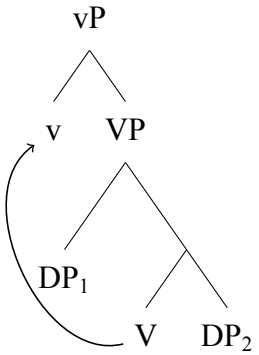
1. Note however, that this is not the only possible explanation for the facts mentioned above. Following Penka (2011), Zeijlstra (2022b) has argued that English *n't* is not the realization of a semantic negative operator but rather a semantically non-negative marker of the presence of a negative operator. There is a semantically negative, covert operator that must be positioned above the surface position of *n't*. If that is the case, it follows trivially that in (8b) the subject must take scope below negation, but in (8a) it doesn't. The licensing of NPI subjects would follow in the same way.

The same kind of reasoning also applies to the cases discussed in Szabolcsi (2011) of aspectual raising verbs like *begin*, modal auxiliaries and potentially also intensional raising verbs where the position of the verb determines whether they are scopally ambiguous with respect to the subject or not. Also, Keine and Bhatt (2016) show in German long passive constructions that when the lower verb undergoes head movement into a matrix verb, arguments and modifiers of this lower verb must also take scope above this matrix verb. Again, here head movement does not (necessarily) feed semantics. Rather, the scopal properties of other elements depends on the surface position of the verb. In this sense, these data do not provide evidence that head movement itself must be semantically active.[2]

1.2.3 Phrasing

Another challenge for head movement, our third dimension, is noted by Kalivoda (2018). He notes that, given a Match Theoretic view of the mapping from syntax to prosody (Selkirk 2011), standard approaches to ditransitive constructions, which are commonly assumed to involve head movement low in the clause, are predicted to have a particular prosodic structure, where both internal arguments should form a joint prosodic phrase — this we name the *Phrasing Problem*.

(9)



As Kalivoda notes, two of the most commonly assumed models of deriving prosodic constituents from syntactic constituents — namely *Align* (Truckenbrodt 1995) or *Match* (Selkirk 2011) — make undesirable predictions about the typology of prosodic phrasing of ditransitives when head movement of the sort schematized above is assumed.[3] *Match* makes a pernicious prediction that the

2. These are not the only claims that head movement feeds semantics. More indirect claims have been made as well, for instance by Gribanova 2017, who argues, following Hartman 2011, that head movement may affect the way constituents can meet the semantic identity requirements that license ellipsis, but also here the effect is indirect and there are other ways to derive the relevant facts without alluding to LF reflexes of head movement. Other indirect relations between head movement and LF effects consist of order preservation effects (like Holmberg's Generalization), where phrasal movement appears dependent on head movement. If such phrasal movement can feed semantics, head movement should apply in Narrow Syntax as well. At the same time, such effects can also be understood in terms of order preservation at PF, where the surface position of a head must remain linearly fixed with respect to the surface position of some phrasal element along the lines of Fox and Pesetsky 2005 (see also Dikken 2007 for more discussion).

3. This is true even of theories that would assign a special status to certain nodes, thereby allowing the VP out of which head movement has taken place to be ignored by the mechanism for translating syntactic structure to a prosodic representation. Kalivoda evaluates one such proposal made by **selkirk2017syntactic<empty citation>** and finds that it overgenerates. As far as we can tell other proposals that allow VP to be ignored in this way, such as Kratzer and Selkirk 2007 likely give rise to the same overgeneration problem.

following prosodic phrasing for ditransitives should be possible and perhaps widely attested, yet, as Kalivoda shows in a comprehensive survey of such patterns, no language is known to exhibit such phrasing. For the tree above, after head movement has taken place, there is a syntactic constituent containing only the two internal arguments, to the exclusion of the verb; *Match* theory would lead us to expect this constituent to map to a prosodic phrase, yet no such phrasing is known to exist.

(10) (V (DP_1 DP_2))

In contrast to the overgeneration problem that befalls *Match* theory, Kalivoda notes that *Align*-based approaches undergenerate in the case of patterns like those found in Ewe ditransitives, schematized below. The problem for such an approach is that there is no syntactic edge between the verb and DP_1 that should map to the sort of prosodic boundary Ewe tone sandhi seems to be sensitive to.

(11) (V) (DP_1) (DP_2)

These problems aside, Kalivoda notes that the most common patterns for prosodic phrasing of ditransitives are those shown below, where the verb either phrases together with the linearly closest argument, or with both internal arguments. He notes that accounting for the commonality of these patterns on an *Align* or *Match* based theory require auxiliary constraints, which have most often been cashed out as involving syntactic relationships such as selection, and suggests that such constraints are undesirable for reasons related to modularity.

(12) ($_\varphi$ v+V DP_1 DP_2)

(13) ($_\varphi$ v+V DP_1) ($_\varphi$ DP_2)

Kalivoda takes these facts to motivate a different approach to the mapping of syntax to prosody, which we take no particular stance on here. What we wish to note is that the three mappings above — those in (12)-(13) as well as the Ewe pattern in (11) — are exactly what we might expect were head movement not to take place at all, supplemented by other assumptions about deriving prosodic constituency from syntactic constituency common to Kalivoda's model as well as *Match* and *Align*-based models. Moreover, the phrasing found in (10) under a Match theoretic approach would be unexpected given a model where head movement does not take place. This, foreshadowing the proposal to be introduced later in this paper, would be possible only if heads could be linearized to the left of their specifiers.

1.2.4 Locality

The fourth dimension in which head movement is different from A- or Ā movement concerns locality. Head movement is strictly local. It is impossible to adjoin a head X to a head Y if some head Z intervenes. This is known as the Head Movement Constraint (HMC), which goes back to Travis (1984). A- or Ā movement, by contrast, does not have to target the closest specifier (albeit A-movement does target the closest A position). We call this the *Locality Problem*.

This difference between head movement on the one hand and A- and A'-movement on the other have led to the development of three schools of thought. One places head movement squarely in the syntax, with the differences between head movement and phrasal movement being attributed to other factors; the other is to delegate parts of head movement to a distinct module of the grammar.

A third approach takes head movement — at least of the classical sort — out of narrow syntax in its entirety. Below we discuss these briefly, and relate them to the proposal we wish to make.

1.3 Approaches to head movement and their problems

Here, we'll briefly review several schools of thought on head movement and how they relate to the problems mentioned beforehand. For a more detailed overview, we refer the reader to Dékány (2018).

One sort of approach locates head movement chiefly in the syntax. This could be done by adopting an approach where Merge is allowed to apply freely (see Safir (2019) for an instance where this logic is made explicit). Another mechanism to resolve the *Cyclicity* problem has been proposed in terms of the so-called reprojection approach to head movement (Koeneman 2000, Fanselow 2004, Suranyi 2008). Under such an approach, a verb does not adjoin to a higher head, but rather adjoins to the phrase it is the head of and projects again. Crucial under this approach is that heads enter the derivation with all the relevant features (e.g., [V], [v], [T]). Every time a different feature then projects, thus yielding a TP dominating a vP, which in turn dominates a VP, albeit that it is not clear why such a head should first project a VP, then a vP, and then a TP. As for all these approaches, while it is possible to define a cyclic principle regulating Merge that will allow for both head movement and phrasal movement, it is not clear that such an approach escapes the other problems for syntactic head movement mentioned beforehand, most notably the *Vacuity Problem* and the *Phrasing Problem*.

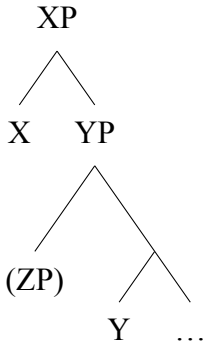
Another way to circumvent the *Cyclicity* problem within narrow syntax would be to assimilate head movement to phrasal movement (see Massam (2000) for such an analysis of a verb-initial language). While we do not dismiss such an approach for languages like Niuean and comparable verb-initial languages, we suspect that a full-blown extension of such an account for all instances of head movement will ultimately be unsatisfactory. Another approach along these lines would be Roberts (2010), who argues that head movement reflects the same underlying Agree mechanism as phrasal movement. For him, when the features of a goal are copied on the probe, these features can also be spelled out on the probe. In that case, he argues, the goal incorporates into the probe. However, this incorporation may still suffer from the *Cyclicity Problem* and the *Locality Problem* as incorporation violates the Extension Condition and probe-goal relations do not have to apply in a strictly local fashion (here, Roberts argues that Extension Condition and the Head movement Constraint do not always apply). Since Agree can feed semantics, it remains also unclear under this proposal how the *Vacuity Problem* is resolved, as well as the *Phrasing Problem* if the goal really incorporates in the probe.

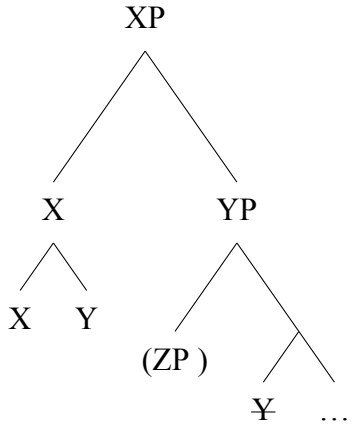
A second sort of approach places portions of head movement in the syntax and other portions in the morphology. The most famous example of this sort of approach comes from Matushansky (2006). Matushansky suggests that head movement takes place in two steps. Heads may undergo movement to specifier positions, with a subsequent morphonological operation — M-merger — being responsible for assimilating the head in specifier position together with another nearby head. While this approach solves the *Cyclicity Problem*, it runs afoul of the three other problems mentioned before. It does not lead us to expect head movement to be semantically vacuous, so it is still subject to the *Vacuity Problem*, it predicts the unattested prosodic phrasings that characterize the *Phrasing Problem*, and the *Locality Problem* is not addressed under this proposal either.

A third sort of approach locates head movement in the postsyntax. Thereby, the differences

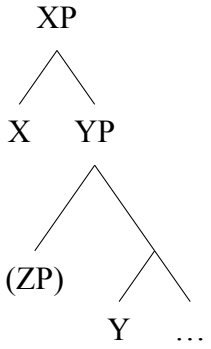
between head movement and non-head movement are captured by attributing the properties of one to a distinct module from the other (see Chomsky 2015, 2021 for proposals along these lines). Outsourcing head movement to PF trivially explains why it no longer feeds LF, resolving the *Vacuity Problem*. It may also avoid the *Cyclicity Problem* as there is no reason to a priori expect operations at PF to be subject to the same cyclic principles as the narrow syntax, although it is not clear why cyclicity should work differently in the PF component of grammar. The *Phrasing Problem* might also be worked around if the prosodic structure is derived from a PF representation, rather than the syntactic tree proper (although we do not know of any claims made in the literature which directly address this juncture between PF-level rebracketing and prosodic phrasing). This third sort of approach thus seems more promising, as it furnishes a straightforward solution to the three problems discussed beforehand. Indeed, we will develop just such an approach. However, we wish here to take issue with many approaches to post-syntactic head movement, namely that the structures which they operate on are similar in nature. Taking Harizanov and Gribanova (2019) as an example, consider the operation they propose for instances of head movement at PF: given two heads, where the first is the head of the phrase which serves as the complement to the second, one of the two might be adjoined to the other.

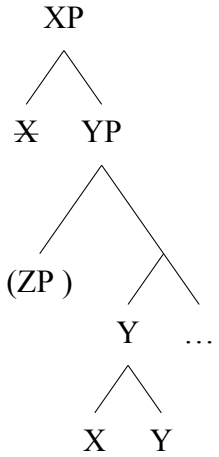
(14)





(15)





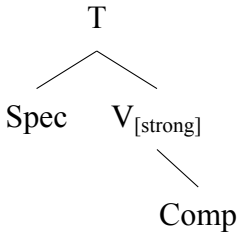
There are good reasons to be suspicious of approaches such as these, however. It is not clear that much is to be won by just taking a set of putatively syntactic operations and declaring them to be part of a distinct module. We are therefore suspicious of modular approaches to grammar where two supposed modules of grammar truck in the same vocabulary — here, minimal and maximal projections, and adjunction to them — and produce very similar outputs — here, novel constituents where the properties of one of the two daughters of the novel constituent are inherited.

An other approach along these lines — which also places head movement out of the purview of the narrow syntax — pioneered by Brody (2000) involves a revision of the syntactic apparatus. Doing so allows for a linearization algorithm which may produce head movement effects as a byproduct of linearization. Crucial assumptions for this type of approach are enumerated below.

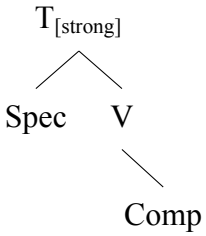
(16) a. A node may represent both a head and its phrasal projection.
b. Phrasal specifiers precede the head they are a specifier of, phrasal complements follow heads they are a complement of; head complements precede heads they are a complement of, head specifiers follow heads they are a specifier of.
c. Language specific parameters may designate heads in an iterated complement relationship as morphological words.
d. Heads may be marked as "weak" or "strong", and this determines where a morphological word is linearized with respect to phrasal elements.

Below we provide Mirror Theoretic representations for an English-like language and a French-like language. Their syntax is identical, with one DP serving as the specifier of T (indicated as a left branch on the tree), and another serving as the complement of V (indicated as a right branch on the tree). Both are assumed to be languages where T and V form a morphological word. The chief difference between the two is whether T or V is marked as strong. In an English-like language (17), the exponents of T and V will precede the complement of V, but not the specifiers of T and V, while in a French-like language, (18), the exponents of T and V will precede the complement of T (namely, both the specifier and complement of V).

(17)



(18)



Mirror Theoretic approaches fare well with regard to three of the problems discussed before: Cyclicity falls out straightforwardly from the architecture, as the approach does not assume a structural alternation for head movement, and so need not constrain it with something like the Head Movement Constraint. Likewise, as head movement is part of the algorithm for translating structure to a string, Vacuity is expected. Locality is likewise baked into the system in (16).

Such approaches may run afoul of the *Phrasing Problem*, though, depending on how the mapping from syntax to prosody is formulated for Mirror Theoretic representations. On a straightforward approach where "[strong]" is taken to indicate that such a position is privileged for the PF

branch, either the overgeneration problem inherent to Match-theoretic approaches should arise, or the undergeneration problem inherent to Align-theoretic approaches should arise.[4]

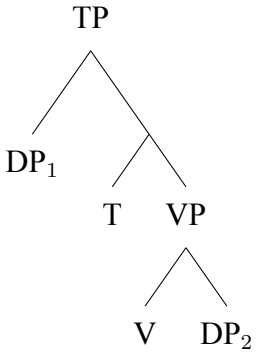
Our aim, then, is to sketch what a post-syntactic approach to head movement might look like without making a direct appeal to operations of the sort discussed above. We will take an altogether different approach, and place much of the responsibility for ensuring that two heads are in a local relationship in the hands of a different PF mechanism without alluding to any kind of movement, either in narrow syntax, or at PF. Instead, we suggest that head movement effects might be treated as the consequence of a particular linearization algorithm that allows a wider range of possible linear orders for a given maximal projection than is commonly assumed.

The theory developed here resembles that proposed by Brody and others in developing a system where some head movement effects result from a theory of linearization. We will see that ours is distinct from theirs in a number of ways. The theory to be developed makes certain predictions not only about where heads may or may not be pronounced with respect to other heads, but also bars phrases from appearing in certain positions in a way that a mirror-theoretic approach does not — allowing us to understand, for instance, differences in the possible interpretations of scrambled sentences in different languages, discussed in §4.

2 Linearization and head movement

Following Kayne (1994) and much subsequent work, it is often assumed that there is a strict one-to-one mapping of syntactic structure to linear order. For instance, if one has a tree like the following, only one order of terminal nodes will be admitted, namely: DP_1 T V DP_2. Under such strict views of linearization, *any* deviation from this linear order must reflect a difference in syntactic structure — head-finality, for instance, must involve movement of constituents within the clause.

(19)



Even if a less strict approach is taken — for instance, by assuming something like a head parameter — the result is the same: any one tree must map to exactly one linear order. Linearization requires an algorithm, either as a strict mapping or by means of a parameter. But if we admit a head parameter into the system that is not uniform across all heads in a language, in essence allowing the features of certain lexical items to determine linear order, we might wonder whether a fully feature-driven view of linearization is possible. We will pursue this line of analysis here. More

4. Moreover, it is unclear whether a Command-theoretic approach like that developed by Kalivoda could be straightforwardly recast to be compatible with Mirror-theoretic syntactic objects.

specifically, we will suggest that the relative order of heads and phrases dominated by the projection of that head are, at least in principle, all freely linearizable. Restrictions on the ordering of elements arise from lexically specified features analogous to the aforementioned features encoding the head parameter, and a highly cyclic approach to the linearization of elements in the clause. Our core focus, naturally, will be to capture certain word order patterns generally attributed to head movement.

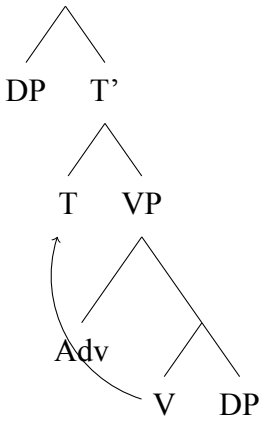
2.1 Linearization

As discussed before, a well-known set of facts commonly attributed to head movement involves differences between languages with respect to possible orderings of the head, complement, and adjunct. For expository purposes, we restrict ourselves to fairly simple syntactic representations free of intermediate projections, but we note that the analysis presented here is, of course, compatible with enriched representations of this sort. As shown in Section 1, and repeated here, certain adverbs may appear between the verb and object in French, but not in English (at least when it concerns lexical verbs).

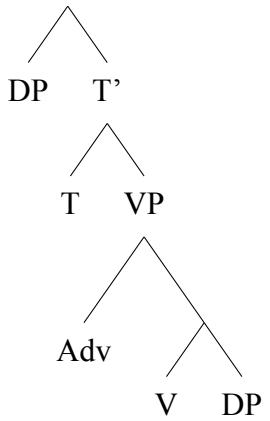
(20) a. Jean embrasse souvent Marie
J. kisses often M.
"John often kisses Mary." French
b. *John kisses often Mary.

A common analysis of facts like these, following Joseph Embley Emonds (1970) and Pollock (1989), involves a parametric difference between English and French cashed out in terms of head movement: in French, V moves to T (21), while in English V does not (22).

(21)



(22)



The strategy here, then, is similar to that commonly assumed for A and Ā-movement. The expected linear order where the adverb precedes the verb is not present in French as a result of the verb undergoing displacement to a head which takes the VP, including the adverb, as its complement. The relevant linear order reflects derived constituency.

But as said before, the question is indeed whether linearization always needs to reflect (derived) constituency. Going back to the head parameter, the idea is that linear order is not fully predictable

on the basis of the hierarchical order of the syntactic objects, but only partially so, and additional linearization constructions come from features on particular heads. If V has some head-initial feature, it must be linearized to the left of its complement (i.e. to the left of the material that makes up the complement); if V instead has a head-final feature, it must be linearized to the right of it.

Head-initial and head-final features are, however, not sufficient to determine the linear order between a head and its specifier(s) and/or adjuncts. Unless additional mechanisms are assumed (like anti-symmetry or other linearization features) the order of a head with respect to these non-complements (both specifiers and adjuncts) is underdetermined. The question is whether that is a good or a bad thing.

But before addressing that question, it should be noted that specifiers and adjuncts behave differently when it comes to (rigid) linearization. Specifiers generally have a specific position in the linear order (often to the left of their heads); here underspecification would be problematic:

(23) a. She likes her
b. *Likes her she

But adjuncts are relatively free and can often appear both to the left and to the right of the verb and its complement, as shown below:

(24) a. She often eats spinach
b. She eats spinach often

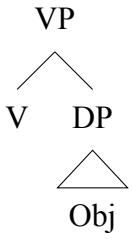
From the perspective of adjuncts, underspecified linearization constraints would actually result in the observed word order freedom.

There are good reasons to assume, though, that the fixed order of specifiers is due to other factors. After all, it is not clear that universally specifiers are always on the left; rightward specifiers have bee proposed in the literature (see, for instance, Aissen 1992; Chung 1998; Tvica 2017; Dobashi and Otaki 2025 for examples and discussion). That suggests that it is not an inherent property of specifiers that they be linearized to the left of their sisters. The question then arises as to why specifiers generally appear on the left. Multiple options are available. For one, specifiers could tend to be leftward for reasons like the EPP. Alternatively, certain affixes need something metrically strong supporting them or a prosodic alignment requirement (Richards 2016). Another reason for the leftness of specifiers in head-initial phrases could be that certain heads are barred from appearing phrase initially (see Sabbagh 2014, who discusses 2nd position clitics of various sorts). And finally, as argued for in Zeijlstra 2022a, if specifier positions are landing sides for movement, and movement is leftward, such specifiers positions should be to the left of their complements as well. Irrespective of the exact reason, we take the general leftness of specifiers to be something that follows on independent grounds and argue that a general theory of linearization can leave the position of non-complements in principle relatively free. In §4, we make an explicit proposal for why derived specifiers tend not to appear on the right, grounded in the feature-based theory of linearization we put forth here.

The question, then, arises how the linear order of heads, complements and non-complements (i.e. specifiers and adjuncts) is determined. To answer that question, we will discuss our assumptions step by step and look first at cases of English lexical verbs, where no head movement effects can be observed. After that, we will look at cases where head movement effects do play a role.

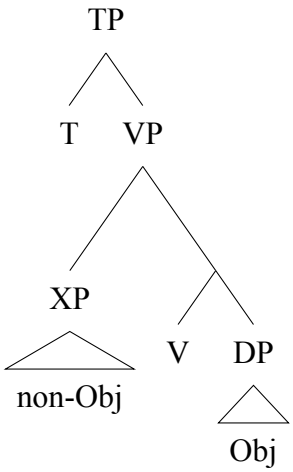
Our first assumption is that linearization takes place whenever a head and complement are merged. The reason is that when a head and its complement merge, it is only then that the head parameter can be properly implemented. As we will see, the fact that linearization is highly cyclic also makes what looks like head movement highly local. Hence, when a VP consisting of a head V and a complement DP is linearized, V is either linearized to the left of (all the material inside) DP or to the right of it. In English, V, being head-initial, ends up to the left of its complement.

(25)



The next step of linearization applies when the next head is merged, which we take to be T:

(26)



It is already clear from the above that V must appear to either the left or the right of its complement. Since English is overall head-initial, V appears to the left. The same applies to T. This way, the following linearization statements are established:

(27) a. T < VP
b. V < DP

As VP contains XP, V and DP, the list of established linearization conditions can be extended as follows:

(28) a. T < XP
b. T < V
c. T < DP
d. V < DP

Note, however, that this expansion of statements is not part of the linearization algorithm proper. The algorithm we propose is relatively sparse, in that it does not produce a total ordering of terminal nodes: some other mechanism is at play in translating the sparse statements in (27) to a total ordering of nodes like that in (28).

For this, we adopt the following rule:

(29) If an element A precedes/follows an element B, everything inside A also precedes/follows everything inside B.

As a consequence of this algorithm, the order T < V < DP is already established; the only question is where XP needs to be pronounced. In this light, it is important that the relative position of XP does not have to be restricted to one position. As we saw, in the case of adjuncts, it can be either in between T and V, or after the DP. In other words, the adjunct needs to be linearized at the edge of the VP. The word *edge* is deliberately chosen.

But what is the proper definition of "edge"? It may be intuitive to think of the edge of a phrase as the outer layer (either linear or structural) of some phrase. But outer is configurationally defined; a position is an outer position of a phrase if the next element up no longer belongs to that phrase. Let's therefore define a linear edge as follows.

(30) An constituent XP that is part of a phrase BP in a configuration [A [BP]] is at the edge of BP if it is
   a. linearized as closely to A as possible with respect to other linearization statements,
   b. or as far away as possible from A with respect to previously established linearization statements.

This way, XP is either linearized to the immediate right of T, or to the immediate right of the DP that serves as a complement of V. This gives us two sets of possible linearization conditions:

(31) a. T < XP
   b. T < V
   c. T < DP
   d. V < DP
   e. XP < V
   f. XP < DP

or

(32) a. T < XP
   b. T < V
   c. T < DP
   d. V < DP
   e. V < XP
   f. DP < XP

This is indeed in full accordance with the examples shown before. If XP is an adjunct, both orders are fine; if XP is a specifier, the second order must be ruled out independently, for reasons

we have mentioned earlier in this section. This allows us to formulate the following linearization algorithm:[5,6]

(33) **Within a cycle of linearization**
   a. A head is linearized to the left or right of its complement (and everything inside it).
   b. A non-complement inside the complement of a head is linearized at either the left or right edge of that complement.

This derives the English cases with lexical verbs that do not display head movement effects.

Note that the process described in (33) is to a large extent top-down. First, the order of the highest head is determined with respect to its sister. Next, the order of elements within that order are determined. For this procedural reason — namely, that the order of the just-Merged head and its complement are determined first — we have arrived at our definition of edge, given before. The order of the head and its complement are the first thing to be determined in a linearization cycle. Defining "edge" relative to that head, rather than with respect to the phrase it is adjoined to, minimizes the amount of information the linearization mechanism has reference to: edge as defined makes no actual reference to syntactic structure, merely to linearization statements generated in the present and prior cycles of linearization.

In cases where there is more than one specifier/adjunct, our system predicts that the highest one will always be linearized either adjacent to the head (modulo head movement effects, which we discuss later) or as far away as possible from it. This prediction is borne out given the existence of certain mirror effects that Koster 1974 famously observed. Concretely, Koster observed that in Dutch VPs with two specifiers/adjuncts, only two of four possible linear orders are fine when they are both linearized at the same side of the verb, as the following examples show:

(34) a. Ik zou willen dat Jan [[tijdens de pauze]$_1$ [[in de stad]$_2$ rondliep]].
   I would want that John during the break in the city walked.around
   'I wish that John walked around in the city during the break.'
   b. ??dat Jan [[in de stad]$_2$ [[tijdens de pauze]$_1$ rondliep]].
   c. dat Jan [[rondliep [in de stad]$_2$] [tijdens de pauze]$_1$]
   d. ??dat Jan [[rondliep [tijdens de pauze]$_1$] [in de stad]$_2$]

The structural order is the one where the adjunct *tijdens de pauze* ('during the break') is structurally higher than the argument *in de stad* ('of his father'). That means that when the two phrases precede the head, the higher one should precede the lower one. However, when the two phrases follow the head, the higher one should follow the lower one.[7]

5. Note that to the extent that one wishes to think about this in terms of phasality, this means that every head is in some sense a (PF) phase: the moment a head is merged, the complement will be linearized with respect to it. Whether this follows from phasality or from something else, is something we are agnostic about. The same holds for the workings of phasality or its alternatives on the LF side. The only thing that is crucial for us, is that linearization is highly cyclic.

6. There is an open question as to how the highest specifier in a tree is linearized with respect to its sister. As far as we can tell, a similar open question stands for phase-theoretic approaches to the mapping of syntax to the interfaces, discussed in Grohmann et al. (2017). It is likely that the right answer to one question could be taken as an answer to the other. We leave this a topic for future research.

7. Post-verbal adjuncts in English appear to behave differently — their order is apparently free, and the scope of these adjuncts does not seem to track linear order. See Takano (2003) for some discussion of approaches to these facts.

This pattern follows from our analysis. The moment the VP is linearized with respect to the higher head T (which we take to be head-initial, but nothing depends on that), T's sister, the VP, must be linearized. For now, let's ignore cases, where the adjuncts appear on different sides of the verb. Now, the first constituent the linearization algorithm targets is the phrase containing the higher adjunct (*tijdens de pauze*). That adjunct must then either appear to the immediate right of T, or at the very end of the phrase. The next phrase (the argument *aan zijn vader*) then must either follow the higher one, if that one is linearized immediately after T, or appear to the immediate left of the adjunct if that one is linearized in phrase-final position. The argument *aan zijn vader* cannot appear in the position to the immediate right of T or in phrase-final position itself, as then it would contradict the earlier-established linearization statements that require the adjunct to be immediately adjacent to T or as far away as possible inside the VP.[8]

Of course, more needs to be said about the placement of heads in English. For one, lexical verbs are marked for tense and agreement, which we do not attribute to linear adjacency between V and T. Rather, we assume that tense morphology is not hosted on T but rather on V itself, which then Agrees with T (see also Bjorkman (2012) for proposals that Agree between T and V may be responsible for determining the way V is inflected). Evidence for that may also come from examples like:

(35) Wolfgang played tennis on every Sunday. Von Stechow (2006)

The intended interpretation of (35) is one where past tense outscopes the distributive quantifier *every Sunday*, which in turn outscopes the lexical verb *play*, yielding the paraphrase in (36a). The scopal order where past tense would take scope at its surface position, i.e., under *every Sunday*, amounts to the reading in (36b), which is absent (Von Stechow 2003, 2005; Zeijlstra 2012).

(36) a. There exists a past interval t such that for every Sunday in t, Wolfgang plays tennis.
b. *For every Sunday, there exists a time before it such that Wolfgang plays tennis at that time.

This suggests that the semantic locus of tense does not coincide with the surface position of the tense marker. Hence, the tense marker must be realized in a position below T (and arguably on V).

The idea that tense is expressed on V is, of course, by no means new. In fact already the earliest proposals of Affix Hopping (Chomsky 1965) derive that, and in Distributed Morphology it is captured by means of M-merger (Marantz 1988; Halle and Marantz 1993). However, as neither Affix Hopping nor M-merger find a natural place in our proposal, we derive these facts by Agree and do not need to allude to additional (PF) operations for it.

Another thing that needs to be addressed concerns auxiliaries. As is well known, auxiliaries (*be*, *have*, modals) display certain head movement effects. Take, for instance, the following examples, where the auxiliary verb appears to the left of the adverb *often*.

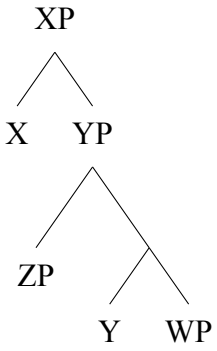
(37) a. Mary has often left
b. Mary may often leave

8. Note that when the adjuncts appear at different sides of the verbs, there is no issue, as both can appear in an edge position in either order.

Have is clearly verbal and therefore must head a VP. Also, a modal like *may* can take scope below *often* and must thus head a projection that is c-commanded by it. Despite this, they appear to the left of the adverb. Their surface position, therefore, does not reflect their base position, which would commonly be attributed to head movement. We might then wonder how such head movement effects can result from our linearization algorithm. To answer this question, we will first focus on a language that manifests head movement effects for all verbs, namely French. After that, we will return to the question raised by English auxiliaries.

First, let's recapitulate the linearization algorithm we have developed, and the conceptual motivations behind it. We've suggested that, first, the relative position of a head and its complement are decided; subsequently, specifiers/adjuncts of the complement are linearized within that complement. This ordering mirrors the relative position of the unlinearized elements in the tree in (38).

(38)



We can think of the linearization algorithm as a strategy for establishing order between two unordered sisters, proceeding from the top down. First, the relative order of X and YP are established. Next, the relative order of XP and the other material in the YP is established.

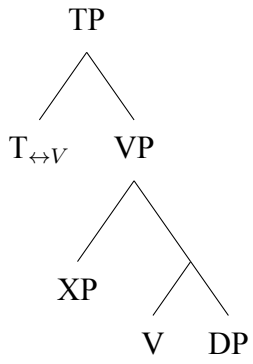
2.2 Adjacency diacritics

Now, let's look at French. As shown before, in French, the finite verb needs to precede adverbs that are located at the VP-boundary, unlike what is the case with English lexical verbs. An adjunct thus follows the verb but may still precede or follow the object.

(39) a. Jean embrasse souvent Marie
J. kisses often M.
b. Jean embrasse Marie souvent
J. kisses M often.
"John often kisses Mary."

Following classical analyses of head movement, we assume that in French T is subject to a condition that it is linearly adjacent to V. However, unlike classical analyses of head movement, we do not assume that this triggers head adjunction to T. We assume that in French T is equipped with a feature, which we notate as $[\leftrightarrow V]$, which merely informs the linearization mechanism to linearize V and T in a strictly adjacent fashion (see Ryan 2010). It ensures that the tense marker and the verb end up being adjacent so that the tense affix is not stranded. With this in mind, let us now look at how linearization will take place exactly. Take the tree in (40).

(40)



First, after merger of V and DP, the following linearization construction is derived following the algorithm introduced earlier in this section, based on the fact that French is head-initial.

(41) a. V < DP

Then, upon merger with T, an additional instruction is generated. Next to the instruction that says that T should precede VP, there is an additional instruction, T«V, which states that T should be strictly adjacent to V.[9] We thus derive the following rules:

(42) a. T < XP
b. T « V
c. T < DP
d. V < DP

The final question is then how XP must be linearized. What we know is that such elements have to be spelled out in the edge of T's complement, i.e. XP must be linearized as closely to T as possible with respect to other linearization conditions, or as far away as possible from T with respect to other linearization conditions. Since, V must be strictly adjacent to T, the closest possible linear position for XP to T is then to the immediate right of V. The position furthest away is immediately after DP. Thus, the following statements hold:

(43) a. T < XP
b. T « V
c. T < DP
d. V < DP
e. V < XP
f. XP < DP

or

9. One caveat is in order there. The results so far may suggest that T should be prefixal on V instead of suffixal (given that T precedes V in (40b)), contrary to fact. However, we assume, that the internal order of V and T is something that is determined by the morphology/vocabulary insertion, and that vocabulary insertion applies after linearization, following Embick 2010 and references therein.

(44) a. T < XP
b. T « V
c. T < DP
d. V < DP
e. V < XP
f. DP < XP

Note that by the sole innovation of a diacritic like $_{\leftrightarrow}V$ the linearization component can linearize the verb in a position that does not *appear* to reflect its base position. But crucially, no head movement has taken place, neither in syntax nor at PF. In French, T is uniformly equipped with $_{\leftrightarrow}V$. One could even say that $_{\leftrightarrow}V$ is connected to the feature [T]. What this derives is that T and V are linearly adjacent.

But this does not always have to be the case. It may very well be that only T with some particular feature has a property like $_{\leftrightarrow}V$. This is what we assume is the case for English auxiliaries. We assume that T[Aux] has a diacritic $_{\leftrightarrow}V$, so that whenever T merges with a VP (or any other phrase) with a head containing feature [Aux], this head ends up strictly adjacent to T. As we see below, both examples have a T head carrying [Aux]. Since T[Aux] has a diacritic $_{\leftrightarrow}V$, this requires that the highest verb be linearly adjacent to T, and therefore gets linearized to the left of *often*. Note, also, that in case of multiple auxiliaries, only the highest needs to be dislocated. In fact, only the highest may be dislocated: in (45b), for instance, the linear order of *been*, and the phrase it heads, will have already been determined with respect of a higher head which takes that phrase as a complement, prior to the introduction of T into the structure.

(45) a. Mary has often left
b. Mary has often been working

Note that the presence or absence of auxiliaries must then correspond with the presence or absence of [Aux] on T. If [Aux] is present, there must be an auxiliary in the clause, which then ends up adjacent to T. If [Aux] is absent, no auxiliary may be present. One way to derive this is by appealing to selection — we could propose, for instance, that this would be blocked by the selectional properties of T, with T[Aux] being incompatible with lexically headed VPs, and auxiliaries requiring to be selected by T[Aux]. Alternatively it could be assumed that every auxiliary must stand in some Agree relation with T[Aux] and vice versa. Then it follows that if no auxiliary is present, T[Aux] cannot survive, nor can an auxiliary survive without T[Aux] being present. Hence, only (highest) auxiliaries can and must appear to the left of the rest of their VP.

Also note that this derives do-support. Since the negative marker *n't* can only be attached to auxiliaries, the sentence must contain an auxiliary. One can implement this by assuming that *n't* also is compatible with the selectional requirements of $T_{[Aux]}$ or that the negative marker carries a feature that needs to agree with $T_{[Aux]}$.[10] In addition the negative head must have a diacritic $_{\leftrightarrow}V$. Then it follows that a sentence with sentential negation must contain an auxiliary verb, either *be*, *have*, a modal or *do*. In a sentence like *Mary doesn't leave*, *n't*, the complex unit V-*n't* (where the

10. And in the case where the phrasal negative marker *not* is used, the negative head still carries such an agreement feature.

negative marker suffixes on the verb), then further satisfies the adjacency requirement encoded by the $\leftrightarrow V$ diacritic on T as well.[11]

Naturally, the question arises why certain heads have adjacency features and others not. We assume that ultimately such diacritics reflect what underlies the Stray Affix Filter (Lasnik 1981). To avoid that some affix would be left without a head, an adjacency critic ensures that its host ends up at the right place. And similarly, if some head is realized by a free morpheme there is no need for an adjacency diacritic. One may wonder whether this does not give rise to a look-ahead effect, given that at the stage of the derivation where linearization takes place, Vocabulary Insertion has not taken place yet. However, whether a particular head (or a head with some particular feature) has an adjacency diacritic is a matter of learnability. Children, based on the language input, determine which head carries which feature. If a bound tense/agreement morpheme, as in French, always is attached to the verb, that is a cue for the first language learner to postulate a diacritic $\leftrightarrow V$ on T/Agr. In a language like English, there is no such cue present. However, a child would not be able to acquire a language like French, with T being realized by bound morphology without $\leftrightarrow V$, as that would reflect in an impossible grammar. The language input for the child can only be compatible with possible grammars. The fact that some form is morpho-phonologically bound forms a cue for the language learner that a particular head has some adjacency diacritic.

To sum up, what underlies our proposal are two mechanisms for determining the linear order of elements. One is universal yet subject to parametric variation: it involves the incremental establishment of linearization statements between two unlinearized elements. When one of the unlinearized elements is a terminal node, the setting of a parameter determines whether or not it precedes or follows its non-terminal sister. When neither of the unlinearized elements is a terminal node, the choice of the relative order of the sisters is free. The other is language specific: certain elements are specified to always appear in an immediately adjacent position to other specified elements. The theory we present, then, is generally quite restrictive: possible mechanisms for determining linear order make reference to one structural notion (sisterhood) and one linear notion (immediate adjacency), and nothing more.

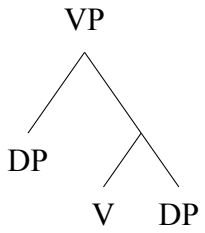
2.3 Solutions to previous problems

So far, we have established an account of head displacement effects that crucially involves no such thing as head movement. Heads do not undergo movement and stay at their base position when they are linearized. This mechanism naturally resolves the four problems that traditional accounts of head movement were facing: the *Cyclicity Problem*, the *Vacuity Problem*, the *Phrasing Problem*, and *Locality Problem*. As for the *Cyclicity Problem*, since there is no head movement going on whatsoever, this problem is trivially resolved. As for the *Vacuity Problem*, since linearization is a PF phenomenon, no semantic feeding is expected in the first place.

As for the *Phrasing Problem*, since heads do not undergo movement, the syntactic structure for a ditransitive should give rise to the prosodic structures discussed by Kalivoda (2018). Consider the following representation, alongside the assumption that V is to be linearized at the left edge of the VP, as a consequence of a relationship with some higher head.

11. Strictly speaking, one could argue that in this case $\leftrightarrow V$ on T is not fully satisfied as *n't* still intervenes. However, since *n't* ends to the right of the auxiliary, this requirement is ultimately fulfilled upon Vocabulary Insertion. Alternatively, one could also argue that higher heads in the extended projection, such as the head of NegP, also carry a feature [V], and therefore count as verbal as well.

(46)



Suppose we adopt an approach to the syntax-prosody interface like that proposed by (Selkirk 2011).

(47) **Match-XP**
For each XP, there is a phonological phrase ϕ dominating the exponents of the terminals XP dominates.

On such an approach, we are led to expect the verb and both of its arguments to phrase together, as shown below. The VP maps to a φ, as do both of the arguments of the verb.

(48) (V DP_{φ} DP_{φ})$_{\varphi}$

Crucially, we do not expect the phrasing shown below, which we would on a head-raising approach where V has moved to some higher position. There is no XP which dominates the two arguments to the exclusion of the verb, so such a phrasing will always violate Match-XP.

(49) (V (DP_{φ} DP_{φ})$_{\varphi}$)$_{\varphi}$

The question that remains, then, is why the DPs themselves are not encapsulated in their own prosodic phrases. We can understand this as the result of a family of Binarity requirements on prosodic phrasing, which prefer prosodic phrases which consist of exactly two prosodic words (see Itô and Mester (1992) and Selkirk (2000) for more discussion of this point). Kalivoda's survey focuses on ditransitive constructions where the arguments of the ditransitive verb each consist of one prosodic word — while (50a) most closely satisfies Match-XP, it is in violation of Binarity, as the φ corresponding to each DP consist of a single phonological word each. (50b-c) come closer to respecting Binarity.

(50) a. (V DP_{φ} DP_{φ})$_{\varphi}$
b. (V DP DP)$_{\varphi}$
c. (V DP)$_{\varphi}$ (DP)$_{\varphi}$

It is useful here to spell out what the family of Binarity constraints might look like, following Selkirk (2011).

(51) a. BinMin
A φ should dominate no fewer than two other elements.
b. BinMax
A φ should dominate no more than two other elements.

We can see that both (50b-c), the most common patterns found by Kalivoda, as different strategies for satisfying the Binarity family above. (50b) violates BinMax while satisfying Match-XP and BinMin. (50c) violates BinMin while satisfying Match-XP and BinMax. The common patterns from Kalivoda's survey fall out as an interaction between Match-XP and the Binarity family. Crucially, the expected interactions only arise given an input structure where raising of V has not taken place.

Finally, as for the *Locality Problem*, the Head Movement Constraint follows as linearization is strictly local. To see this, assume that T has a diacritic $\leftrightarrow V$, but that v lacks any such diacritic. Then, specifiers and adjuncts of V can intervene in between V and v, and there is no way in which V could be guaranteed to end up adjacent to T. In fact, in a harmonic system, v will always appear in between T and v, hence T-V adjacency will never be available either. The only case where v would not appear in between T and V (and where T and V could end up adjacent to each other) are cases where T and V are head-final and v is head-initial, or the other way round. Both configurations are, however, independently ruled out, as they violate the Final-Over-Final Condition (FOFC, Biberauer, Holmberg, and Roberts (2014)). Hence, there is no way for T to end up adjacent to v, unless v has a $\leftrightarrow V$ and T a diacritic $\leftrightarrow v$. That derives the exact locality footprint of what looks like head movement.[12]

2.4 So far

So far, we have established an account of head displacement effects that crucially involves no such thing as head movement. Heads do not undergo movement and stay at their base position when they are linearized. As shown above, this mechanism naturally resolves the four problems that traditional accounts of head movement were facing: the *Cyclicity Problem*, the *Vacuity Problem*, the *Phrasing Problem*, and *Locality Problem*.

We do so by making a number of assumptions, most notably that heads can be equipped with a diacritic that states how it is to be linearized with respect to its complement as well as whether it needs to be adjacent to some other head. These assumptions, as well as the ancillary assumptions we make, all have been independently proposed and/or motivated in the literature before. As for the former, anybody who does not assume a specific linearization algorithm such as the LCA, must

12. We might, of course, wonder about the possibility of apparent non-local head movement fed by phrasal movement. Suppose we had a language (similar to German, to be discussed in more detail in the following section) where C is head-initial but lower projections are head-final. Suppose also that VP raises to spec,CP and C has an adjacency requirement with V. In such cases, we might expect to find non-local head movement, where fronting of the VP allows V to satisfy an adjacency requirement with C without necessarily satisfying this requirement with lower functional projection (such as v, and presumably also T).

(52) [[O V] $C_{\leftrightarrow V}$ …[[O V] v]]

For such a language, we might expect at least some functional material to be stranded in a clause final position, while whatever inflectional information we attribute to the relevant head — C in the example above — should appear on the verb. A potential example of this comes from Urarina, an OVS language isolate of Peru described at length in Olawsky (2011). While the verb is reasonably morphologically complex in Urarina, negation has a special status in the language: it is expressed using a clause-final clitic Olawsky (2011, p. 834). This could be understood as a consequence of "long" head movement allowed by the derivation in (52). We hope that future research reveals clearer examples of this sort.

assume head complement linearization diacritics, and the adjacency diacritic is essentially an implementation of the Stray Affix filter. Hence, our overall theoretical apparatus is not in any way richer, if not poorer, than in alternative approaches to head movement.

At the same time, there are also many questions left open. One of them concerns so-called "mixed headedness" patterns. These, we will discuss in the next section.

3 Mixed headedness

So far we have looked at harmonically-headed languages, where every head is either head-initial or head-final. In this section we discuss a potential challenge for the theory developed here — that of "mixed headedness". Those are languages where one head that is head-initial dominates another head that is head final. Take, for instance, German (53). German is a language with head-final VP and TP but with head-initial CP. In addition, in main clauses, the finite verb/auxiliary 'raises' to C. This way, the highest verb (either lexical or an auxiliary) shows up in a clause-initial position, preceding the object, but the rest of the verbs cluster in the hierarchically determined order in clause-final position (see Den Besten 1989 and many others).

(53) a. Ich sehe Marie.
I see Marie
"I see Mary."

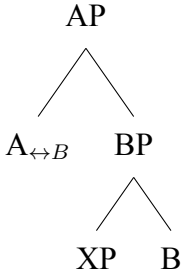
b. Ich habe Marie gesehen.
I have Marie seen
"I have seen Marie."

c. Ich werde Marie gesehen haben.
I will Marie seen have
"I will have seen Marie."

d. dass Ich Marie gesehen haben werde.
that I Marie seen have will
"that I will have seen Marie."

Apart from cases like (53d), the theory of linearization developed so far does not straightforwardly account for such cases where a head-final verb appears to 'raise' to C. The reason is that in such cases, the linearisation mechanism yields conflicting statements. To illustrate this, let's look at the tree below, including only a higher head-initial AP dominating a lower head-final BP, where B should be string-adjacent to A:

(54)



Here, B, being head-final, should be linearised to the right of its complement XP. At the same time, it should be strictly adjacent to A, and A should be linearised to the left of BP and thus of XP:

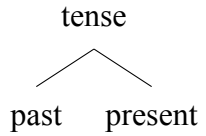
(55) a. A < B
b. A < XP
c. XP < B
d. A « B

This is a contradictory set of linearization statements: (55d) precludes the possibility of the statements (55b) and (55c) in conjunction. In general, the system we have developed does not straightforwardly account for cases where one head may appear to the left or right of its complement in one context and on the other side in another. Below, we present a possible solution to this problem.

One approach we could consider involves a non-Kaynean version of an asymmetry between head-initiality and head-finality, inspired by markedness theory. In many cases, markedness effects are not taken to be the result of one of two features that stand in a geometrical sisterhood relation (see Harley 1994), one being more prominent than the other, but rather the result of two features standing in a geometrical mother-daughterhood relation.

To see this, take present vs. past tense. One could argue that a geometry of tense features would involve a superfeature [Tense] with two daughters, [Past] and [Present]. Here, [Past] and [Present] would stand in a symmetric sisterhood relation and the alleged markedness of [Past] should follow from something else. Alternatively, one may assume that the geometry actually involves a mother feature [Tense] and a daughter [Past]. [Tense] is unspecified with respect to anteriority, [Past] is not. The fact that [Tense] generally denotes a present tense would then be a pragmatic blocking effect: anteriority must be expressed by using the more specific [Past] feature, hence only if the speaker does not intend to make an anteriority claim will she use underspecified [Tense]. Such an asymmetry between past and present has been proposed independently by many (including Sauerland 2002; Altshuler and Schwarzschild 2013 and many others). If [Past] and [Present] indeed stand in asymmetric feature-geometric relation, it follows naturally that the more specific feature is also the more marked one.

(56) **"Differentiated present" approach**

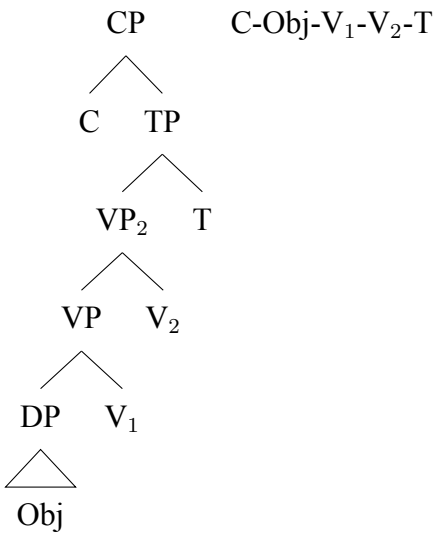


(57) **"Underspecified present" approach**

The same reasoning could apply to opposite featural diacritics like < and >. They could be real duals of each other, with opposite head-directionality 'values', but they could also be taken to stand in an asymmetric relation. We could assume instead that there are two diacritics. One is < (giving rise to head-initiality) and one is •. • is underspecified. Since < is more specific than •, we take • to have a default interpretation of > unless this violates other linearization statements (just like [tense] receives a non-past interpretation as long as that is compatible with the rest of the meaning; otherwise it could be interpreted as denoting anteriority (see Sauerland 2002)). Hence, H < ComP

means H should precede its complement, H • Comp means that H follows Comp unless this violates other linearization statements.[13,14]

Now take German C-Obj-V-Aux, where Aux is to be treated as the morphological realization of an adjacent V and T.

(58) 

Now assume C is head-initial and that V and T are head final. This gives rise to the following linearisation statements:

(59) a. C < TP
b. T • VP_2
c. V_2 • VP_1
d. V_1 • Obj

This means that T and V may strictly speaking appear either to the left or right of their complement, but unless anything contradicts this, TP and VP will be linearised in a head-final way.

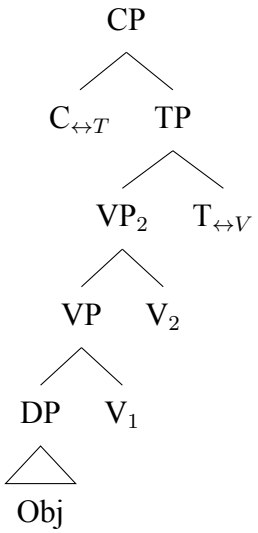
Now, let's also assume the following linearisation conditions $T_{\leftrightarrow V}$ and $C_{\leftrightarrow T}$.

In that case, linearization of the following tree, where T should be adjacent to V ($T_{\leftrightarrow V}$) and C to T ($C_{\leftrightarrow T}$) would block interpretation of • as being a head-final diacritic.

13. See **fukuiSaitoOrderPhraseStructure1998**; Fukui and Saito 1992; Joseph E. Emonds 2009; Haider 2013 for earlier proposals that take head-finality to be the default head-complement linearization

14. A reviewer notes that this might motivate an Optimality Theoretic approach, rather than the particular algorithm for linearization we have specified here. We are relatively confident that an OT grammar could be written that would produce similar results to the algorithm we have utilized, and think that exploration of this idea may prove fruitful. However, we shy away from doing so in this paper — one expectation that we might have on an Optimality Theoretic approach is that learners should be able to rerank the constraints, giving rise to a multiplicity of potential linearization mechanisms. The expectation then would be that these mechanisms have empirical bite — that languages can be found where this clearly is the case. We are less confident that such a claim is correct.

(60)



To see this, let's go through all the relevant steps.

First step: V_1 • Obj, will result in V ending up to the right of Obj, unless overruled later on:

(61) a. V_1 • Obj

Next, V_2 • VP_1 will result in V_2 ending up the right of VP_1, unless overruled later on.

(62) a. V_1 • Obj
b. V_2 • V_1
c. V_2 • Obj

In the third step, $T_{\leftrightarrow V}$ • VP_2, should in principle still result in T appearing to the right of VP_1 (unless overruled later on). The diacritic, will result in a statement that V_2 and T be adjacent. Hence, we have (63), where •• indicates (unspecified) strict adjacency:

(63) a. V_1 • Obj
b. V_2 • V_1
c. V_2 • Obj
d. T •• V_2
e. T • V_1
f. T • Obj

Here, nothing would block an interpretation of • as >, and the following order would appear: Obj < V_1 < V_2 < T. Note that this is indeed the attested order in German embedded (non-V2) clauses (53d).

But now, let's see what happens when matrix head-initial C with the adjacency diacritic for T, $\leftrightarrow$T, gets merged:

(64) a. V_1 • Obj
b. V_2 • V_1
c. V_2 • Obj
d. T •• V_2
e. T • V_1
f. T • Obj
g. C « T
h. C < V_2
i. C < V_1
j. C < Obj

As the reader can see, at this stage, in order not to trigger any contradictions, C must be linearized to the left of TP (and everything in it), while C must be strictly left-adjacent to T and T and V_2 must be strictly adjacent too. This will now force • to be interpreted as (non-default) < in the following set of statements:

(65) a. V_1 • Obj
b. V_2 < V_1
c. V_2 < Obj
d. T « V_2
e. T < V_1
f. T < Obj
g. C « T
h. C < V_2
i. C < V_1
j. C < Obj

The only unspecified order is that between Obj and V_1. Hence, only V_1 will be linearized to the right of the object. Hence, we get the following linearization: C « T « V_2 < Obj < V_1 , which is the attested order. Thus, this way German V2 orders can be derived as well.[15]

Note furthermore that the mirror image of these V2 orders (i.e. V-penultimate orders) cannot be derived given the asymmetry between head-finality and head-initiality. As long as it is hard-wired that head-initial features are more marked than head-final ones and markedness implies specificity relations, only leftward head 'movement' is possible. Of course, this begs the question as to why a mirror system would not be possible (i.e. > and •, where • would encode head-initiality when possible). The idea would be that in such a system rightward head movement effects but not leftward head movement effects should be attested. Whatever the exact nature is of the ban on rightward head movement effects (either in terms of filler-gap mechanisms or as part of inherent restrictions on grammar) is reflected in the availability of < and • only.

15. The question arises when the ordering instruction represented by • is interpreted as being head-final. This should happen at least by the end of the entire clause/extended projection, as it may be determined only at C-level that • must be interpreted as head-initial or not. Whether this is static (involving some kind of phasality) or more dynamic (applying at the moment when it is clear that a head-final interpretation can no longer trigger any contradictory linearization statements later on) is a question we leave open for now. However, in §4.3 we develop an approach to certain types of head movement which are more compatible with the static approach.

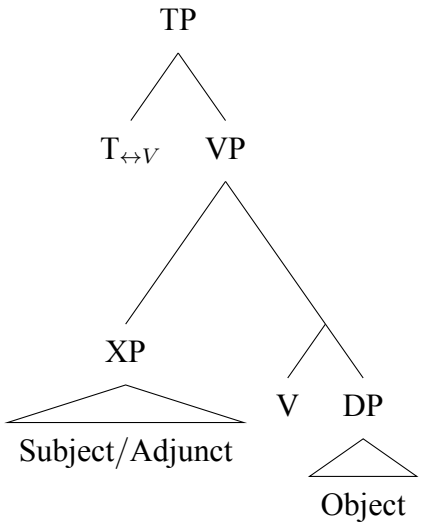
4 Whether or not a language has scrambling

4.1 Scrambling and the linearization system

In this part of the paper we introduce another feature of heads, which requires them to be non-initial, in that they must host a (leftward) specifier of some sort. We suggest that this sort of feature underlies at least some classical EPP effects. That is to say: the need for an element to be non-initial might consistently be satisfied by syntactic movement of a phrase to a specifier position. In doing so, we highlight a potentially interesting property of the system developed here: elements that undergo what looks like head movement may not themselves be allowed to have an EPP property. Building on work by Branan (2021), we show that this derives a difference between Japanese and Tagalog with respect to the semantic effect of scrambling. In both languages — analogous to the cases discussed involving adjuncts in the prior section — complements of the verb and specifiers may be linearized freely with respect to each other, giving rise to semantically vacuous scrambling.[16]

In the previous section, we assumed a simplified clause structure like that given below for French-like languages. There, we saw that complement and adjunct or specifier could be freely linearized with respect to each other, while V was required to precede both of them.

(66)



It turns out this model for French works very well for languages like Tagalog.[17] More specifically, a structure like that above is one where we should expect the verb to appear first, followed by the arguments it selects. Much like adverbs and objects in French, an agent in spec,VP and internal argument as comp,V will be freely ordered with respect to each other. As shown below, this is indeed the case.[18]

16. This is analogous, in some respects, to proposals made by Sauerland and Elbourne (2002) for scrambling in Japanese. There, they suggest that scrambling in Japanese may take place either in the narrow syntax or at PF. In the former case, reconstruction does not take place, while in the latter case the constituent is obligatorily interpreted in its base position. Considered in this light, the theory developed here builds a model of what PF scrambling might look like.

17. We set aside the question of voice alternations in Tagalog for now, and return to them at the end of this section.

18. We gloss *ang* here as a nominative marker, but note that everything said here is compatible with an analysis of ergative-absolutive analysis of Tagalog like that put forth in Aldridge (2004). We gloss *ng* as a genitive marker,

(67) a. L<um>unon ang ina ng mani.
av.swallowed nom mother gen peanut
'The mother swallowed a peanut'.
b. L<um>unon ng mani ang ina.
av.swallowed gen peanut nom mother
'The mother swallowed a peanut'.

We should further expect that the relative order of arguments has no effect on scope or variable binding. As demonstrated below, this is indeed the case.[19] Regardless of the order of a pivot agent and non-pivot theme, for instance, the theme may never take scope over the agent, and thus may not bind a variable in it.

(68) a. *Nag-mamahal ng bawat anak ang kanyang ama
av-love gen every child nom poss. father
"Her$_i$ father loves every$_i$ child."
b. *Nag-mamahal ang kanyang ama ng bawat anak
av-love nom poss. father gen every child
"Her$_i$ father loves every$_i$ child." Rackowski (2002)

This absence of semantic effect is just what we should expect if the free order of arguments in Tagalog is a result of the linearization algorithm developed here, rather than a result of syntactic movement (see, e.g. (Richards 1993) for an analysis of Tagalog scrambling as involving Ā-movement which undergoes obligatory reconstruction). However, given the fact that there are languages — such as Japanese, which we will discuss in further detail below — which do allow scrambling that has a semantic effect, we might now wonder why it would be the case that no learner of Tagalog takes the free order of constituents as a sort of syntactic scrambling. It turns out that the system we have developed has a ready answer for this question when amended to allow us to understand movement dependencies.

To illustrate how this works, we turn to Japanese, and will subsequently return to Tagalog. First, consider the two simple sentences like the following. We assume that Japanese is a head-final language — in the structures below T and V are linearized to the right of their complements. Note, given the logic we applied to the case of adjuncts in French, that the relative order of arguments is free. The subject may be linearized either to the left of V, but to the right of V's complement, thereby placing the subject as close as possible to the higher T, modulo the requirement that T and V be strictly adjacent, or to the left of both V and the complement of V. In the latter case, an SOV order arises, while in the former case an OSV order arises.

following standard convention where the *ng* marker which appears on many non-pivot arguments is analyzed as being identical to a comparable *ng* which appears on arguments of nominals.

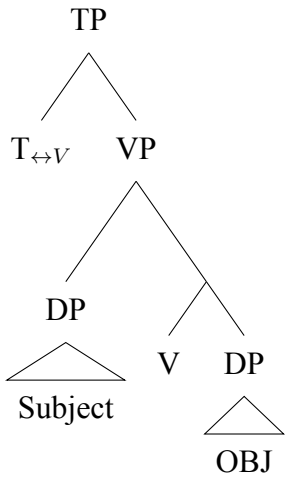
19. See also Branan (2021) for discussion of nuclear stress in Tagalog that is in accord with what has been discussed here.

(69) a. John-ga ringo-o tabe-ta
J.-nom apple-acc eat-pst
"John ate an apple."

b. Ringo-o John-ga tabe-ta
apple-acc J.-nom eat-pst
"John ate an apple."

The following structure is the base structure for Japanese — a clause where nothing has undergone movement.

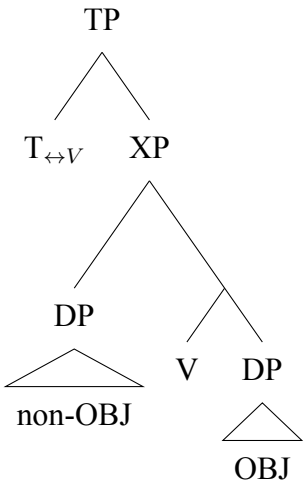
(70)



OBJ < V
V < T
OBJ < T
SUBJ < V
SUBJ < T
SUBJ < OBJ **or** OBJ < SUBJ

Note that the underlying structure for the two sentences is the same, even in the case where the object precedes the subject. Given that Japanese has scrambling, this is somewhat desirable. For Tagalog, the facts are very much the same, modulo the setting of the headedness parameter. As sketched below, V will be linearized before the object, and the subject may be linearized between the verb and object — placing it as close to T as possible, modulo the requirement that nothing appear between V and T — or after both the verb and object — placing it as far away from T as possible. The effect here is much the same as in Japanese: the presence of an adjacency requirement between T and V results in an in-situ subject being linearized freely with respect to the object.

(71)



V < OBJ
T < V
T < OBJ
T < non-OBJ
V < non-OBJ
not-OBJ < OBJ **or** OBJ < Not-OBJ

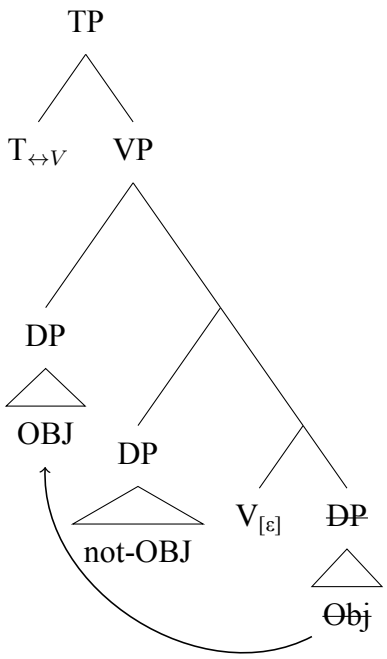
What of the cases where scrambling *does* have a semantic effect in Japanese? At present, our system is unable to account for these cases, and thus requires an amendment. As we have stressed throughout, our approach to linearization is *feature driven* — the basic insight is that the individual properties of the categories lexical items belong to are by and large the factors which determine the relative order of specifiers, heads, and their complements within a given maximal projection. Here, we would like to introduce another feature of lexical items, which we call [ε] (where [ε] is reminiscent of (if not identical to) EPP, in that it is a feature which drives movement).[20] For the system developed here, the order of specifiers with respect to other elements in their phrase is in principle free. We suggest that the [ε] feature demands that a particular specifier of a phrase appear in the highest position. Features of this sort, we suggest, can also be triggers for movement of elements c-commanded by the head bearing the relevant feature: whichever element undergoes movement driven by [ε] must be leftmost in the relevant phrase.[21]

For the cases of scrambling in Japanese which do have a semantic effect, we assume a structure like the following. Here the object has undergone movement to a specifier position, driven by the needs of [ε]. Consequently, the object must be linearized at the left edge of VP, allowing only the OSV word order. Here, the object precedes the subject, feeding an interpretation where the object takes scope over the subject for the purposes of evaluating scope and variable binding.

20. Essentially, we adopt a feature-driven approach to scrambling. See Grewendorf and Sabel (1999), and Erlewine and Lim (2023) for some recent discussion of locality conditions on scrambling on such an approach

21. Alternatively, one could assume that there is a higher head present that triggers movement of the relevant phrase into its specifier. This assumption is orthogonal to our proposal.

(72)

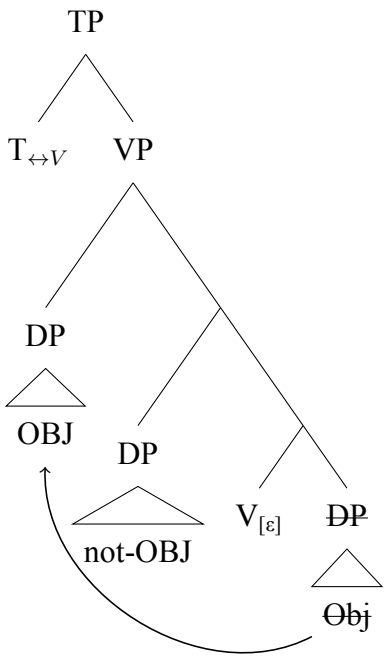


V < T
not-OBJ < V
not-OBJ < T
OBJ < T
OBJ < V
OBJ < not-OBJ

We return now to Tagalog, and highlight an interesting outcome of the analysis proposed for Japanese scrambling which affects the semantics: Tagalog cannot have scrambling which leads to an interpretive difference. Consider the following, where heads are linearized before their complements, T and V have an adjacency requirement, and V has an [ε] feature. The result is a conflict at PF: the [ε] feature on V here requires the object to be linearized in an initial position, before all other elements in the phrase. However, if this takes place, the requirement that T and V be strictly adjacent cannot be satisfied, as the scrambled object precedes V, as a consequence of being moved by [ε].[22]

22. As noted by a reviewer, a straightforward expectation of this is that, in ditransitive clauses, the structurally lowest argument should not be able to precede or follow the highest. This is not the case — these orders are licit in Tagalog. However, the lowest argument, in Tagalog ditransitives, should be understood as a prepositional phrase, following Rackowski (2002) and V. Chen (2025). Aldridge (2003) notes that the distribution of prepositional phrases in Tagalog is surprisingly free - PPs construed with the predicate of a prenominal relative clause may appear after the head of the relative clause. As part of her analysis of relative clauses, Aldridge proposes that PPs in Tagalog may undergo movement to a relatively high position - at least as high as the highest argument in the clause. If we adopt this aspect of her analysis as generally the case in Tagalog, we have a route to understanding why the surprising orders are licit. In brief, the idea would be that movement to this position is not driven by the [ε] feature, but is in a position to disrupt adjacency between two elements in the extended verbal projection, such as T and V. Consequently, it could be linearized either before or after all other arguments in the clause.

(73)



T < V
T < OBJ
T < not-OBJ
V < not-OBJ
OBJ < not-OBJ
OBJ < V

The system presented here should allow Japanese-like scrambling only in languages that are head-final, or head-initial without adjacency requirements between V and T like those found in French. In languages that do have such an adjacency requirement, scrambling will be impossible, as the particular linear order imposed by [ε] is incompatible with the linear order imposed by [↔V] on T.[23]

4.2 Voice alternations in Tagalog

Tagalog is famous for having a robust *voice system*, where one argument receives a particular marker which corresponds with systematic changes in the morphology of its selecting predicate. Much work on Tagalog voice alternations suggests that this privileged element — which we will call the *pivot* — occupies a structurally privileged position above all other arguments in the clause. For instance, for the *patient voice* clause below, there is good evidence that the pivot has undergone movement to a position above the agent: the theme pivot may bind a variable in the non-pivot agent.

23. There is at least some reason to believe this might be on the right track. Otsuka (2005) argues that Tongan (a V-initial language) has scrambling with a semantic effect; Tongan seems to lack an adjacency requirement between V and T, as clitic pronouns may appear in such a position. At least some well-studied head-final languages - such as Korean, Mongolian, and Turkish - behave like Japanese with respect to scrambling. An exhaustive review of this would be well beyond the scope of this paper, but there is reason to believe this proposal may be borne out to some extent.

(74) M<in>amahal ng kanyang ama ang bawat anak
pv.love gen poss father ang every child
"Her father$_i$ loves every child$_i$." Rackowski (2002)

Evidence for movement of this sort comes from facts involving scope and binding, where theme arguments may outscope and bind into agents only when they are the pivot of the clause (Rackowski 2002), as well as a process of quantifier float which may take place only from the pivot of the clause (Kroeger 1993).

This potentially seems at odds with the proposal put forth in this section, depending on the position one assumes non-agent pivots land in. If we adopt a more elaborated clause structure, we might expect movement of this sort to target a position "in the middle of the clause" — which should be ruled out given the proposal put forth here, which leads us to expect movement of this sort to uniformly be banned. Here, we present two possible ways out of this conundrum.

The first would be to deny that the landing site of the pivot is clause-medial in Tagalog. More specifically, we would hope that the pivot lands in a position at or above the highest head in the extended inflectional projection which bears a $\leftrightarrow$ feature. Such proposals have antecedents in the literature — McGinn (1988), Richards (1996), and Sabbagh (2005) all propose that the landing site of the pivot in Tagalog is somewhere high in the left periphery, but differ in their details as to how the pivot appears freely in the post-verbal position. McGinn and Sabbagh both suggest that pivots move to this position, and are then positioned freely after the verb through a post-syntactic operation. Richards suggests that movement to the pivot position in Tagalog is covert movement. Of the three, Richards' proposal is the most straightforwardly compatible with the analysis that we have developed so far. We leave serious consideration of reconciling McGinn's and Sabbagh's proposals with our own a task for future work.

The second would be to deny that the pivot undergoes movement. Such an approach has been posited by Y.-H. Chen (2017). Chen suggests that Agree underlies the relationship between the lexical elements that are realized as voice morphology and the pivot, but that this does not involve displacement of pivots. This approach would be straightforwardly compatible with the theory developed here, with the choice of pivot having no effect at all on the syntactic configuration of arguments with respect to each other. To our mind, the main challenge such an approach faces is not in reconciliation with the theory developed here, but rather in explaining the aforementioned scope and binding facts in a parsimonious way. We leave this too a task for future work.

4.3 T-to-C and [ε]

As part of the analysis of scrambling put forth here, we proposed a particular movement-driving feature, [ε], that deviates from more canonical conceptions of an [EPP] feature. [ε] places two particular restrictions on the head of a phrase: (i) it may not be linearized at the edge of the phrase it heads, rather, the phrase must have a specifier/adjunct, and (ii) that specifier or adjunct must be linearized to the left of that head. A consequence of this is that a head with a strict adjacency requirement can never satisfy this requirement with a head that bears [ε], as the adjacency requirement cannot be satisfied at the same time as the requirement imposed by [ε].

A question then arises: How are we to deal with English sentences like the following? If we attribute the requirement that T host a specifier in English to the presence of [ε], and the presence of the auxiliary before the subject indicates that C has a $_{\leftrightarrow}$T requirement, (75) should run afoul of

the same problem posed for syntactic scrambling in Tagalog: the DP subject and auxiliary should enter an unwinnable competition for being the leftmost element in TP.

(75) Has Mary left?

One move would be to attribute movement to spec,TP in English to a classical [EPP]-feature on T instead of [ε], which does not impose the same linearization requirements as [ε], but only requires that some element is hosted in its specifier. This would unify cases like the example above with earlier cases of alleged head movement effects in French. The subject is in the specifier of TP while the T-head must precede it because of C's adjacency requirement with T.

A potential drawback of this this approach, though, is that it would weaken the proposal made earlier in this section about the unavailability of semantically contentful scrambling in Tagalog-like languages in the sense that it needs to be independently motivated that Tagalog V carries a different movement-triggering feature than English T. This could be motivated, though, by assuming that children are sensitive with respect to whether instances of movement can feed semantics or not. If English children can detect the semantic differences effects triggered by subject movement to Spec,TP this could be a cue to postulate an [EPP] feature on T, instead of [ε]. Tagalog children will, instead, postulate [ε] given that scrambling is semantically vacuous.[24]

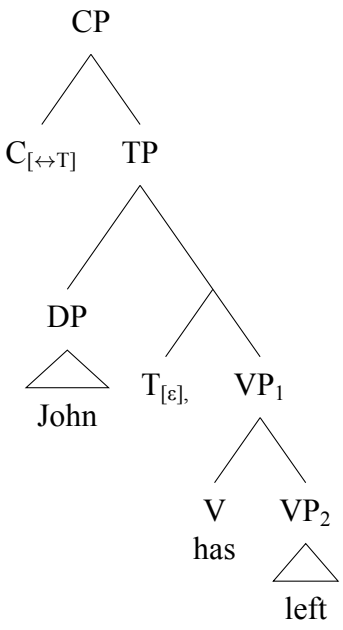
Alternatively, we could suggest that C, at least in English, also acts like intermediate projections of the clause below C in a language like German: it is underspecified in terms of whether it appears to the left or right of its complement; as well as any heads that it satisfies an immediate adjacency requirement with. This, alongside a certain sort of derivational opacity involving the • system proposed in §3, may also allow us to understand why T-to-C of the above sort is allowed.

Previously in this section, we remained uncommited as to how [ε] was evaluated: one option would be to have the requirement that the specifier come first be established as soon a a phrase with an [ε]-bearing head came to have a specifier; another would be to have [ε] play a role in limiting options for the positioning of the specifier when a subsequent round of linearization took place. We suggest that the latter option is more favorable, given the the approach pursued in this paper.

For concreteness, consider the following structure, immediately after C has been merged in the structure. At this point, a cycle of linearization will be triggered. Previously, it would have been established that T precedes VP, as T is head-initial in English. Likewise, given the discussion of auxiliaries in English in §2, it would have also been established that T « *has*.

24. As absence of evidence is not the same as evidence as absence, the fact that scrambling is vacuous in Tagalog does not require that C carries [ε]. It could still be compatible with an [EPP]-feature on V and some other source being responsible for the absence of any semantic effect. Hence, how do we know that the Tagalog child would converge to [ε] on V instead of [V]? One reason could be that the child will always opt for the strongest candidate when it has the choice between two options for the target grammar. Since, [ε] is more restricted than [EPP], both English children would start postulating [ε] and only children would receive conflicting input necessitating them to switch to [EPP].

(76)



a. T « *has*
b. T < VP_1

Now presume that C is underspecified with respect to headedness. Then, it will establish a linear relationship with TP and an immediate adjacency relationship with T, as a consequence of this.

(77) **After linear order of C and other elements is established**
a. T « *has*
b. T < VP_1
c. C • TP
d. C •• T

Next, the position of the DP must be established. The presence of [ε] on T will block the generation of a statement where T < DP and VP < DP — this DP must be leftmost in the TP. We suggest that this results in the generation of the following linearization statements, with the requirement that the specifier be leftmost in TP enforced by [ε] being satisfied.

(78) **After linear order of DP and other elements is established**
a. T « *has*
b. T < VP_1
c. C • TP
d. C •• T
e. DP < VP_2

Why would the statement DP < T not also generate in (78)? The answer would have to do with the presence of the C •• T statement in (78b). At this stage of the linearization process, the position of T is to some extent underdetermined, as its final order will be determined by the way that the • statement is resolved. The idea, then, would be that the leftmost requirement imposed [ε] ignores

elements which stand in a • relationship with other elements, since their linear order is in some sense underdetermined. We must further take this ignorance to be transitive: C stands in a • relationship with T, which stands in a < relationship with VP_1 and a « relationship with *has*; consequently VP_1 and *has* will be irrelevant for the satisfaction of [ε], as their final position is ultimately determined by the resolution of the • relationship into an actual precedence relationship. Since T, VP_1, and *has* all stand, transitively, in a • relationship, they are irrelevant for the evaluation of [ε]. The only element that DP must appear before to satisfy this requirement is VP_2, which does not stand in any relationship of this sort.

The consequence of this is that T may end up preceding DP once the resolution of the • relationship into clear precedence relationships takes place, which takes occurs well after the cycle of linearization triggered by Merge of C. It ends up being the case, given the statements below, that T will precede DP.

(79) **After • statements are resolved**
- a. T « *has*
- b. T < VP_1
- c. C < TP
- d. C « T
- e. DP < VP_2

Why would it be impossible for C and T appear in a head-final position? This is because T does not have an underspecified head parameter like C. When T was Merged and linearized, T was linearized before VP. Linearizing C in a head final position would require T to be immediately adjacent to it, also in a head final position, but T cannot simultaneously appear in this position and precede VP. The only possible resolution of the • statements involving C, then, is one where C precedes TP and immediately precedes T.

What we thus see is that the approach to non-harmonically headed languages developed in the previous section may also allow us to understand T-to-C effects while maintaining the approach to the absence of semantically-contentful scrambling in Tagalog-like languages: the same analysis for English in terms of underspecified directionality diacritics would not be available to the learner of Tagalog, since the extended verbal projection of Tagalog is not normally "split" by an overt, leftward specifier in the way it is in the T-to-C context in English.

In this section, we sketched two possibilities for the analysis of English T-C movement. Evaluating between the two, arguably, depends on more on general architectural considerations rather than strong empirical differences. Hence, we do not further adjudicate between the two types of analysis here.

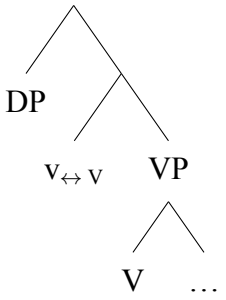
5 Conclusion

In this paper, we have set out to model classical head movement effects as the consequence of a particular linearization mechanism, rather than a structure-altering transformation. On this approach, classical head movement effects involve linearizing a head at the edge of the phrase it heads, with other orders being blocked as a consequence of the PF needs of the next head up. We have demonstrated how this approach works to derive the difference between French-like and English-like

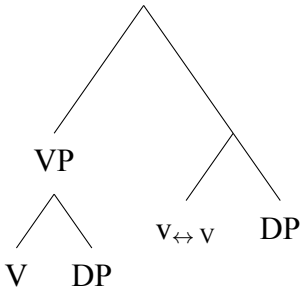
languages. We discussed two potential extensions of the model to account for mixed-headed languages like German. Finally, we saw that the theory leads us to expect scrambling to be possible in some languages but impossible in others.

An open question remains: why do adjacency requirements of this sort exist, if they appear to duplicate the already extant selectional relationship between a selecting head and its complement? One possible reason relates to recent proposals about phrase structure developed by Newman (2021). Newman argues at length that selectional requirements cannot guarantee that a selected element will necessarily be a complement or specifier, and that if that is the case, a variety of facts — such as apparent phrase structure paradoxes — receive a straightforward resolution. If this is on the right track, it is not clear that the duplication between complementation and adjacency is in fact duplication. By way of example, she proposes that a *v* which selects an external argument and internal argument should allow both phrase structures given below.

(80)



(81)



Such structures pose a challenge for classical approaches to head movement (as Newman discusses): in a language that would be argued to have movement of V to v, it is not entirely straightforward how head movement is supposed to take place, given standard assumptions about head movement. If Newman is correct in proposing that VP could be a specifier of vP, then classical head movement of V to v — which is restricted to head-complement relationships — cannot straightforwardly apply. One advantage of the theory developed here is that it is able to straightforwardly ensure that V and v are adjacent: for both structures, v may be linearized in a position adjacent to the V in VP. In (80), this comes about straightforwardly; in (81) this could come about when VP is linearized with respect to v: v would have to be linearized to the left of VP, allowing v and V to be adjacent.

References

Aissen, Judith L. 1992. "Topic and Focus in Mayan." *Language* 68 (1): 43–80. https://doi.org/10.2307/416369.

Aldridge, Edith. 2003. "Remnant Movement in Tagalog Relative Clause Formation." *Linguistic inquiry,* 631–640. JSTOR: 4179253.

———. 2004. "Ergativity and Word Order in Austronesian Languages." PhD diss., Cornell University.

Altshuler, Daniel, and Roger Schwarzschild. 2013. “Moment of Change, Cessation Implicatures and Simultaneous Readings.” In *Proceedings of Sinn Und Bedeutung,* 17:45–62.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. “A Syntactic Universal and Its Consequences.” *Linguistic Inquiry* 45 (2): 169–225.

Bjorkman, Bronwyn M. 2012. “Auxiliary Verb Constructions and the Morphosyntax of Verbal Inflection.” PhD diss., Doctoral dissertation, MIT.

Bobaljik, Jonathan David, and Susi Wurmbrand. 2012. “Word Order and Scope: Transparent Interfaces and the ¾ Signature.” *Linguistic Inquiry* 43 (3): 371–421.

Branan, Kenyon. 2021. “Domains of Ordering.”

Brody, Michael. 2000. “Mirror Theory: Syntactic Representation in Perfect Syntax.” *Linguistic Inquiry* 31 (1): 29–56.

Chen, Victoria. 2025. “The Syntax of Philippine-type Alignment: Insights from Case-Marking.” *Natural Language & Linguistic Theory* 43, no. 3 (August): 1839–1898. issn: 0167-806X, 1573-0859.

Chen, Yen-Hsin. 2017. “A Reexamination of the Philippine-type Voice System and Its Implications for Austronesian Primary-Level Subgrouping.” PhD diss., University of Hawai’i at Mānoa.

Chierchia, Gennaro. 2013. *Logic in Grammar: Polarity, Free Choice, and Intervention.* Oxford University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* MIT Press.

———. 2015. “Problems of Projection: Extensions.” In *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti,* edited by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3–16. Amsterdam: John Benjamins. https://doi.org/10.1075/la.223.01cho.

———. 2021. “Minimalism: Where Are We Now, and Where Can We Hope to Go.” *Gengo Kenkyu* 160:1–41.

Chung, Sandra. 1998. *The Design of Agreement: Evidence from Chamorro.* Chicago: University of Chicago Press.

Dékány, Éva Katalin. 2018. “Approaches to Head Movement: A critical assessment.” *Glossa: a journal of general linguistics* 3 (1): 1–43.

Den Besten, Hans. 1989. *Studies in West Germanic Syntax.* Amsterdam: Rodopi.

Dikken, Marcel den. 2007. “Phase extension: Contours of a theory of the role of head movement in phrasal extraction.” *Theoretical Linguistics* 33.

Dobashi, Yoshihito, and Koichi Otaki. 2025. “An Elsewhere Condition Approach to Kaqchikel VOS.” In *Phonological Externalization.*

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology.* 218. Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.

Emonds, Joseph E. 2009. “Universal Default Right Headedness and How Stress Determines Word Order.” *Lingue e linguaggio* 8 (1): 5–24. https://doi.org/10.1418/29336.

Emonds, Joseph Embley. 1970. “Root and Structure-Preserving Transformations.” PhD diss., Massachusetts Institute of Technology.

Erlewine, Michael Yoshitaka, and Joey Lim. 2023. “Deriving Scrambling with Word Order Freezing.” *Glossa: a journal of general linguistics* 8 (1).

Fanselow, Gisbert, and Damir Cavar. 2001. “Remarks on the economy of pronunciation.” In *Competition in Syntax,* edited by Gereon Müller and Wolfgang Sternefeld, 107–150. Berlin / New York: De Gruyter / Mouton. isbn: 3-11-016945-2.

———. 2002. “Partial deletion.” In *Theoretical Approaches to Universals,* edited by Artemis Alexiadou, 65–97. Amsterdam / Philadelphia: John Benjamins.

Fox, Danny, and David Pesetsky. 2005. “Cyclic Linearization of Syntactic Structure.” *Theoretical Linguistics* 31:1–45.

Fukui, Naoki, and Mamoru Saito. 1992. “Order in Phrase Structure and Movement.” In *Current Issues in Comparative Grammar,* edited by Robert Freidin, 217–242. Dordrecht: Kluwer.

Grewendorf, Günther, and Joachim Sabel. 1999. “Scrambling in German and Japanese: Adjunction versus Multiple Specifiers.” *Natural Language & Linguistic Theory* 17 (1): 1–65.

Gribanova, Vera. 2017. “Head movement and ellipsis in the expression of Russian polarity focus.” *Natural Language and Linguistic Theory* 35 (4): 1079–1121.

Grohmann, Kleanthes, Markus Pöchtrager, Tobias Scheer, Michael Schiffman, and Neven Wenger. 2017. “The Apex Paradox.” *Snippets* 31.

Haider, Hubert. 2013. *Symmetry Breaking in Syntax.* Cambridge: Cambridge University Press.

Halle, Morris, and Alec Marantz. 1993. “Distributed Morphology and the Pieces of Inflection.” In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger,* edited by Kenneth Hale and Samuel J. Keyser, 111–176. Cambridge, MA: MIT Press.

Harizanov, Boris, and Vera Gribanova. 2019. “Whither Head Movement?” *Natural Language & Linguistic Theory* 37, no. 2 (May): 461–522. issn: 1573-0859.

Harley, Heidi. 1994. “Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy.” *Massachusetts Institute of Technology Working Papers in Linguistics* 21:289–320.

Hartman, Jeremy. 2011. “The semantic uniformity of traces: Evidence from ellipsis parallelism.” *Linguistic Inquiry* 42 (3): 367–388.

Homer, V. 2015. “Neg-raising and positive polarity: The view from modals.” *Semantics and Pragmatics* 8:1–88.

Iatridou, Sabine, and Hedde Zeijlstra. 2013. “Negation, Polarity, and Deontic Modals.” *Linguistic inquiry* 44 (4): 529–568.

Itô, Junko, and Armin Mester. 1992. "Weak Layering and Word Binarity." *Ms., University of California, Santa Cruz.*

Jeretič, Paloma, and Gary Thoms. 2023. "Modals, Negation and Movement: A Reassessment." *Glossa: a journal of general linguistics* 8 (1).

Kalivoda, Nicholas. 2018. "Syntax-Prosody Mismatches in Optimality Theory." PhD diss., UC Santa Cruz.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax.* LI Monographs 25. Cambridge, Massachussetts: MIT Press.

Keine, Stefan, and Rajesh Bhatt. 2016. "Interpreting verb clusters." *Natural Language and Linguistic Theory* 34:1445–1492.

Koster, Jan. 1974. "Het werkwoord als spiegelcentrum." *Spektator* 3:601–618.

Kratzer, Angelika, and Elisabeth Selkirk. 2007. "Phase Theory and Prosodic Spellout: The Case of Verbs." *The Linguistic Review* 24 (2-3): 93–135.

Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog.* Stanford, California: Center for the Study of Language (CSLI).

Ladusaw, William A. 1979. "Polarity Sensitivity as Inherent Scope Relations." Doctoral dissertation. PhD diss., University of Texas at Austin.

Lasnik, Howard. 1981. "Restricting the Theory of Transformations." In *Explanation in Linguistics,* edited by N. Hornstein and D. Lightfoot, 152–173.

Lechner, Winfried. 2004. "An Interpretive Effect of Head Movement." In *Phases of Interpretation,* edited by Maria Frascarelli, 45–69. Berlin/New York: Mouton de Gruyter.

Marantz, Alec. 1988. "Clitics, morphological merger, and the mapping to phonological structure." In *Theoretical Morphology,* 253–270.

Massam, Diane. 2000. "VSO versus VOS: Aspects of Niuean Word Order." In *The Syntax of Verb Initial Languages,* edited by Andrew Carnie and Eithne Guilfoyle, 97–116. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.

Matushansky, Ora. 2006. "Head Movement in Linguistic Theory." *Linguistic inquiry* 37 (1): 69–109.

Matyiku, Sabina. 2016. "Semantic effects of head movement in negative auxiliary inversion constructions." In *Proceedings of NELS 46,* edited by Christopher Hammerly and Brandon Prickett, 2:355–364.

McCloskey, James. 1996. "On the scope of verb movement in Irish." *Natural Language and Linguistic Theory* 14:47–104.

McGinn, Richard. 1988. "Government and Case in Tagalog." Edited by Richard McGinn. *Studies in Austronesian linguistics,* 275–293.

Newman, Elise Sophia Bershad. 2021. “The (in) Distinction between Wh-Movement and c-Selection.” PhD diss., Massachusetts Institute of Technology.

Olawsky, Knut J. 2011. *A Grammar of Urarina.* Vol. 37. Walter de Gruyter. Accessed January 12, 2026.

Otsuka, Yuko. 2005. “Scrambling and Information Focus: VSO-VOS Alternation in Tongan.” *The free word order phenomenon: Its syntactic sources and diversity,* 243–280.

Penka, Doris. 2011. *Negative Indefinites.* Oxford University Press.

Pollock, Jean-Yves. 1989. “Verb Movement, Universal Grammar, and the Structure of IP.” *Linguistic inquiry* 20 (3): 365–424.

Rackowski, Andrea. 2002. “The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments.” PhD diss., Massachusetts Institute of Technology.

Richards, Norvin. 1993. “Tagalog and the Typology of Scrambling.” *Unpublished honors thesis. Ithaca, NY: Cornell University.*

———. 1996. “Subjects in Tagalog and Icelandic.” In *Voice in Austronesian,* edited by Marian Klamer, 39:31–49. NUSA Jakarta, Indonesia.

———. 2016. *Contiguity Theory.* Cambridge, Massachussetts: MIT Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals.* Linguistic Inquiry Monographs 59. MIT Press.

Ryan, Kevin M. 2010. “Variable Affix Order: Grammar and Learning.” *Language* 86 (4): 758–791.

Sabbagh, Joseph. 2005. “Non-Verbal Argument Structure: Evidence from Tagalog.” PhD diss., Massachusetts Institute of Technology.

———. 2014. “Word Order and Prosodic-Structure Constraints in Tagalog.” *Syntax* 17 (1): 40–89.

Safir, Ken. 2019. “The A/Ā Distinction as an Epiphenomenon.” *Linguistic Inquiry* 50 (2): 285–336.

Sauerland, Uli. 2002. “The Present Tense Is Vacuous.” *Snippets* 6 (11): 12–13.

Sauerland, Uli, and Paul Elbourne. 2002. “Total Reconstruction, PF Movement, and Derivational Order.” *Linguistic Inquiry* 33 (2): 283–319.

Selkirk, Elisabeth. 2000. “The Interaction of Constraints on Prosodic Phrasing.” In *Prosody: Theory and Experiment,* 231–261. Springer.

———. 2011. “The Syntax-Phonology Interface.” *The handbook of phonological theory,* 435–483.

Von Stechow, Arnim. 2003. “Feature Deletion under Semantic Binding: Tense, Person, and Mood under Verbal Quantifiers.” *Proceedings of NELS 33,* 379–404.

———. 2005. “Semantisches Und Morphologisches Tempus Zur Temporalen Orientierung von Einstellungen Und Modalen.” *Neue Beiträge zur Germanistik* 124:9–54.

———. 2006. “Types of iF/uF Agreement.” University of Tübingen manuscript.

Szabolcsi, Anna. 2011. “Certain Verbs Are Syntactically Explicit Quantifiers.” *Baltic International Yearbook of Cognition, Logic and Communication* 6.

Takano, Yuji. 2003. “How Antisymmetric Is Syntax?” *Linguistic Inquiry* 34 (3): 516–526.

Truckenbrodt, Hubert. 1995. “Phonological Phrases: Their Relation to Syntax, Focus, and Prominence.” Doctoral dissertation. PhD diss., Massachusetts Institute of Technology.

Tvica, Seid. 2017. “Agreement and Verb Movement: The Rich Agreement Hypothesis from a Typological Perspective.” PhD diss., University of Amsterdam.

Zeijlstra, Hedde. 2012. “There Is Only One Way to Agree.” *The linguistic review* 29 (3): 491–539.

———. 2022a. “FOFC and What Left-Right Asymmetries May Tell Us about Syntactic Structure Building.” *Journal of Linguistics* (April): 1–35. issn: 0022-2267, 1469-7742.

———. 2022b. *Negation and negative dependencies.* Oxford University Press.