

## **Coalescence: A unification of bundling operations in syntax**

Brian Hsu (hsu@email.unc.edu)  
*University of North Carolina at Chapel Hill*

### **Abstract**

This article revisits two proposed bundling operations within the architecture of the grammar: Feature Scattering (Georgi & Pianesi 1997), which accounts for variation in the distribution of features across functional heads, and M-Merger (Matushansky 2006), which accounts for head adjunction in head movement. While these mechanisms have been situated in the presyntactic lexicon and a postsyntactic module, respectively, I argue that the two patterns can receive a unified analysis in terms of one syntactic operation called Coalescence, which bundles structurally adjacent heads in a particular configuration. I claim that this eliminates redundancies in the proposed architecture of the grammar while maintaining prior empirical coverage, and sheds new light on long-puzzling properties of head movement. The proposal is illustrated in the analysis of several patterns of head-bundling in the inflectional and clausal domains.

### **1. Introduction**

Syntactic operations applying to heads (defined as syntactically indivisible bundles of features; Matushansky 2006) have played a pivotal role in explaining numerous word order patterns. This paper focuses on the analysis of generalizations of the following form:

- (1) Certain functional category features can be realized either on distinct terminal nodes or on a single terminal node. This variation can be observed both cross-linguistically and within a single language.

This generalization characterizes head movement, which often creates heads that are realized morphologically with individual exponents corresponding to positions within the movement chain. In addition, numerous analyses have proposed that languages vary in the number of heads on which a set functional features is distributed, even in the absence of movement, exemplified by Giorgi & Pianesi's (1997) Feature Scattering Hypothesis.

Both types of patterns can be described as cases where multiple features are *bundled* on a single head. However, the two phenomena have not previously received a unified treatment. Variation in the distribution of features on functional heads has been attributed to parameters of the lexicon (Giorgi and Pianesi 1997; Cowper 2005). Head movement and head-adjunction phenomena have received numerous approaches; broadly speaking, the patterns have been analyzed in terms of various syntactic operations (Travis 1984; Baker 1988; Pollock 1989; Giorgi and Müller 2010), a syntactic process whose output is subject to post-syntactic operations (Embick and Noyer 2001; Roberts 2005; Matushansky 2006; Harizanov 2014), or as a purely post-syntactic or phonological operation (Chomsky 2000; Boeckx and Stjepanovic 2001; Platzack 2013).

In this paper, I argue that both bundling patterns should be unified as the result of a single syntactic operation that applies to terminal nodes during the derivation, *Coalescence*. Informally speaking, Coalescence takes place to prune phrase structure trees by combining weak or unproductive branches of the tree with adjacent stronger ones. This is formalized as a distinction between *recessive* and *dominant* structures. Specifically, I will propose that all category features are first Merged on distinct heads, and that all recessive heads must subsequently be bundled with a dominant head during

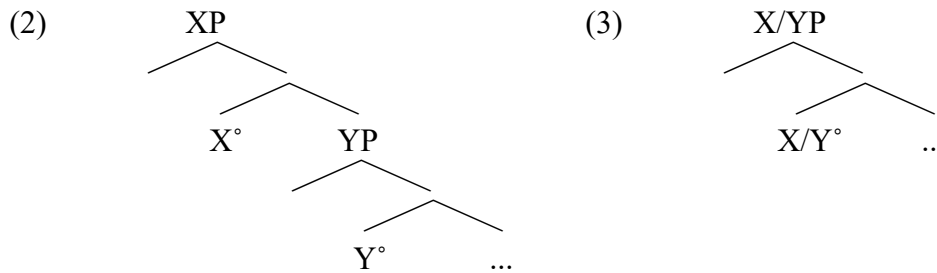
the derivation. Bundling patterns characterized by Feature Scattering are the result of Coalescence operations that apply prior to movement, while Coalescence fed by movement generates head adjunction structures. As Coalescence subsumes Feature Scattering and M-Merger effects, this eliminates redundancies in the proposed architecture of the grammar while maintaining prior empirical coverage. In addition, details of the proposed feature system are shown to permit new explanations for locality restrictions on head movement (Travis 1984) and cases where head movement licenses phrasal movement to the specifier of the target projection (Den Dikken 2007; Kandybowicz 2009; Gallego 2006, 2010).

In Section 2, I review the main claims of the Feature Scattering Hypothesis (Giorgi & Pianesi 1997) and recent approaches to syntactic head movement to show that both types of bundling share the same structural definition and locality restrictions. Section 3 outlines the formal properties of Coalescence, in context of a theory of head movement and phrasal movement that extends and refines the proposal of Matushansky (2006). Several case studies are presented in Section 4. Section 5 presents ways in which the proposed approach to head movement conforms with Minimalist desiderata, and discusses some of its implications for theories of affixation. Section 6 concludes the paper.

## **2. Bundling before and after syntax?**

### *2.1 Feature Scattering*

The Feature Scattering Hypothesis (Giorgi and Pianesi 1997) introduces a parameter on the distribution of category features. As a schematic example, the two features [X] and [Y] can either occur on separate heads  $X^\circ$ ,  $Y^\circ$  (2), or bundled on a single head  $X/Y^\circ$  (3).



The proposal contrasts with standard “cartographic” analyses in which each feature occurs on a single head in a universal hierarchy of projections (Cinque 1999; Kayne 2005; Cinque and Rizzi 2009), i.e. (2) is the only possible configuration of features [X], [Y]. While both approaches are compatible with the observation that extended projections contain an intricate, possibly universal hierarchy of features (Cinque 1999), the Feature Scattering Hypothesis finds support in instances where restrictions on movement are predicted in a structure with syncretic heads, but not one with empty projections. (Bobaljik and Thráinsson 1998; Erlewine 2016; Douglas 2017; Hsu 2017). Although not always couched in the same terms, this basic approach has been applied to variation in the realization of Tense, Aspect, and Mood projections (Giorgi & Pianesi 1997), subject positions in the IP domain (Poletto 2000), Infl/Agr° heads (Iatridou 1990; Speas 1991; Bobaljik 1995; Thráinsson 1996; Bobaljik and Thráinsson 1998; Ouhalla 1991), Voice° and Causative° heads (Pylkkänen 2002), deixis and reference in the nominal domain (Panagiotidis 2014; Höhn 2016), and positions in the extended complementizer domain (Douglas 2017; Hsu 2017).

In the proposal of Giorgi & Pianesi (1997), languages share a universal inventory of features, but differ in how individual features [F] are packaged onto the lexical items that enter syntactic derivations (a similar proposal is articulated in Cowper 2005). In other

words, the relevant locus of variation is within the lexicon, rather than syntactic derivations. To account for the observation that languages maintain a universal hierarchy of projections, regardless of the number of realized projections in a given structure, Giorgi and Piansesi (1997) propose that languages share a universal, hierarchically ordered set of features, and that the order in which they are checked (in this case, by Merge) must reflect this hierarchical order.

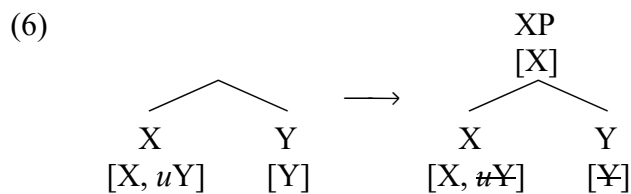
- (4) UNIVERSAL ORDERING CONSTRAINT: Features are ordered so that given  $F_1 > F_2$ , the checking of  $F_1$  does not follow the checking of  $F_2$ . (Giorgi & Piansesi 1997)

This constraint restricts the possible feature bundles that can be found within and across languages. To illustrate, the hypothetical ordered features  $[X] > [Y] > [Z]$  can be instantiated in the structural configurations below. One consequence of this metacondition is that a given set of features can be bundled on a head only if they are contiguous in the universal checking order. For instance,  $[Z]$  and  $[X]$  cannot appear on a bundled head  $Z/XP$  to the exclusion of  $[Y]$ .

- (5)
- |    |           |       |         |     |
|----|-----------|-------|---------|-----|
| a. | $[ZP$     | $[YP$ | $[XP$   | ... |
| b. | $[Z/YP$   |       | $[XP$   | ... |
| c. | $[ZP$     |       | $[Y/XP$ | ... |
| d. | $[Z/Y/XP$ |       |         | ... |
| f. | $*[Z/XP$  |       | $[YP$   | ... |

In addition, Giorgi & Piansesi propose an intimate link between the degree of bundling and the number of specifier positions necessary to license items in the numeration. For instance, if a given derivation contains a phrase that must occur in the specifier of a projection containing a feature  $[X]$ ,  $[X]$  cannot be bundled with feature  $[Y]$ , where  $[Y]$  follows  $[X]$  in the universal checking order.

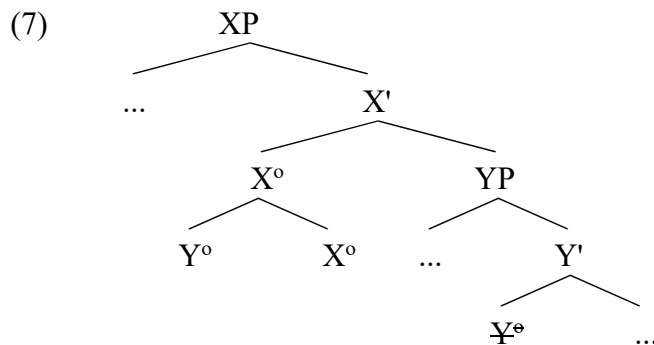
Let us contrast this system with how the universal order of feature checking is typically accounted for in a theory in which each category feature is realized on a separate head. Consider the proposal that the order of feature checking reflects a type of category selection (Svenonius 1994; Julien 2002; Matushansky 2006; Di Sciullo and Isac 2008). In this approach, each head contains an interpretable category feature, and optionally an uninterpretable category feature. Merge is triggered in order to check uninterpretable category features, in the configuration in (6). Uninterpretable category features enable heads to select the categories of their complements. Consequently, the universal order of feature checking is encoded in the atoms of the syntactic derivation. In this view, category features can be defined as features of syntactic objects that trigger External Merge (Di Sciullo and Isac 2008).



On the other hand, if some categories can be bundled onto single heads in the lexicon prior to numeration of the syntactic derivation, the restriction on the order of feature checking has to exist independently of lexical entries as a stipulated metacondition on bundling in the lexicon, such as the Universal Ordering Constraint. Restrictions on the order of feature checking are thus duplicated both in the syntax and in the lexicon, a redundant specification in separate modules of the grammar. In this work, I propose that this redundancy is dispensed with if the bundling of category features takes place during the derivation.

## 2.2 Head movement

Prior to the emergence of the Minimalist Program, head movement was largely analyzed as a single movement operation that creates an adjunction structure in which both the moved head and attracting head are dominated by a head-level projection (7), and that successive steps of head movement produce additional branching structure within complex heads (Travis 1984; Baker 1988). The structure is used to account for the observation that head movement tends to create morphologically complex forms in which features of the movement target position appear as affixes on the exponent of the moved head (Baker 1988; Julien 2002).



Key aspects of this “traditional” analysis become suspect in the bottom-up, derivational theory of syntax developed in the Minimalist Program, and the problematic aspects of head movement become particularly apparent when compared with phrasal movement (Chomsky 1994, 2000; Mahajan 2001). First, traditional head movement fails to extend the root node of the tree (Chomsky's 1995 Extension Condition), and as a result the moved head does not c-command its lower position. Head movement is also subject to a locality restriction not characteristic of phrasal movement; informally stated, head

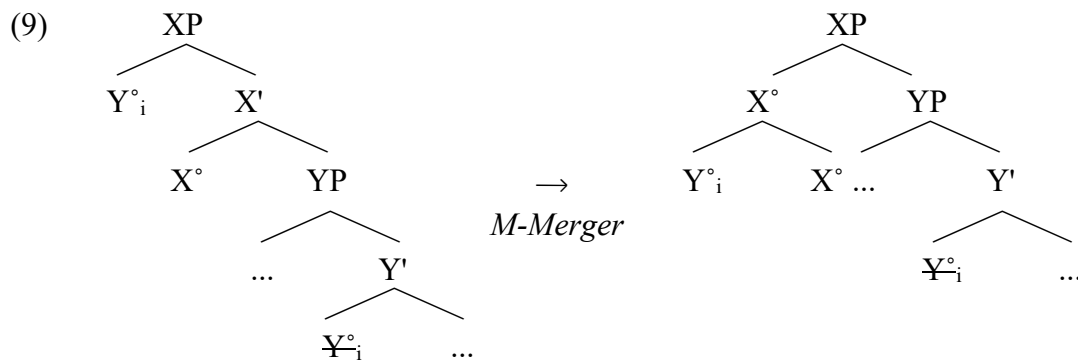
movement from  $X^\circ$  to  $Y^\circ$  cannot not skip an intervening head  $Z^\circ$  (Travis's 1984 Head Movement Constraint).

On the other hand, there is evidence for analyzing head movement as a syntactic operation. First, some instances of head movement have clear semantic effects, often related to scope and the licensing of negative polarity items (Lechner 2006; Matushansky 2006; Roberts 2010; Hartman 2011; Matyiku 2017). Furthermore, some types of head movement are necessary to license phrasal movement to the specifier of the target projection (Gallego 2006, 2010; Den Dikken 2007; Kandybowicz 2009), which is unexpected if head movement is postsyntactic but phrasal movement is not. One pattern of this type is German verb second (considered in more detail in Section 4.1), where phrasal movement to Spec, CP occurs only if C is filled by head movement of a tensed verb or auxiliary.

- (8)    a.  $[_{CP} \quad XP \quad [_{C'} \quad T+C \quad [_{TP} \dots \bar{T}]]$   
           b.  $[_{CP} \quad (*XP) \quad [_{C'} \quad C \quad [_{TP} \dots T]]$

Matushansky (2006) proposes a theory of head movement that addresses the incompatibilities of traditional approaches with a bottom-up derivation. The structure produced by traditional head movement (7) is argued to be built in two steps. First, a lower head is attracted to the specifier of the highest head in the derivation, as occurs in phrasal movement, satisfying the Extension Condition and c-command condition on movement.<sup>i</sup> The two heads are subsequently bundled into a complex head by an operation called M-Merger.





Matushansky proposes that head movement differs from phrasal movement in the featural trigger: phrasal movement is triggered by Agree (feature valuation), while head movement is triggered by c-selectional features of the attracting head. Because c-selection is a local relation between heads and their complements, this predicts the Head Movement Constraint. It remains unresolved, however, why c-selection does not always result in head movement. In other words, why do languages vary in head-movement paths within extended projections? The problem here reflects the difficulty in identifying a featural trigger in many instances of head movement that is distinct from the need for heads to be concatenated, i.e. bundled (Baker 1988; Julien 2002; Roberts 2005).

Lastly, M-Merger is proposed as a PF morphological operation that applies after Spell-out, rather than a syntactic one. The key motivation for this claim is to account for the impossibility of excorporation; Spell-out is expected to render the internal structure of the complex head opaque to later syntactic operations. This leaves the resulting head-adjunction structure immune to the Extension Condition and c-command condition on movement, which remain satisfied in the narrow syntax. However, the proposal requires an architecture of the grammar in which syntactic and morphological operations are interleaved throughout the derivation, and in which constituents created by the

morphological component [ $X^0$  [ $Y^0$   $X^0$ ]] remain accessible to later syntactic operations. I argue in Section 5 that the inaccessibility of subparts of the complex heads need not require an analysis of bundling as an operation that takes place after Spell-out. Rather, this inaccessibility is accounted for if no head-internal branching structure is created when heads are bundled.

### *2.3 Structural resemblances*

Feature Scattering and M-Merger are similar bundling mechanisms posited to operate in the presyntactic lexicon and postsyntactic morphology, respectively. While it may turn out that similar operations can take place in various components of the grammar, the pursuit of a theory that minimizes the complexity of grammatical derivations motivates a more unified analysis. Here, I argue that key structural similarities between Feature Scattering and M-Merger suggest that both types of bundling can be attributed to a single operation that applies in a single component of the grammar, syntax.

First, consider the locality conditions that constrain bundling in both Feature Scattering and M-Merger. Under the Feature Scattering Hypothesis, a pair of features that are adjacent in the universal checking order must be realized either in immediately adjacent projections or bundled in a single projection, and features that license specifiers are not bundled with features that immediately follow them in the universal checking order. Similarly, M-Merger applies to adjacent heads in an asymmetric c-command relation, with no intervening specifier. In this sense, both types of head-bundling are restricted by the same condition of head adjacency (to be defined in Section 3.1). The

upshot of this is that it is possible to generate both patterns from a bundling operation that applies during the syntactic derivation to structurally adjacent heads.

We can also note a similarity in the motivation for both types of bundling. Patterns in which head movement is needed to license a phrasal specifier in the target projection (Gallego 2006, 2010; Den Dikken 2007; Kandybowicz 2009) suggest that the target projections of head movement are unable to independently license a specifier, and that this ability is inherited from the lower, moved head. While it may not be initially obvious to look at Feature Scattering in these terms, we see again that the number of functional projections that instantiates a given set of features is determined by the number of specifier positions that are necessary to host lexical items in the numeration; A feature [X] that licenses a specifier cannot be bundled on the same projection as a feature [Y], where the checking of [Y] must take place after checking of [X]. In both M-Merger and Feature Scattering patterns, bundling has the effect of licensing “deficient” category features, which cannot host phrasal specifiers or be realized with a standalone head, by associating them structurally with a “prominent” feature that can.<sup>ii</sup> In essence, this can be viewed as a type of *prominence based licensing* pattern found in a range of phonological domains (Itô 1988; Goldsmith 1990; Steriade 1995; Walker 2011; among others).

Given the identical structural constraints on bundling proposed for Feature Scattering and M-Merger, I propose that their effects can be unified as a single syntactic operation. In this proposal, all category features enter the derivation as distinct heads, but can be concatenated into complex heads as the result of a syntactic operation that I will call *Coalescence*. The effects of Feature Scattering are the result of Coalescence operations

that apply prior to movement, while Coalescence fed by movement generates the effects of M-Merger.

### **3 Coalescence**

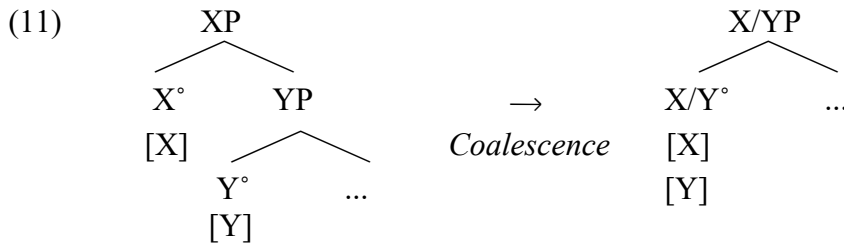
The proposal is framed within a bottom-to-top derivational theory of syntax (Chomsky 1995, 2000, 2008). Each derivation begins with a selection of lexical items to be manipulated, a.k.a. the numeration. Lexical items are then combined into a hierarchical constituent structure by internal or external Merge. Syntactic operations are triggered by features of lexical items, and syntactic variation arises from differences in the featural properties of lexical items (Borer 1984; Chomsky 1995). Based on the desiderata made in Section 2.1, I propose that all lexical items contain exactly one interpretable category feature [F] (i.e. features whose selection triggers External Merge). I assume that lexical items can enter the derivation with other interpretable and uninterpretable features (e.g. the lexical item containing [Tense] can also contain [ $u\phi$ ], [pres]). In this section, I define a bundling operation Coalescence and a feature system that determines when the operation applies.

#### *3.1 Defining Coalescence*

We first consider the structural definition of this bundling operation. I propose that Coalescence, similar to M-Merger, applies to structurally adjacent heads, i.e. heads that are not separated by a specifier. I define this relationship (a.k.a. head adjacency) as follows:<sup>iii</sup>

- (10)  $\alpha$  and  $\beta$  are head-adjacent iff:
- (i)  $\alpha$  and  $\beta$  are minimal projections,
  - (ii)  $\alpha$  asymmetrically c-commands  $\beta$ ,
  - (iii) there is no node  $\kappa$  that asymmetrically c-commands  $\beta$  and is asymmetrically c-commanded by  $\alpha$ .

In this configuration, Coalescence creates a single node that contains all features associated with the individual heads.<sup>iv</sup>



I depart from M-Merger and prior approaches to head adjunction by proposing that no branching structure is present in the newly formed head. There are several motives for this move, elaborated in Section 5. First, the absence of internal branching directly accounts for the impossibility of excorporation from bundled heads, without having to assume that bundling triggers Spell-Out to PF, as proposed for M-Merger. In Section 5, I show that this representation additionally obviates incompatibilities of traditional adjunction structures with the Chain Uniformity condition (Chomsky 1994).

### 3.2 Dominance and Recession

We now turn to defining the structural environment that triggers Coalescence that can account for the patterns attributed to Feature Scattering and M-Merger. I propose that this can be accounted for by proposing a binary distinction between *dominant* and *recessive* category features. Informally, recessive features are those that must occur on a bundled head (e.g. [Asp] obligatorily realized with [T] in T/AspP, or [T] that must be head-

adjoined to V). Dominant features are those that do not need to be bundled, and are potential ‘hosts’ for recessive features.

(12) A functional category feature is either *dominant* [ $F_D$ ] or *recessive* [ $F_R$ ]

I will assume that this is a distinction among formal features visible to syntactic operations, and not directly predictable from other factors. For instance, although there is a tendency for heads that undergo bundling to have null or affixal representations in surface phonology, it does not appear possible to state a phonological characterization on the types of phonologically dependent or affixal items that trigger head bundling to the exclusion of reduced forms that do not. For example, while triggers of bundling are often affixal in the sense that their exponents must be linearly adjacent to particular types of morphemes, the same generalization applies to bound roots (ex. Spanish *habl-* ‘speak’) that pattern as dominant heads throughout the derivation.<sup>v</sup> Furthermore, not all syntactic heads that have affixal exponents undergo head bundling (Section 4.2 returns to this point in contrasts between contracted negation and reduced auxiliaries in English).

In addition to the distinction between dominant and recessive category features, a distinction is required between *dominant* versus *recessive heads*. While category features are lexically specified as to whether they are dominant or recessive, the status of a head is determined by its featural composition. A head that contains at least one dominant feature is dominant (13), whereas a head that contains only recessive features is recessive (14). In all following examples, dominance is indicated with subscript  $_D$ , recessiveness with  $_R$ .

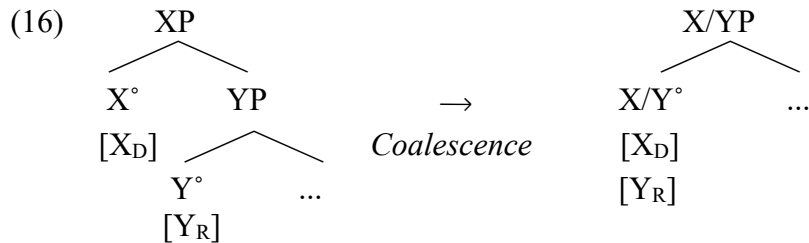
(13)  $X/Y^{\circ}_D$   
     $[X_D]$   
     $[Y_R]$

(14)  $X^{\circ}_R$   
     $[X_R]$

I assume that the grammaticality of surface forms is determined not only by principles of the syntax proper, but by well-formedness requirements of the Conceptual-Intentional (LF) and Articulatory-Perceptual (PF) interfaces (Chomsky 1993, 1995, 2000). I propose that the recessive heads are not legitimate PF objects, and Coalescence applies to ensure that no recessive heads remain at the end of the syntactic derivation. I will refer to this inviolable restriction as the Dominance Condition.

- (15) **DOMINANCE CONDITION**  
All terminal nodes of the syntactic representation contain a dominant feature.

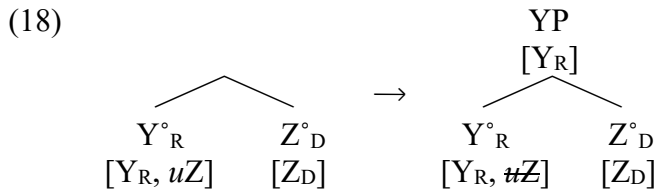
Coalescence applies only in the configuration of head-adjacency in (16), where a dominant head c-commands a recessive one. This has two important consequences: Merge of a recessive head cannot trigger Coalescence with lower heads, and pairs of dominant heads can not be bundled.



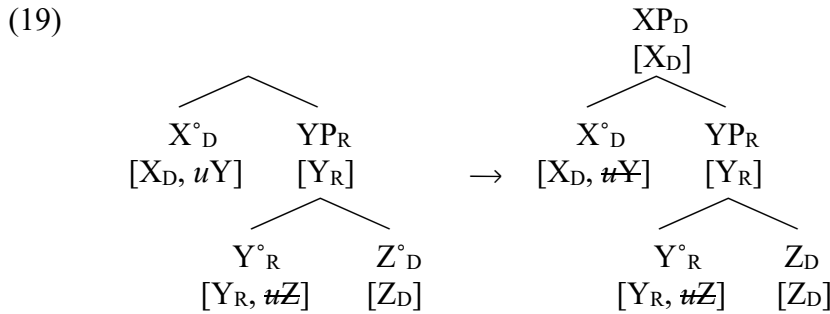
Here, I illustrate the derivation of a bundled  $\text{X/Y}^\circ$  head. I assume that the order of projections arises from the distribution of uninterpretable c-selectional features (Svenonius 1994; Julien 2002; Matushansky 2006; Di Sciullo and Isac 2008). I propose that the inventory of category features and their c-selectional properties are identical regardless of whether they are dominant or recessive.

- (17) *Numeration:*  
 X [ $X_D$ ,  $uY$ ]  
 Y [ $Y_R$ ,  $uZ$ ]  
 Z [ $Z_D$ ]

In the first step,  $Y^\circ$  c-selects  $Z^\circ$ .  $Z^\circ$  is dominant, as it contains a dominant category feature.  $Y^\circ$  is a recessive head as it contains only recessive category features.  $Y^\circ$  Merges with  $Z^\circ$  to check [ $uZ$ ] on  $Y^\circ$ , and the features of  $Y^\circ$  project to the newly formed node.

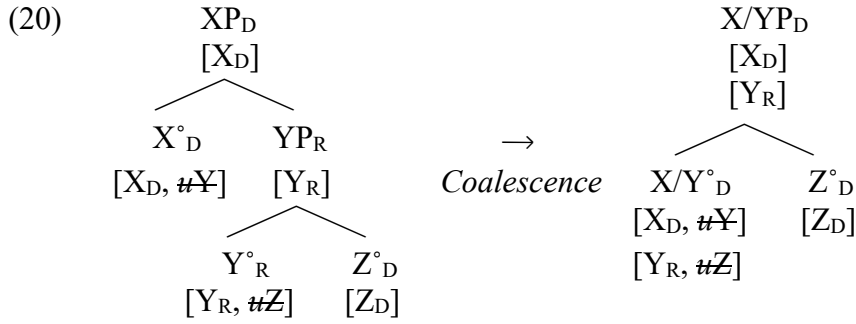


$X^\circ$ , whose category feature is dominant, then enters the derivation. Merge applies to check its [ $uY$ ] feature and creates a new phrase headed by X.



Because this step creates two heads in structural adjacency where the lower head is recessive, Coalescence applies. Since the head formed by X and Y contains one dominant category feature ( $[X_D]$ ), the head is dominant and no recessive heads remain in the workspace.





At this point, the question arises of how the label of the highest node is to be determined.

Now that the projecting head contains multiple interpretable category features, one expects both features to project to the root node (Di Sciullo and Isac 2008). A potential concern here is that the structure runs afoul of the expectation that syntactic operations do not alter the labels of syntactic objects that they manipulate, the No Tampering Condition (Chomsky 2008).<sup>vi</sup> In particular, Coalescence appears to alter the label of the root node.

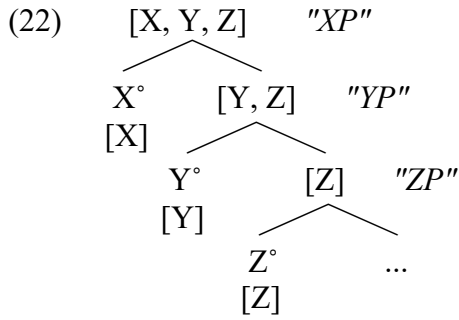
The problem can be avoided by adopting Keine's (2015) proposal of Category

Percolation, proposed independently to explain certain properties of extended projections (van Riemsdijk 1988, 1998; Grimshaw 1991, 2000). Specifically, Category Percolation is the claim that all heads inherit the category features of all lower heads within the same extended projection.

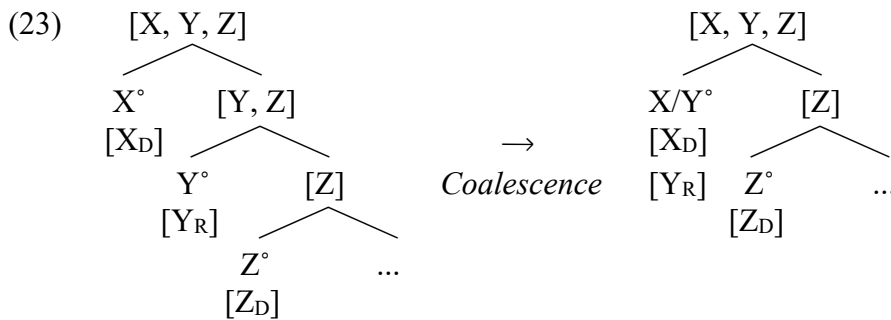
(21) CATEGORY PERCOLATION

Given an extended projection  $\Phi = \{\Pi_n > \Pi_{n-1} > \dots \Pi_1\}$ , the categorial features of  $\Pi_m$  percolate to  $\Pi_{m+1}$ . (Keine 2015)

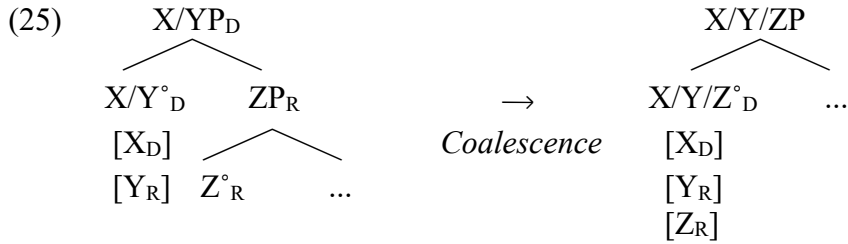
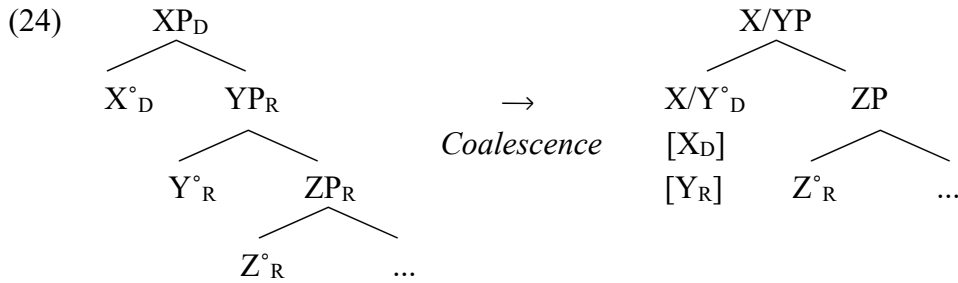
This is schematically represented in (22) for a structure in which the features  $[X]$ ,  $[Y]$ , and  $[Z]$  are in the same extended projection.



What is key for the present analysis is that if phrasal nodes contain all features of lower heads within the same extended projection, Coalescence will not produce root nodes whose labels are distinct from those that would exist if no bundling applied. This is shown in (23), where the highest phrasal node contains features [X, Y, Z]; The highest node contains the same features even if Coalescence applies to  $X^\circ$  and  $Y^\circ$ , because the set of features in "YP", [Y, Z], is a subset of those contained in "XP", [X, Y, Z].

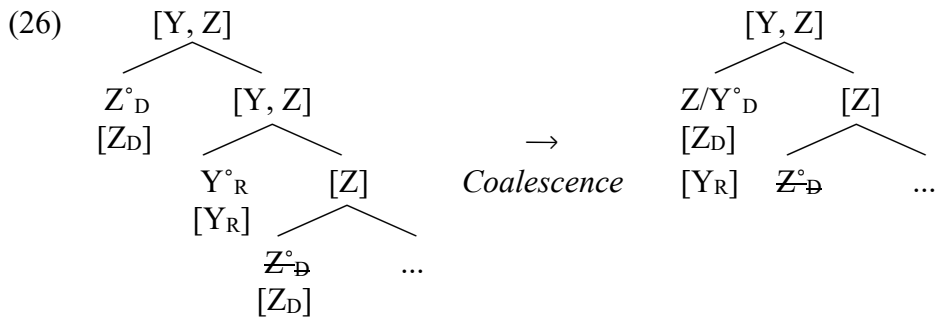


Lastly, complex heads that contain more than one recessive category feature are generated by successive “top-down” applications of Coalescence. At each step in (24) and (25), Coalescence applies to the topmost dominant head in the workspace and the immediately lower recessive head. For representational simplicity, checked features and features associated with non-head nodes will be omitted in the remainder of the paper.



### 3.3 Coalescence and head movement

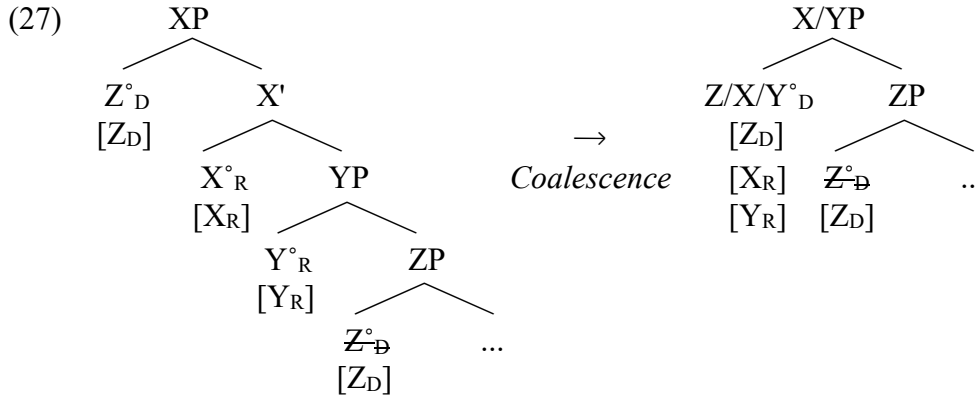
The previous examples show how Coalescence applies following external Merge of a dominant head. Coalescence can also be triggered following internal Merge of a dominant head lower in the tree. This is schematically shown in (26), in which dominant  $Z^\circ_D$  first moves to the specifier of recessive  $Y^\circ_R$ , providing the structural adjacency necessary for  $[Y_R]$  to be bundled into a dominant head. Again, Category Percolation ensures that the root node following Coalescence is featurally identical to the root node prior to movement.



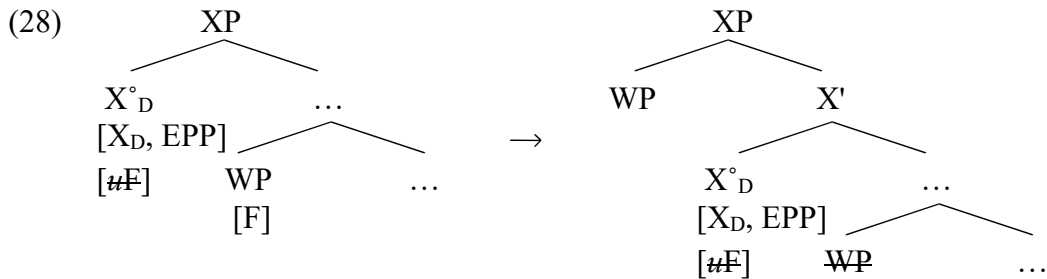
Section 4 discusses several cases where languages vary in whether the application of Coalescence to a recessive head is fed by internal or external Merge.

In the system of Matushansky (2006) both head movement and phrasal movement consist of movement of a lower item to the specifier of the attracting head. Both types of movement are thus given a single structural characterization that satisfies the Extension Condition. While phrasal movement is triggered by Agree initiated by a probe on the target head, head movement is triggered by c-selection features of the target head (thus deriving the head movement constraint). However, it is unexplained why c-selection does not always trigger head movement.

In the case of head movement, I propose that movement of a dominant head to the specifier of a recessive one is not triggered by a probe or selecting property of the target head, but is a Last Resort operation that ensures that the Dominance Condition is satisfied. Locality restrictions on head movement can be derived under additional assumptions that the target head attracts the closest dominant head that contains a subset of its category features (maintaining Category Percolation). Even in instances where one or more recessive heads intervene between the source and target positions when movement takes place, iterative application of Coalescence ensures that the two positions of the moved head will be in adjacent projections, essentially deriving the Head Movement Constraint (Travis 1984). This is illustrated below for a structure in which the dominant head  $Z^{\circ}_D$  closest to the target  $X^{\circ}_R$  crosses an intervening projection headed by  $Y^{\circ}_R$ .



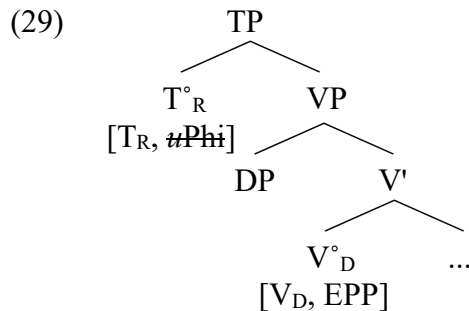
Following Matushansky (2006), I maintain that phrasal movement requires a [ $\mu$ F] probe on the target head to Agree with a constituent that it c-commands. In addition, I propose that phrasal movement takes place only if the attracting head has the EPP property, defined as the ability to trigger Merge of an item that does not contain a subset of its own category features. Specifically, uninterpretable [ $\mu$ F] probes can be checked if they c-command a suitable goal, but will only trigger movement of the goal if the probing head also contains an [EPP] feature.<sup>vii</sup> To account for the generalization that only dominant heads license specifiers, I propose that only lexical items that contain a dominant category feature can also carry an [EPP] feature. The necessary configuration for phrasal movement is shown in (28) below.



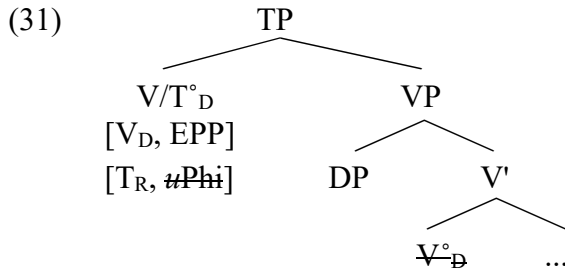
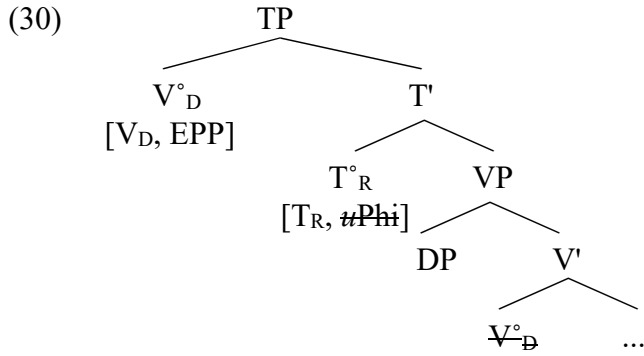
A key consequence of the dissociation of [EPP] features from particular probes is that an [EPP] feature can be associated with more than one probe during the derivation,

specifically in derivations where a dominant head moves and undergoes Coalescence with recessive heads that contain distinct [ $\mu$ F] probes. Here, I will make the ancillary assumption that checked uninterpretable features need not be immediately deleted from the derivation (Pesetsky & Torrego 2000).

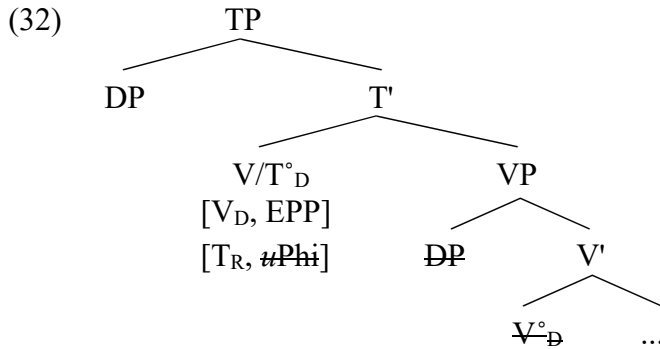
The basic workings of head movement, Coalescence, agreement, and phrasal movement can be succinctly illustrated with a schematic example of Romance-style V-to-T head movement. In the first step (29), a recessive  $T^\circ$  head with a phi-feature probe is Merged upon the completion of VP, which includes a dominant  $V^\circ$  head and a subject DP in Spec, VP. At this point, [ $\mu$ Phi] on  $T^\circ$  is checked by agreement with the subject DP. However, phrasal movement of the subject does not take place since T lacks  $^\circ$  [EPP].



In this case, the Dominance Condition is satisfied by moving  $V^\circ$  to Spec, TP (30). As dominant  $V^\circ_D$  now c-commands recessive  $T^\circ_R$ , Coalescence applies to create a bundled V/ $T^\circ$  head (31).

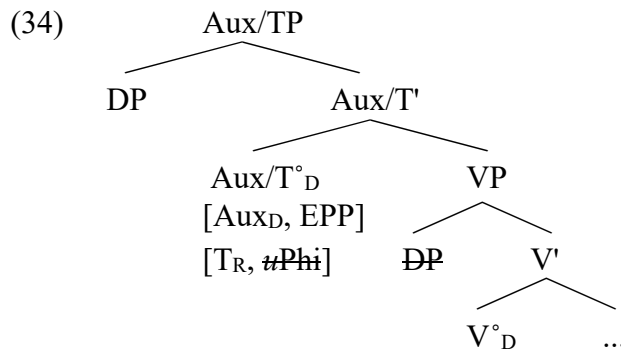
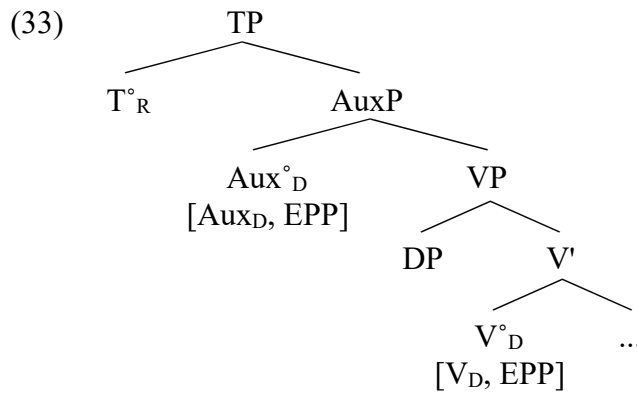


At this point, the bundling of  $V^\circ$  and  $T^\circ$  into a single head  $V/T^\circ$  enables  $[\#Phi]$  probe of  $[T]$  to inherit the use of the EPP feature associated with  $[V_D]$ , thus triggering movement of the subject to Spec, TP.



In Section 4.1, I argue that this ‘delayed gratification’ pattern in which probes do not trigger movement of their goals until later steps of the derivation accounts for the flexible discourse interpretations of sentence-initial items in verb second phenomena (Fanselow 2009; Fanselow and Lenertová 2010).

Here, I will briefly discuss implications of the proposals for the analysis of languages like English, in which subjects appear in Spec, TP, while main verbs remain in VP. In clauses with a modals or an auxiliary, the highest item moves to T° (Pollock 1989), so we can posit that English T<sub>R</sub> is recessive when first Merged, and that modals and auxiliaries are dominant heads with the [EPP] property. This first-Merge clause structure is shown in (33), and the result of Aux-to-T movement followed by Coalescence and [EPP] inheritance is given in (34).



Because subjects also occupy Spec, TP in clauses without a modal or auxiliary, [Tense] must occur on a dominant head with an EPP feature. In order to maintain that [T<sub>R</sub>] is recessive in all clauses, a plausible solution is to posit that clauses without overt auxiliaries nonetheless contain a dominant Aux°<sub>D</sub> with [EPP] (cf. Emonds 1970; Pollock



1989; Wilder and Ćavar 1994a for similar instantiations of a null Aux or null *do* proposal).<sup>viii</sup>

It should be noted that the Last Resort analysis of head movement does not extend to proposed cases "long head movement" where head movement appears to skip over a potential intervener [ $X^{\circ}_i \dots [Y^{\circ} \dots [X^{\circ}_i \dots]]$ ] (Rivero 1991; Borsley et al. 1996; Roberts 2010 and references therein). However, these patterns can potentially arise under conditions where the target projection has a probing dominant head with an EPP feature whose goal can be a head. In the best-known cases of participle fronting in many Slavic languages (Lema and Rivero 1989; Wilder and Ćavar 1994; Bošković 1995; Embick and Izvorski 1995; Migdalski 2006, a.o.) and Breton (Rivero 1991; Borsley et al. 1996; Roberts 2010), the target position can be filled by either a head or a phrase. Crucially, the present proposal predicts that long head movement cannot take place to satisfy the Dominance Condition by triggering Coalescence. There is indeed no evidence that the moved and target heads (e.g. fronted participle + auxiliary in Slavic) can move together in later steps of the derivation.

- (35) a. Az sŭm                      čel                      kniga-ta (*Bulgarian*: Migdalski 2006)  
           I    be.AUX.PRES.1SG read.PART.F.SG book-the
- b. Čel                      sŭm                      kniga-ta  
           read.PART.F.SG be.AUX.PRES.1SG        book-the  
           'I have read the book'
- (36) a. Al levr    en    deus    lennet Tom        (*Breton*: Borsley et al. 1996)  
           the book 3SG.M has    read    Tom  
           'Tom has read the book.'
- b. Lennet    en    deus    Yann al levr  
           read        3SG.M has    Yann the book  
           'Yann has read the book.'

## 4 Case studies

This section presents several studies of variation in the realization of functional projections of the inflectional and complementizer domains, which illustrate key aspects of the proposals on Coalescence and movement. It considers two instances in which languages permit a given head to have both dominant and recessive variants, the Kashmiri V2/V3 alternation and the realization of sentential negation in English. In addition, within-language variation in the realization of the Catalan past perfect is attributed to differences in the distribution of dominant and recessive features.

### *4.1 Relaxed Verb Second Effects*

To illustrate how Coalescence can account for Feature Scattering analyses, I present a reinterpretation of Hsu's (2017) analysis of verb second (V2) effects, whose aim is to generate a range of attested 'strict' and 'relaxed' V2 patterns. In well-known 'strict' V2 patterns like that of Standard German and Dutch, finite verbs move to a C-domain projection, and are preceded by a single phrase. Traditionally, these patterns have been analyzed as the result of verb movement from Infl to C, followed by movement of exactly one phrase to Spec, CP (den Besten 1983).

$$(37) \quad [_{CP} \quad XP \quad V-C^{\circ} \quad [_{IP} \dots V^{\circ} \dots \cancel{XP} \dots$$

However, Rizzi (1997) and later research in the Cartographic Program have produced a range of evidence that the traditional "CP" contains a series of hierarchically ordered functional projections associated with clause typing and information structure features, as illustrated in the 'core' structure of Rizzi (1997). If the expanded inventory of left-peripheral features is universally present, then the generation of strict V2 languages

requires language-specific restrictions on the number of features that can be simultaneously instantiated in phrase structure.

(38) [ForceP ... [TopicP ... [FocusP ... [FinitenessP ... [IP

In addition, there is a variety of ‘relaxed’ V2 languages that show verb movement to the C-domain, and allow more than one phrase to simultaneously precede the verb (either optionally or obligatorily). In these languages, the ordering of pre-verbal constituents is restricted based on discourse properties, consistent with the order in (38) (Poletto 2002; Poletto 2014; Benincà and Poletto 2004; Walkden 2017; Wolfe *to appear*). For example, Ingush, whose V2 pattern otherwise resembles that of Standard German (Nichols 2011), permits V3 orders in which a topic XP and focused XP precede the verb, in the order *Topic + Focus + V*.

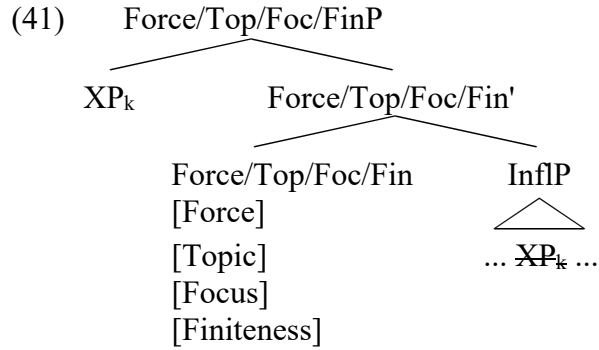
(39) Ingush: Topic + Focus + V  
 [Jurta jistie] [joaqqā sag] ull cymogazh jolazh.  
 town.GEN nearby AGR.old person lie.PRS sick.CVB.SIM AGR.PROG.CVB.SIM  
 'In the next town an old woman is sick.' (Nichols 2011: 683)

A further articulated structure is found in languages like Old Italian, which additionally permits frame-setting adverbials in the highest, clause-initial position.

(40) Old Italian: Frame-setting adverb + topic + focus + V  
 [e per volontà de le Virtudi] [tutta questa roba] [tra' poveri]  
 and by will of the virtues all this stuff among poor.PL  
 dispense.  
 distribute  
 'And according to the will of the virtues, distributed all these goods among the poor.' (Poletto 2014: 16)

Hsu proposes that the primary parameter that drives variation in the ‘strictness’ of V2 patterns is in the number of heads on which left-peripheral features are distributed. Strict

V2 patterns are generated when all left-peripheral features are bundled on a single head, which attracts a single specifier.

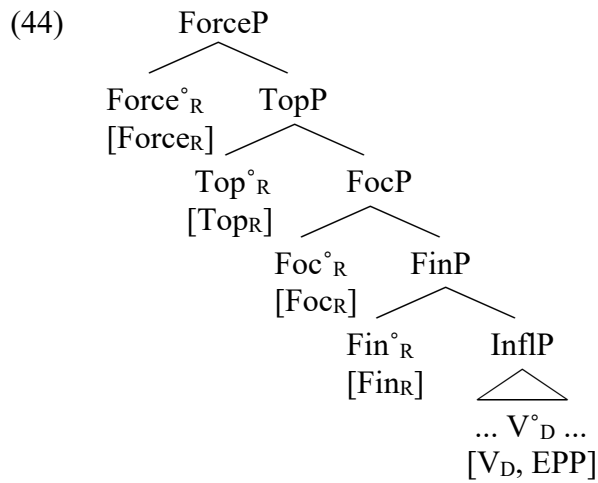


Relaxed V2 patterns are generated by when left-peripheral features are distributed across multiple heads, whose hierarchical order remains consistent with those proposed in ‘one feature - one head’ (Cinque and Rizzi 2008) Cartographic analyses (see also Kayne 2005).

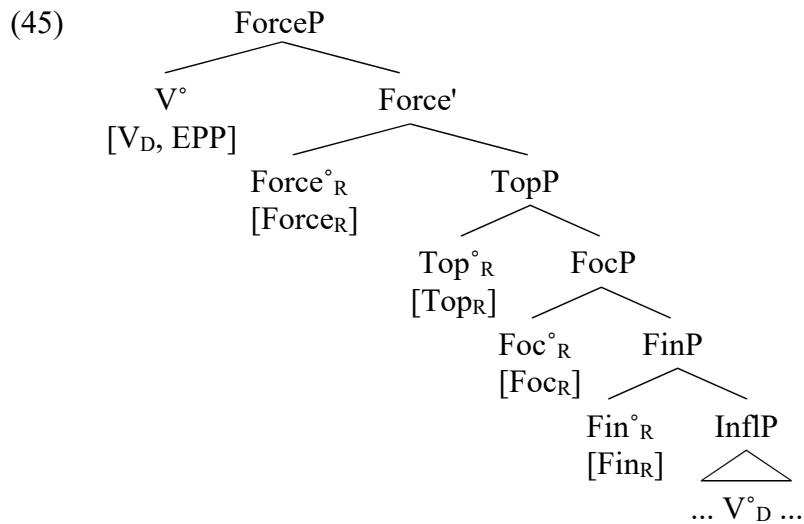
(42) Ingush *Topic Focus V ...*  
 $[_{\text{Force/TopP}} \text{XP}_{\text{top}} \quad [_{\text{Foc/FinP}} (\text{XP}_{\text{foc}}) \quad \text{V-Foc/Fin}^\circ \quad [_{\text{IP}} \dots \text{V}^\circ]$

(43) Old Italian *Frame-setter Topic Focus V ...*  
 $[_{\text{ForceP}} \text{XP}_{\text{frame}} \quad [_{\text{TopP}} (\text{XP}_{\text{foc}}) \quad [_{\text{Foc/FinP}} (\text{XP}_{\text{top}}) \quad \text{V-Foc/Fin}^\circ \quad [_{\text{IP}} \dots \text{V}^\circ]$

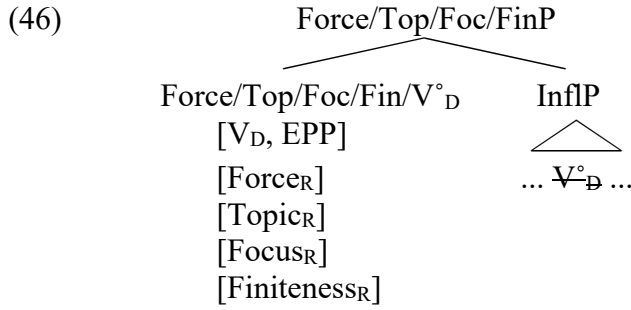
I now present a derivational analysis of these patterns in terms of Coalescence. To generate the idealized strict V2 pattern, suppose that all left-peripheral category features (including Force) are recessive, and a recessive head corresponding to each feature is Merged into the derivation.



If the derivation were to end here, the structure would not be interpretable at PF, as it contains recessive heads. One strategy to avoid this (the only one available in main clauses) is Last Resort movement of  $V^{\circ}_D$  to Spec, ForceP.

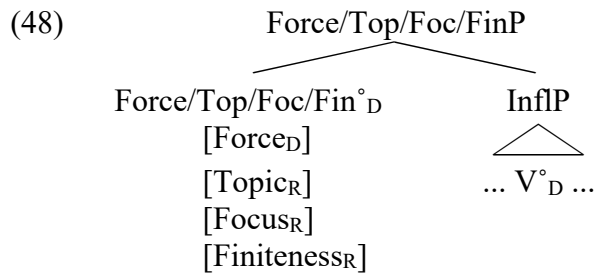
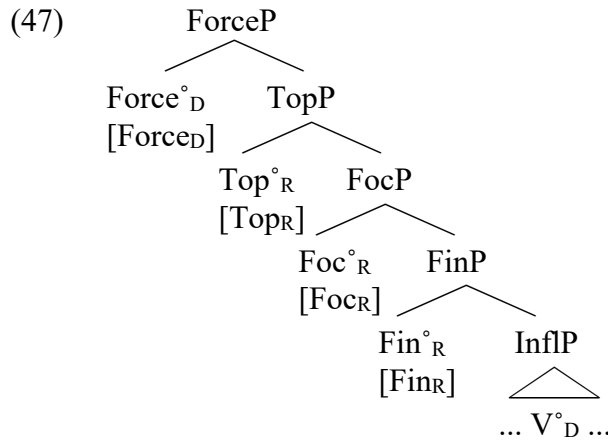


This produces the environment for Coalescence, which applies iteratively until all recessive features are bundled into a dominant head.



The association of the [EPP] property with [V<sub>D</sub>], rather than a left-peripheral feature, accounts for the fact that subsequent phrasal movement to Spec, ForceP has several possible pragmatic functions. In Standard German the first position can be occupied by either a given information topic, contrastive focus, or pragmatically unmarked subject (Mohr 2009). Suppose that German permits [Topic<sub>R</sub>] to be associated with a [*u*Topic] probe, [Focus<sub>R</sub>] with [*u*Contrast], and [Finiteness<sub>R</sub>] with [*u*D], and that these can be checked via Agree with a constituent lower in the clause.<sup>ix</sup> However, phrasal movement occurs only if the probe is found on a dominant head with [EPP], a situation that arises only as the result of verb movement. Assuming that there are no priority restrictions in German on which probe triggers movement in concert with [EPP], we predict that either focus, topicalization, or subject movement can take place in the final step, thus deriving the “unrestricted edge feature” property of verb second clauses (Fanselow 2009; Fanselow and Lenertová 2010).

Now suppose that rather than moving V°<sub>D</sub>, a dominant Force°<sub>D</sub> head can be externally Merged (47). This allows Coalescence to apply without the need for verb movement out of InflP (48).



Even if [Topic<sub>R</sub>], [Focus<sub>R</sub>], and [Finiteness<sub>R</sub>] are associated with probes, agreement cannot trigger phrasal movement, because [Forced] is not associated with the [EPP] property.

The availability of both internal Merge and external Merge options satisfy the Dominance Condition accounts for 'asymmetric' V2 patterns like Standard German, in which embedded clauses can contain an overt complementizer (49a) or show V2 order (49b), but not both (49c). Specifically, *dass* is the realization of a dominant Force° head directly Merged in the left periphery that does not have an EPP property. V2 clauses are generated by movement of dominant V<sub>D</sub>°, with subsequent movement to first position permitted by [EPP] on V<sub>D</sub>°.<sup>x</sup>

- (49) a. Er sagte [dass er morgen kommt]  
He said that he tomorrow comes

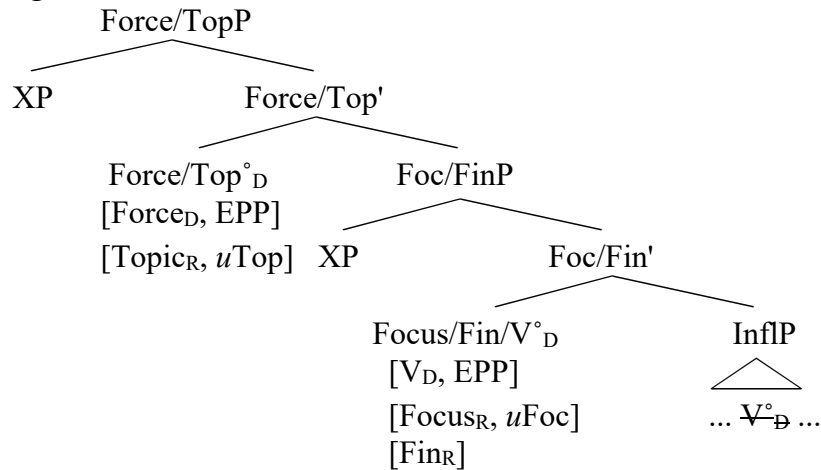
- b. Er sagte [er kommt morgen]  
 He said he comes tomorrow  
 'He said that he is coming tomorrow.'

- c. \*Er sagte [dass er kommt morgen]  
 He said that he comes tomorrow

(Holmberg 2015 after den Besten 1983)

As the number of dominant category features increases, fewer applications of Coalescence take place, leaving a greater number of independent heads in the left periphery. The Ingush V3 pattern in which topics and foci can simultaneously precede the verb crucially requires the [Topic] and [Focus] features to be realized on separate dominant heads. This occurs if [Focus<sub>R</sub>] compels movement of the verb, creating a lower dominant head, and if a [*u*Top] probe is associated with a dominant head above FocusP.

(50) *Ingush V3*



Further support for the distinction between dominant and recessive heads is found in languages that contain both dominant and recessive variants of the same category feature. This accounts for an otherwise puzzling V2/V3 alternation in Kashmiri. Kashmiri exhibits a strict V2 requirement in declarative main clauses, in which no more than one preverbal phrase is permitted (Bhatt 1999; Manetta 2011; Munshi and Bhatt 2009).



- (51) a. laRk-as dyut rameshan raath kalam  
 boy gave Ramesh yesterday pen  
 'It was a boy to whom Ramesh gave a pen yesterday'
- b. \*tem raath dyut akh laRk-as kalam  
 he yesterday gave one boy pen (Bhatt 1999)

There is some inconsistency in descriptions of the information structure characteristics of Kashmiri V2. In the descriptions of Bhatt (1999) and Manetta (2011), non-subjects in the first position of declarative clauses must be focused, not topicalized. On the other hand, Munshi and Bhatt (2009) report that first-position non-subjects can be either topics or foci, though the interpretations must be distinguished by intonational contour.

In light of these patterns, the word order of wh-questions in Kashmiri is initially puzzling. As is typical of languages with V2, wh-phrases obligatorily precede finite verbs in Kashmiri. However, Bhatt (1999) and Manetta (2011) report that if the clause contains eligible topics, one of them is preferably placed before the wh-phrase. This pattern is unexpected both because it appears to be an obligatory deviation from strict V2, and because topicalization to a preverbal position is not possible in declarative clauses in their descriptions.

- (52) a. tse kyaa dyutnay Rameshan  
 you what gave Ramesh  
 'As for you, what is it that Ramesh gave?'
- b. ?kyaa dyutnay Rameshan tse  
 what gave Ramesh you (Bhatt 1999)

Munshi & Bhatt (2009) similarly note the acceptability of [XP<sub>topic</sub> XP<sub>wh</sub> V ...] order in questions, but describe topicalization to first position as optional, rather than obligatory.

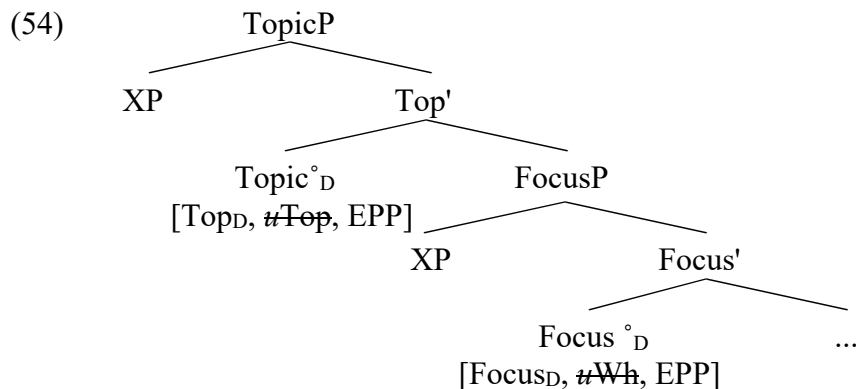
While I cannot explain the source of disagreement between these descriptions, I will proceed with a working assumption that they are correct descriptions of related but

distinct grammars, and describe how each system can be accounted for in the present proposal.

At first glance, the Kashmiri V2/V3 alternation appears puzzling in that the realization of the high topic position *depends* on the presence of a *wh*-phrase. However, it can be straightforwardly accounted for if Kashmiri has two types of Focus heads – a dominant version with a [*uWh*] probe in declarative clauses, and a recessive version with a [*uContrast* probe] in interrogative clauses. In contrast, languages like Ingush, where V3 is available in both declaratives and interrogatives, [Focus] is always dominant regardless of its other properties.

Restricting our attention to the structural realization of topic and focus, consider how the derivation proceeds if the Numeration contains the dominant [*uWh*] Focus head. Since both Topic° and Focus° are dominant as Merged, Coalescence can not apply, leaving [Topic] and [Focus] realized in separate projections. Because both heads contain an EPP feature, this creates the V3 word order in interrogative clauses.

- (53) Numeration for V3  
 [Focus<sub>D</sub>, *uWh*, EPP]  
 [Topic<sub>D</sub>, *uTop*, EPP]



In Kashmiri declarative clauses, the Focus head with a [*u*Contrast] probe is recessive. Coalescence applies once the dominant Topic° head is Merged. This bundles the [Topic] and [Focus] features into a single projection, leaving only one position available for movement.

- (56)
- TopicP

Topic<sup>°</sup><sub>D</sub>

FocusP

[Topic<sub>D</sub>, #~~Top~~, EPP]

Focus<sup>°</sup><sub>R</sub>

[Focus<sub>R</sub>, #~~Contrast~~]

...

→

Topic/FocusP

Topic/Focus<sup>°</sup><sub>D</sub>

...

[Topic<sub>D</sub>, #~~Top~~, EPP]

[Focus<sub>R</sub>, #~~Contrast~~]

35

Although there are other languages that permit  $XP_{Topic} XP_{Wh} V$  verb-third order, but not  $*XP_{Topic} XP_{Focus} V$ , such as Badiotto Rhaetoromance (Poletto 2002) and Yiddish (Diesing 2005), the opposite pattern remains unattested, suggesting that no language can contain both  $[Focus_D, uContrast]$  and  $[Focus_R, uWh]$ . Such patterns can be explained by the proposal of Cinque (1999) that all functional features can be associated with either a marked or default value. Suppose that  $[Focus]$  always has  $[uContrast]$  as its default value and  $[uWh]$  as its marked value. The apparent restriction is that it is not possible for a category feature with a marked value to be recessive, while the same feature with the unmarked value is dominant. It is possible however, for both instances of a category head to be recessive or for both to be dominant (as is the case for  $[Focus]$  in Ingush).

#### 4.2 English negative contraction

English negation is another case where a head has both dominant and recessive varieties. English has a 'full' negative morpheme (orthographic *not*) and a contracted form (orthographic *n't*). In many contexts, the two forms appear to be in free variation, with the contracted form apparently derived by optional phonological reduction.

- (57) a. Michael did not make a mistake.  
b. Michael didn't make a mistake.

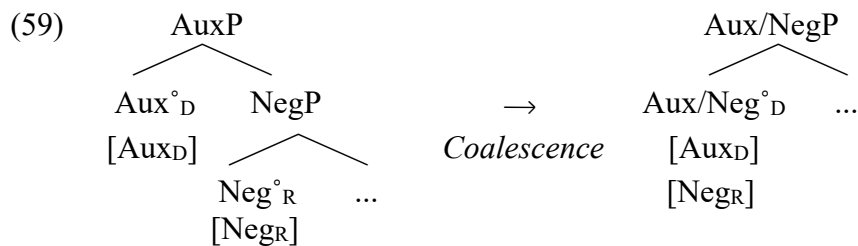
As noted by Zwicky and Pullum (1983) however, the distribution of the two forms is constrained by syntactic factors, and the use of a particular form is obligatory in certain contexts. For example, consider the case of negative inversion. In English, auxiliary verbs raise to a pre-subject position in interrogative contexts. If the negation morpheme raises along with the auxiliary, use of the affixal form is obligatory (58a). This gives the effect of contraction feeding raising. On the hand, only the full form is possible if the negative

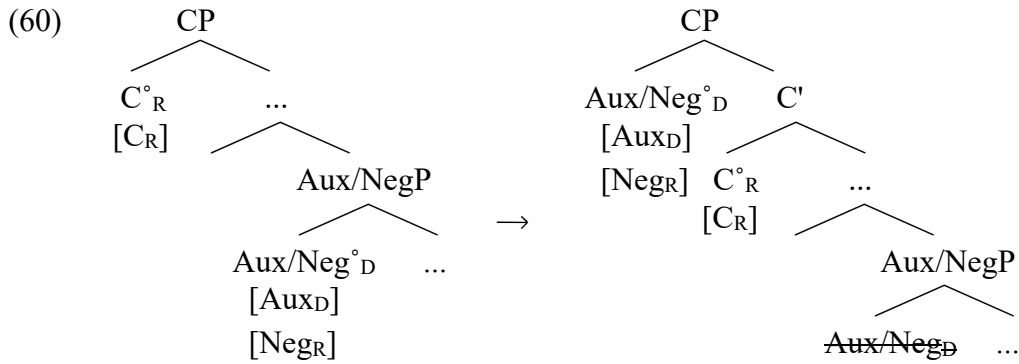
remains in a post-subject position (58b). Under an approach where the affixal form is derived by an operation that applies after the syntactic derivation, the obligatory use of the affixal negative when it raises with the auxiliary is unexpected.

- (58) a. Didn't Lindsay host the gala? (cf. \*Did not Lindsay host the gala?)  
 b. Did Lindsay not host the gala? (cf. \*Did Lindsay n't host the gala?)

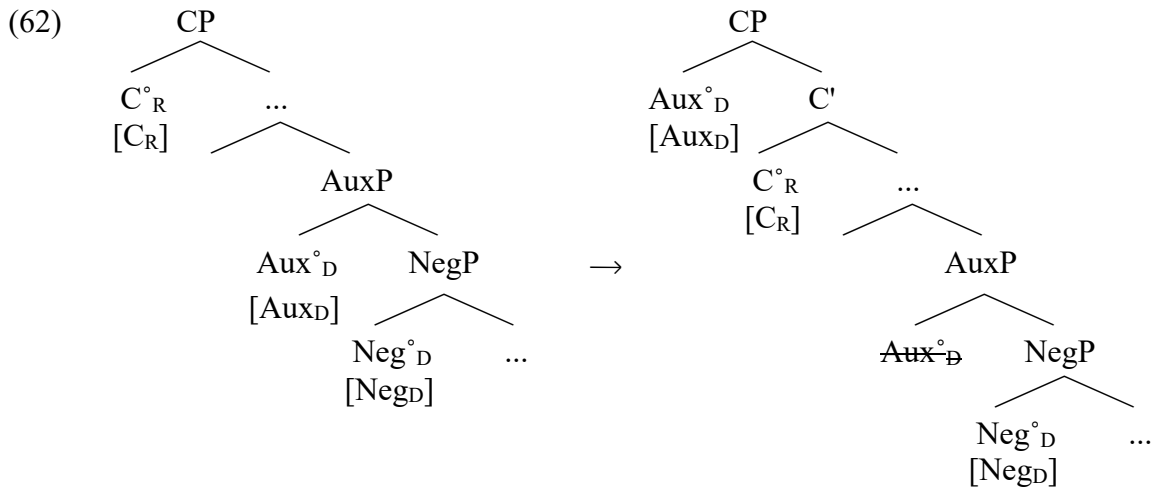
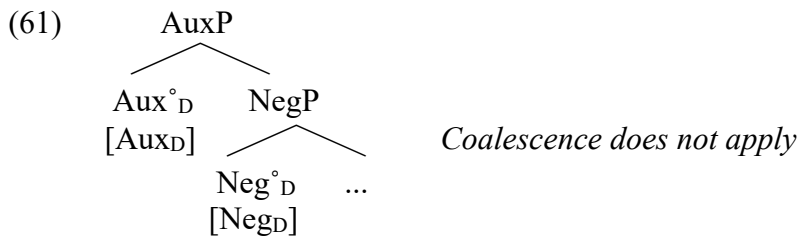
Matushansky (2006) makes the key observation that the different distribution of the full and contracted forms is explained if contracted negation is formed by head-bundling during the derivation. Specifically, she proposes that bundling by M-Merger optionally applies to  $\text{Neg}^\circ$  and  $\text{Aux}^\circ$  as soon as the items are Merged, and that M-Merged  $\text{Neg}^\circ$  corresponds to *n't*. If  $\text{Aux}^\circ$  and  $\text{Neg}^\circ$  are M-Merged prior to movement, both negation and the auxiliary undergo movement together when  $\text{Aux}^\circ$  is attracted to  $\text{C}^\circ$ . If M-Merger does not apply, the auxiliary moves alone.

This analysis has a straightforward reinterpretation using Coalescence. The different distribution of the full and contracted forms is accounted for if the full form enters the derivation with a dominant category feature  $[\text{Neg}_D]$ , while the contracted form is first Merged with a recessive feature  $[\text{Neg}_R]$ . The recessive head undergoes Coalescence once the dominant auxiliary is Merged (59). Thus, when recessive  $\text{C}^\circ$  attracts the closest dominant head, it attracts bundled  $\text{Aux}/\text{Neg}^\circ$  (60).





On the other hand, the dominant  $\text{Neg}^\circ$  head that corresponds to the full form does not undergo Coalescence when the auxiliary is Merged. In inversion contexts, interrogative  $\text{C}^\circ$  thus attracts only  $\text{Aux}_D^\circ$ , leaving  $\text{Neg}_D^\circ$  in its first-Merge position (62).



Some additional constraints on the distribution of contracted negation are worth noting. First, negative contraction is not possible after non-verbal items, non-finite verbs, or auxiliaries, as in the following examples (based on Zwicky and Pullum's ex. 11, 14):

- (63) a. Buster doesn't try {\*n't/not} to pay attention; he just can't help it.  
 b. For him {\*n't/not} to use an alliance-approved assistant is the last straw.  
 c. Should you have {\*n't/not} held the door for the man in the pricy suit?

Rather than a constraint on the possible clitic hosts for *n't*, the unavailability of *n't* in these examples indicates that contraction is sensitive to the contrast between sentential negation and constituent negation (Klima 1964). Among other differences in their distribution, sentential negation occurs only immediately following tensed auxiliaries and *do*, while constituent negation precedes a range of negated phrases. From this, we can posit that sentential [Neg] is recessive, whereas constituent [Neg] is dominant.

Second, negative contraction differs in key ways from auxiliary reduction (Zwicky 1970; Anderson 2008), which affects English auxiliary forms *is*, *has*, *would*, *had*, *have*, *am*, *are*, and *will*. First, reduced auxiliaries do not form syntactic units with their hosts in the manner of contracted negation. For instance, reduced auxiliaries do not pattern as if they raise with elements that they are affixed to.

- (64) a. Who's going to Phoenix?  
 b. \*Who's do you think ~~who's~~ going to Phoenix?

Whereas contracted negation only follows tensed auxiliaries, reduced auxiliaries *is*, *has*, *would*, *had*, are generally unrestricted by the category or phrase structure status of the items that they precede: *no touching's allowed by the guards / the role that he auditioned for's been written out.*<sup>xi</sup>

These contrasts between contracted negation and reduced auxiliaries highlight the impossibility of relying on surface exponence to determine the dominance or recession of a corresponding category feature. However, there is evidence from their syntactic patterning that auxiliaries correspond to dominant heads. First, the highest auxiliary in the

clause head-moves to TP, which subsequently triggers subject movement to spec, TP. Furthermore, the possibility of quantifier float with all auxiliaries in a sequence (Sportiche 1988) indicates that all auxiliary heads carry [EPP], a unique property of dominant heads.

### 4.3 *The Catalan Perfect*

I proposed in Section 3 that recessive heads can be eliminated either by External Merge or Internal Merge of a dominant item, and that this accounts for the 'asymmetric' V2 alternation where overt complementizers and V2 are in complementary distribution. Here, we consider another case in which a language varies in whether a recessive head is eliminated by External Merge or movement: the realization of the past perfect in Catalan.

Oltra-Massuet (2013) notes some dialects of Catalan express the past perfect either in a synthetic form, in which the subject and tense and aspect morphemes are realized as suffixes to a lexical verb, and an analytic form, in which subject agreement is realized on an auxiliary verb, *anar* 'to go.'

(65) purific-ares  
purify-2SG.PST.PERF  
'you purified'

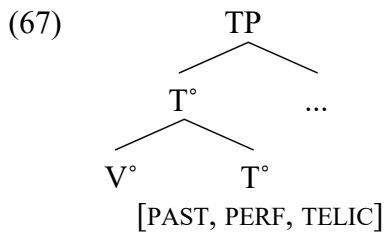
(66) vas purificar  
AUX.2SG purify  
'you purified'

An unusual and important property about the alternation is that there is no apparent semantic difference between the two ways of forming the past perfect. As described by Oltra-Massuet, "these forms do not express different lexical or truth-conditional semantics, nor do they show different morpho-syntactic functions, and individual speakers use some subset of them without distinction." It is further clarified that both

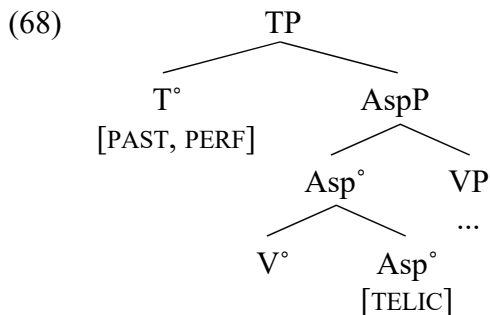


within and across speakers, variation depends on the lexical items and conjugations used; speakers do not probabilistically use both forms for any given verb and conjugation pair.

Oltra-Massuet proposes that the form of the syncretic past perfect is derived by head-movement of the verb through Aspect° and Tense° heads. Abstracting away from the distinction between V and little *v*, not relevant to the present analysis, she proposes the following structures. In synthetic forms like *purificares*, the verb moves to a Tense° head that carries specifications for past tense, perfective aspect, and telicity.

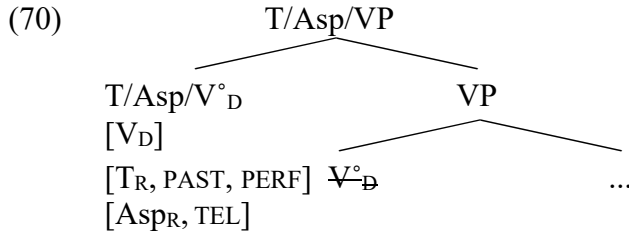
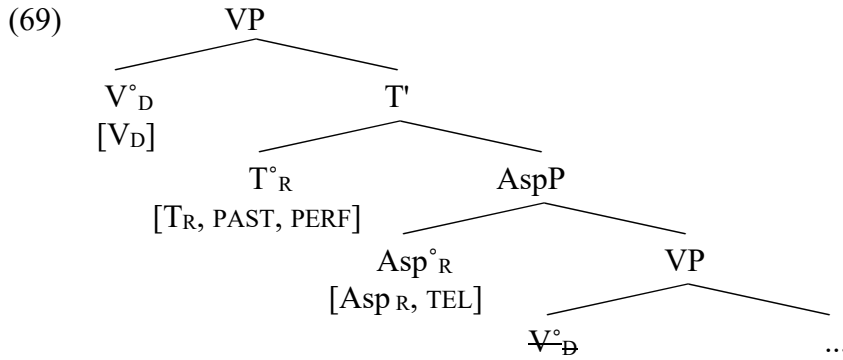


In the analytic past prefect, Oltra-Massuet proposes that [+PAST, +PERF] are associated with T°, while the [+TELIC] feature is contained in a separate Aspect° head. Furthermore, verb movement stops at Aspect°, and the V-Asp° head is pronounced as the participle.

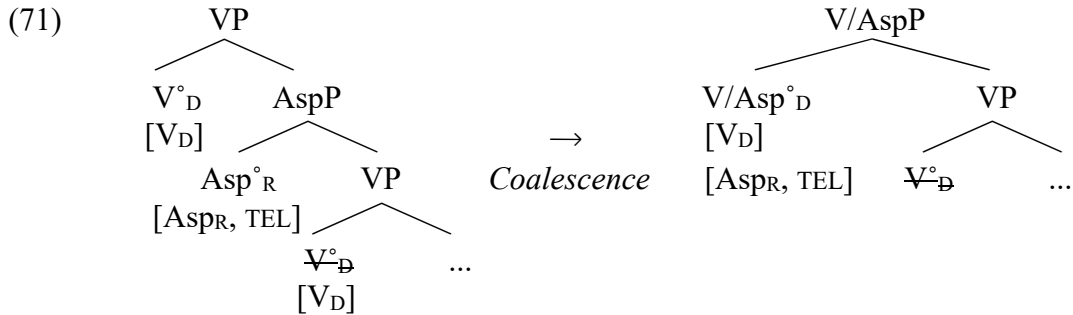


Because T° is obligatorily a suffix, it is proposed that the realization of this structure triggers *go*-support, the insertion of an *anar* auxiliary verb that supports the inflection of both tense and subjecthood.

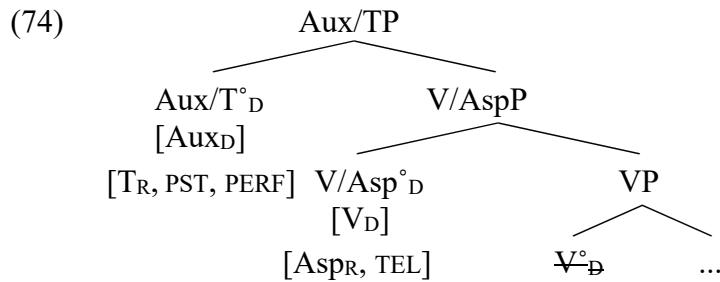
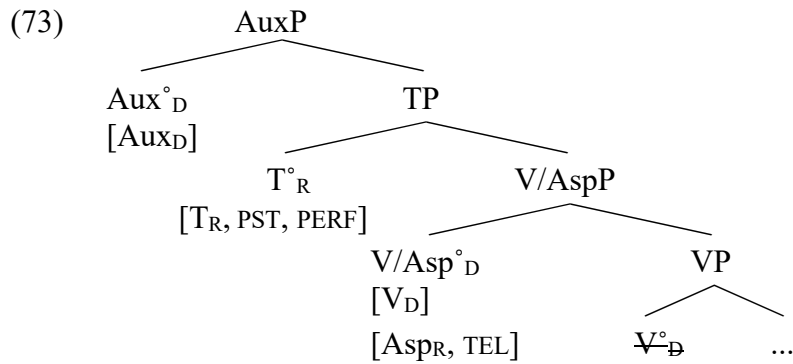
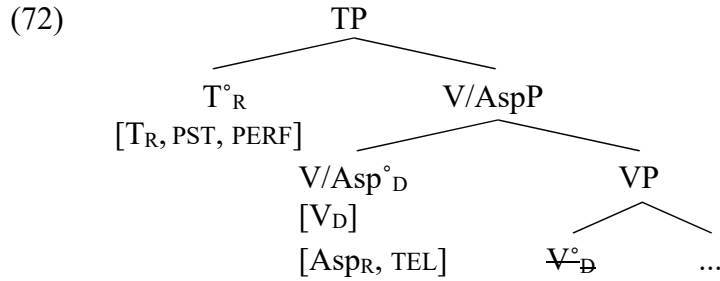
In the present framework, the distinction between the synthetic and analytic forms does not result from differences in the status of category features [Tense] and [Asp], which are recessive in all derivations, but in whether Coalescence is fed by external or internal Merge. In the derivation for the syncretic form,  $V^\circ$  raises to  $T^\circ$ , and is then bundled with the recessive  $T^\circ$  and  $Asp^\circ$  heads.<sup>xii</sup>



Suppose that in these dialects of Catalan, it is also possible for verb movement to proceed only as far as Spec, AspP. At this point, Coalescence bundles  $V^\circ$  and  $Asp^\circ$ .



In the next step, the recessive  $T^{\circ}_R$  head is Merged (72). At this point, a dominant auxiliary head corresponding to *anar* is externally Merged (73). I will remain agnostic as to the categorial status of the auxiliary, labeling it simply as [Aux]. After the auxiliary enters the derivation, Coalescence bundles  $Aux^{\circ}_D$  and  $T^{\circ}_R$  (74).



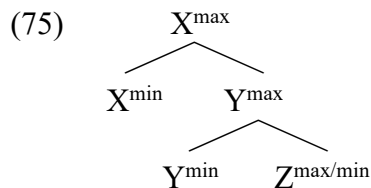
Catalan thus permits two strategies for satisfying the Dominance Condition, depending in this case on the verbal predicate involved. In the synthetic derivation,  $T^{\circ}_R$  is bundled with a dominant verb that moves to its specifier, while the analytic derivation a dominant auxiliary head is externally Merged.<sup>xiii</sup>

## 5 Complex heads as minimal projections

We now return to the claim that Coalescence leaves no internal branching structure within the complex head, accounting for the inability for later syntactic operations to affect subparts of bundled heads. This section discusses other benefits and implications of this proposal. First, it shows how that the elimination of branching structure allows head movement to satisfy the Uniformity Condition (Chomsky 1994). Second, I discuss how Mirror Principle effects can be captured in the absence of head-internal branching structure.

### 5.1 Coalescence and Phrase Structure Uniformity

A key insight of Bare Phrase Structure (Chomsky 1994) is that the phrase structure status of a syntactic object (head vs. phrasal) can be determined uniquely by its position within the syntactic tree. Specifically, an object whose label does not project is a maximal projection (notationally  $XP$  or  $X^{\max}$ ), while an object whose label is not identical to that of a node that it dominates is a minimal projection ( $X^{\circ}$  or  $X^{\min}$ ). An object with both properties is simultaneously maximal and minimal. The three options are schematized in (75).

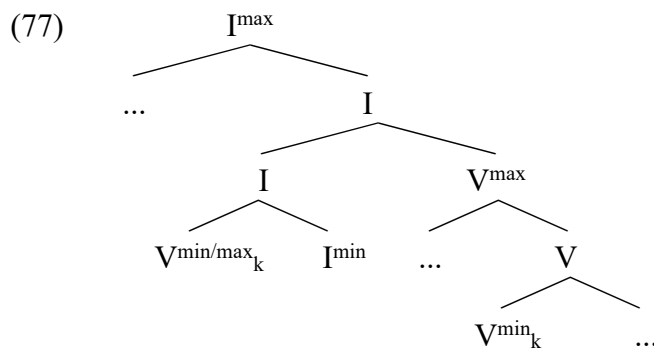


Projections that are neither minimal nor maximal, corresponding to  $X'$  levels in X-bar Theory, are proposed to be inaccessible to syntactic operations and thus cannot undergo movement. To rule out movements deemed to be impossible, such as head-to-specifier

movement or XP adjunction to heads, Chomsky posits a uniformity condition on movement chains:

- (76) UNIFORMITY CONDITION  
A chain is uniform with respect to phrase structure status

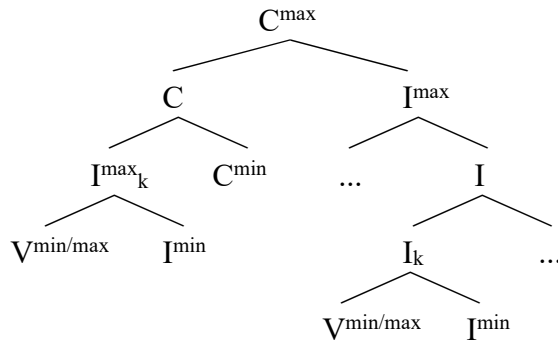
Chomsky observes that the traditional structure associated with head movement violates the Uniformity Condition. Consider the structure that is the predicted result of  $V^\circ$ -to- $\text{Infl}^\circ$  head movement, where  $V^\circ$  adjoins to  $I^\circ$ .



The two copies of V differ in their phrase structure status, in violation of the Uniformity Condition. Because the higher copy of V does not project, it is both minimal and maximal, while the lower copy is minimal because it projects.

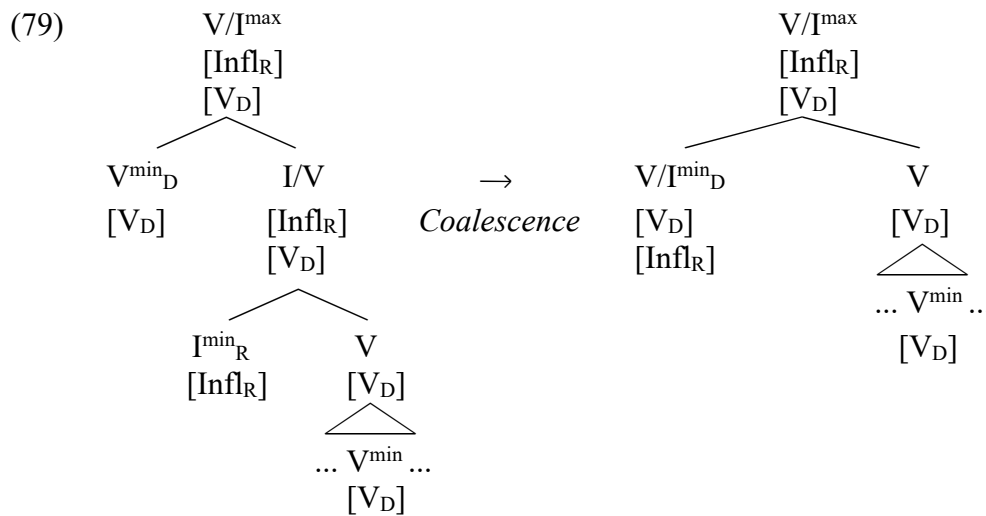
As observed by Harley (2013), successive-cyclic head movement proves even more problematic for the Uniformity Condition. Consider the representation that results from movement and adjunction of the V-Infl head created in (77) to a higher C head.

(78)



The more serious problem in this case is that the lower copy in this structure is neither minimal nor maximal. Since intermediate projections are proposed to be invisible to syntactic operations, movement of the complex head is predicted to be impossible.

Consider how violations of the Uniformity Condition and reference to intermediate projections are avoided in the present proposal. This rests on two claims: First, that root nodes inherit the category features of all lower heads within their extended projection (Category Percolation); Second, that complex heads formed by Coalescence include no internal branching structure. This is shown in the derivation of a bundled V/Infl head. In the first step, dominant  $V^{\min}$  moves to the specifier of  $I^{\max}$ . Because the highest node contains both [V] and [Infl] features in accordance with Category Percolation, it is a maximal projection of both categories ( $I^{\max}$  and  $V^{\max}$ ). This remains unchanged after Coalescence applies. However, the bundled head is now a minimal projection of both [V] and [Infl]. The two links of the movement chain are thus minimal projections of [V], in accordance with the Uniformity Condition, and the bundled head c-commands its lower position.

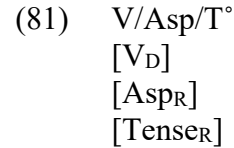
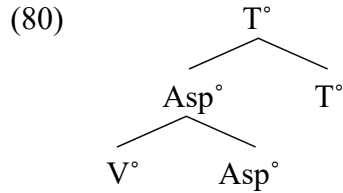


Later head movement of the complex head will similarly produce only minimal projections, and there is no need to posit movement of intermediate-level projections. Coalescence and Category Percolation therefore provide a means to generate head movement within syntax in accordance with Minimalist considerations, namely the Extension and Uniformity Conditions.

## 5.2 Affix ordering

This section considers the implications of Coalescence with respect to theories of affix ordering, particularly in contrast with "classic" views of head adjunction that maintain head-internal branching structure. I argue that the tendency for affix order to reflect the hierarchy of projections can be maintained in the absence of branching structure, so long as c-selection features are accessible to the affix linearization process. In brief, there is no empirical necessity based on patterns of morphological exponence for heads to have internal branching structure in their syntactic representation, lending support to the proposal that featurally complex heads are minimal projections.

As an illustration of a standard analysis of head movement, cyclic head movement of  $V^\circ$  through the  $\text{Aspect}^\circ$  and  $\text{Tense}^\circ$  inflectional heads produces the complex head (80) in accordance with the Right Hand Head Rule (Williams 1981), which places the projecting head at each level on the right.



Taken with the proposal that the association of phonological content with syntactic representations (a.k.a. vocabulary insertion) targets terminal nodes of the tree, the structure derives the Mirror Principle generalization that affix order reflects the hierarchical orderings of their corresponding projections (Baker 1985). Prefixes can be generated by positing a postsyntactic operation Local Dislocation (Embick & Noyer 2001), which alters the linear ordering of sister nodes prior to vocabulary insertion.



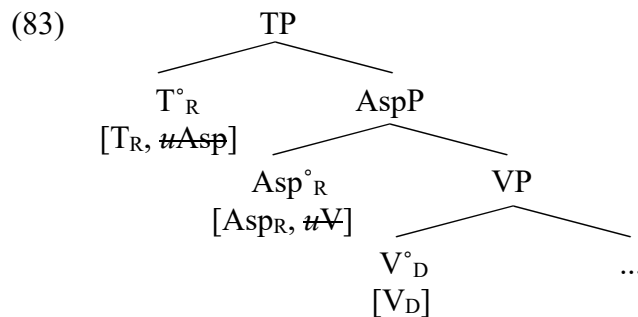
If Local Dislocation applies between sister nodes of complex heads from the bottom up, a  $V/\text{Asp}/\text{T}^\circ$  head can reordered as  $[[V-\text{Asp}]-\text{T}]$ ,  $[[\text{Asp}-V]-\text{T}]$ ,  $[\text{T}-[V-\text{Asp}]]$ ,  $[\text{T}-[\text{Asp}-V]]$ . However, it is not possible to generate orders where the Mirror Principle ordering is violated among the affixes on a single side of the root (Harley 2013), in particular  $*[\text{Asp}-\text{T}-V]$ ,  $*[V-\text{T}-\text{Asp}]$ .

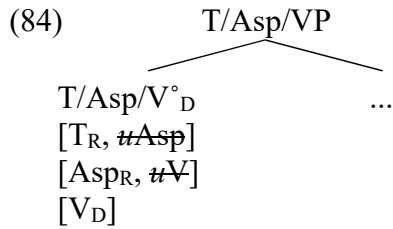
On the other hand, I have proposed that from the same derivation, Coalescence produces a single terminal node that contains the features of each of the bundled heads



(81). While it is possible to formulate vocabulary insertion rules that target bundles of features (Anderson 1982, 1992), this view nonetheless requires an alternative way of predicting the Mirror Principle generalization on affix order, which is no longer directly accessible from the syntactic constituent structure.

The Mirror Principle can be understood as a preference for affix order to reflect the derivational history of a head containing multiple category features. Alternatively put, the order of affixes reflects the order in which their corresponding heads entered the derivation via external Merge. This information is accessible if c-selection features are not immediately deleted from the syntactic representation upon checking, and are accessible to the postsyntactic linearization operation. This conclusion is consistent with the analysis of ‘delayed EPP’ patterns in Section 4, which require checked agreement probes to remain accessible to later stages of the derivation. To illustrate, in the simplified inflectional projection structure [TP [AspP [VP ... ]]] (83), external Merge of each functional head is triggered by c-selectional uninterpretable features [ $uV$ ] on  $Asp^\circ$  and [ $uAsp$ ] on  $T^\circ$ . Subsequent head movement of  $V^\circ$  and Coalescence then create the bundled head in (84).





The Mirror Principle can thus be stated as a preference for linearization to reflect c-selection relations among features of a bundled head, such that the exponent of a feature must be equidistant or closer to the word edge than the exponent of a feature that it c-selects.

It should be noted that this preference on linearization is violable, as languages can place restrictions on the orderings of morphemes that contradict what is expected based on their scopal or derivational properties. A well-known Mirror Principle violation is described for Bantu languages by Hyman (2003) and Good (2003), in which derivational suffixes largely occur in a fixed 'CARP' order. Notably, this results in surface affix orderings that violate the Mirror Principle. In Chichewa, both causativized applicatives and applicativized causatives occur with the same suffix order V-CAUS-APPL, creating a mismatch in scope and linear ordering for causativized applicatives.

(85) *'CARP' template:* Verb stem - Causative - Applicative - Reciprocal - Passive

- (86)
- |   |   |              |
|---|---|--------------|
| <p>a. <b>lil-its-il</b><br/> cry-CAUS-APPL<br/> 'cause to cry with'<br/> [[cause cry] with]</p> | <p>b. <b>takas-its-il</b><br/> stir.with-CAUS-APPL<br/> 'cause to stir with'<br/> [[stir with] cause]</p> | (Hyman 2003) |
|---|---|--------------|

Furthermore, Mirror-Principle preferences on affix ordering can be violated in favor of more apparently phonological restrictions. For instance, suffix order in Pulaar (a.k.a. Fula) generally follows a fixed 'TDNR' order in which more sonorous suffixes occur

further from the root (Arnott 1970; Paster 2005, 2006). While this ordering typically obeys the Mirror Principle by default, and some ordering reversals are permitted to directly reflect semantic scope (Paster 2005, 2006), certain fixed orders do lead to Mirror Principle violations. For instance, the modal suffix *ir*, which introduces a manner/instrumental argument, and the repetitive aspect suffix *it* occur in a fixed order, even if repetitive aspect takes scope over the modal (87b).

- (87) a. o      udd-**it-ir**-ii                      baafal   ŋgal   sawru   wodndu  
           3.SG   close-REPET-MODAL-PST   door   the   stick   different  
           'He closed the door again with a different stick'
- b. mi      udd-**it-ir**-ii                      baafal   ŋgal   sawru  
           3.SG   close-REPET-MODAL-PST   door   the   stick  
           'I closed the door with a stick again.' (same stick)                      (Paster 2005)

Lastly, affix ordering and placement can be sensitive to prosodic properties of stems (Ussishkin 2007; McCarthy and Prince 1993). If prosodic structure is built on the PF branch, this indicates that processes that affect affix ordering apply after syntactic structure undergoes Spell-Out. The generalization that emerges is that affix ordering is determined by a number of competing functional preferences (cf. Manova and Aronoff 2010, Rice 2011 for an overview), including those based on parsimony with syntactic derivations or semantic scope, phonological well-formedness, or morpheme-specific restrictions.

The case studies in this paper have focused on movement and bundling patterns observed within the complementizer and inflectional domains of the clause. Important questions remain in determining whether the approach can be extended to syntactic approaches to derivational morphology, in which acategorical roots must combine with

category-defining functional heads (Marantz 2007; Embick and Marantz 2008). For example, we predict that structures that contain more than one category-defining head like *gloriousness* [*n* [*a* [ $\sqrt{\text{GLORY}}$ ]]] (Embick & Marantz 2008) to be bundled via Coalescence into a single terminal node. It would have to be explained why *gloriousness* must pattern as a noun with respect to later syntactic operations, if it corresponds to a single node containing both [*n*] and [*a*] features. The relative accessibility of features within terminal nodes can potentially be stated in terms of the c-selection relations used to account for affix order, but detailed consideration of the approach will have to await future work.

## 6 Conclusion

In this paper, I have argued that a variety of bundling processes that affect heads should be understood as the result of a single syntactic operation, Coalescence. Although Coalescence presents a non-trivial addition to the set of syntactic operations assumed in standard Minimalism, it permits a unified analysis of Feature Scattering phenomena and head movement, which in prior approaches have been attributed to separate bundling operations in the lexicon and in a postsyntactic module. In particular, I have argued that this accounts for structural similarities between the two patterns, and for derivations in which there is a feeding relationship between head bundling and phrasal movement. Furthermore, application of Coalescence to head movement avoids the primary problems posed by traditional head-adjunction models, namely, violations of the Extension and Uniformity Conditions, and defining a featural trigger for movement.

## References

- ABOH, ENOCH OLADÉ. 2006. Complementation in Saramaccan and Gungbe: The Case of C-type Modal Particles. *Natural Language & Linguistic Theory* 24.1–55.
- ANDERSON, STEPHEN R. 1982. Where's Morphology? *Linguistic Inquiry* 13.571–612.
- ANDERSON, STEPHEN R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- ANDERSON, STEPHEN R. 2008. English Reduced Auxiliaries Really are Simple Clitics. *Lingue e Linguaggio* 7.169–186.
- ARNOTT, D. W. 1970. *The nominal and verbal systems of Fula*. Oxford: Clarendon Press.
- BAKER, MARK. 1985. The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry* 16.373–415.
- BAKER, MARK. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- BENINCÀ, PAOLA.; and CECILIA POLETO. 2004. Topic, Focus, and V2. *The Structure of CP and IP*, ed. by Luigi Rizzi, 52–75. Oxford: Oxford University Press.
- BHATT, RAKESH. 1999. *Verb movement and the syntax of Kashmiri*. Dordrecht: Kluwer.
- BOBALJIK, JONATHAN DAVID. 1995. Morphosyntax: The syntax of verbal inflection. Massachusetts Institute of Technology.
- BOBALJIK, JONATHAN DAVID.; and HÖSKULDUR THRÁINSSON. 1998. Two heads aren't always better than one. *Syntax* 1.37–71.
- BOECKX, CEDRIC.; and SANDRA STJEPANOVIC. 2001. Head-ing toward PF. *Linguistic Inquiry* 32.345–355.
- BORER, HAGIT. 1984. *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht: Foris.
- BORSLEY, ROBERT D.; MARIA-LUISA RIVERO.; and JANIG STEPHENS. 1996. Long head movement in Breton. *The syntax of the Celtic languages: a comparative perspective*, ed. by Robert Borsley and Ian G. Roberts, 53–74. Cambridge: Cambridge University Press.
- BOŠKOVIĆ, ŽELJKO. 1995. Participle Movement and Second Position Cliticization in Serbo-Croatian. *Lingua* 96.245–266.
- CHOMSKY, NOAM. 1993. A minimalist program for linguistic theory. *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, MA: MIT Press.
- CHOMSKY, NOAM. 1994. Bare Phrase Structure. *MIT Occasional Papers in Linguistics* 5.
- CHOMSKY, NOAM. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- CHOMSKY, NOAM. 2000. Minimalist inquiries: the framework. *Step by Step: essays on Minimalist syntax in honour of Howard Lasnik*, ed. by Roger Martin, David Michaels, Juan Uriagereka, and Samuel Jay Keyser, 89–115. Cambridge, MA: MIT Press.
- CHOMSKY, NOAM. 2008. On phases. *Foundational Issues in Linguistic Theory*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.
- CINQUE, GUGLIELMO. 1999. *Adverbs and functional heads: a cross-linguistic perspective*.

- Oxford: Oxford University Press.
- CINQUE, GUGLIELMO.; and LUIGI RIZZI. 2008. The Cartography of Syntactic Structures. *CISCL Working Papers on Language and Cognition 2*, ed. by V. Moscati, 43–59. Siena: University of Siena.
- CINQUE, GUGLIELMO.; and LUIGI RIZZI. 2009. The Cartography of Syntactic Structures. *The Oxford Handbook of Linguistic Analysis*, ed. by Bernd Heine and Heiko Narrog, 1–21. Oxford: Oxford University Press.
- COWPER, ELIZABETH. 2005. The geometry of interpretable features: Infl in English and Spanish. *Language* 81.10–46.
- DEN BESTEN, HANS. 1983. On the interaction of root transformations and lexical deletive rules. *On the formal syntax of the Westgermania*, ed. by Werner Abraham, 47–131. Amsterdam: John Benjamins.
- DEN DIKKEN, MARCEL. 2007. Phase extension contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33.1–41.
- DI SCIULLO, ANNA MARIA.; and DANIELA ISAC. 2008. The Asymmetry of Merge. *Biolinguistics* 2.260–290.
- DIESING, MOLLY. 2005. The Upper Functional Domain in Yiddish. *Focus on Germanic Typology*, ed. by Werner Abraham, 195–209. Berlin: Akademierverlag.
- DOUGLAS, JAMIE. 2017. Unifying the that-trace and anti-that-trace effects. *Glossa: a journal of general linguistics* 2.60. 1–28.
- EMBICK, DAVID.; and ROUMYANA IZVORSKI. 1995. On long head movement in Bulgarian. *ESCOL '94*, ed. by Janet M. Fuller, Ho Han, and David Parkinson, 104–115. Ithaca, NY: CLC Publications.
- EMBICK, DAVID.; and ALEC MARANTZ. 2008. Architecture and Blocking. *Linguistic Inquiry* 39.1–53.
- EMBICK, DAVID.; and ROLF NOYER. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32.555–595.
- EMONDS, JOSEPH E. 1970. Root and structure-preserving transformations. Massachusetts Institute of Technology.
- ERLEWINE, MICHAEL YOSHITAKA. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34.429–479.
- FANSELOW, GISBERT. 2003. Münchhausen-style head movement and the analysis of verb second. *UCLA Working Papers in Linguistics*.40–76.
- FANSELOW, GISBERT. 2009. Bootstrapping verb movement and the clausal architecture of German (and other languages). *Advances in comparative Germanic syntax*, ed. by Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger, and Florian Schäfer, 85–118. Amsterdam/Philadelphia: John Benjamins.
- FANSELOW, GISBERT.; and DENISA LENERTOVÁ. 2010. Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory* 29.169–209.
- GALLEGO, ÁNGEL J. 2006. Phase sliding. Ms., Universitat Autònoma de Barcelona and University of Maryland.
- GALLEGO, ÁNGEL J. 2010. *Phase Theory*. Amsterdam/Philadelphia: John Benjamins.
- GEORGI, DOREEN.; and GEREON MÜLLER. 2010. Noun phrase structure by reprojection.

*Syntax* 13.1–36.

- GIORGI, ALESSANDRA.; and FABIO PIANESI. 1997a. The syntactic and semantic background. *Tense and Aspect : From Semantics to Morphosyntax*. Cary, NC: Oxford University Press, USA.
- GIORGI, ALESSANDRA.; and FABIO PIANESI. 1997b. *Tense and Aspect: From Semantics to Morphosyntax*. New York: Oxford University Press.
- GIUSTI, GIULIANA. 2002. The functional structure of noun phrases: A bare phrase structure approach. *Functional structure in DP and IP: The cartography of syntactic structures, Volume 1*, ed. by Guglielmo Cinque, 54–90. Oxford: Oxford University Press.
- GOLDSMITH, JOHN. 1990. *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- GOOD, JEFFREY C. 2003. Strong linearity: Three cases studies towards a theory of morphosyntactic templatic constructions. University of California, Berkeley.
- GRIMSHAW, JANE. 1991. Extended projections. Brandeis University.
- GRIMSHAW, JANE. 2000. Extended projection and locality. *Lexical Specification and Insertion*, ed. by Peter Coopmans, Martin Everaert, and Jane Grimshaw, 115–133. Amsterdam: John Benjamins.
- HAIDER, HUBERT. 1988. Matching Projections. *Constituent Structure: Papers from the 1987 GLOW Conference*, ed. by Anna Cardinaletti, Guglielmo Cinque, and Giuliana Giusti, 101–121. Dordrecht: Foris.
- HARIZANOV, BORIS. 2014. Clitic doubling at the syntax-morphophonology interface. *Natural Language & Linguistic Theory* 32.1033–1088.
- HARLEY, HEIDI. 2013. Getting morphemes in order: Merger, Affixation and Head Movement. *Diagnosing Syntax*, ed. by Lisa L Cheng and Norbert Corver, 44–74. Oxford: Oxford University Press.
- HARTMAN, JEREMY. 2011. The Semantic Uniformity of Traces: Evidence from Ellipsis Parallelism. *Linguistic Inquiry* 42.367–388.
- HÖHN, GEORG F.K. 2016. Unagreement is an illusion: Apparent person mismatches and nominal structure. *Natural Language and Linguistic Theory* 34.
- HSU, BRIAN. 2017. Verb second and its deviations: An argument for Feature Scattering in the left periphery. *Glossa: a journal of general linguistics* 2.35. 1-3.
- HYMAN, LARRY M. 2003. Suffix Ordering in Bantu: A Morphocentric Approach. *Yearbook of Morphology 2002*, ed. by Geert Booij and Jaap van Marle, 245–281. Dordrecht: Springer.
- IATRIDOU, SABINE. 1990. About Agr(P). *Linguistic Inquiry* 21.551–577.
- ITÔ, JUNKO. 1988. *Syllable Theory in Prosodic Phonology*. New York: Garland Publishing.
- JULIEN, MARIT. 2002. *Syntactic Heads and Word Formation: A Study of Verbal Inflection*. New York: Oxford University Press, USA.
- JULIEN, MARIT. 2005. *Nominal phrases from a Scandinavian perspective*. Amsterdam: John Benjamins.
- KANDYBOWICZ, JASON. 2009. Embracing Edges : Syntactic and Phono-Syntactic Edge Sensitivity in Nupe. *Natural Language & Linguistic Theory* 27.305–344.
- KAYNE, RICHARD. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

- KAYNE, RICHARD S. 2005. Some Notes on Comparative Syntax, with Special Reference to English and French. *The Oxford Handbook of Comparative Syntax*, ed. by Guglielmo Cinque and Richard S. Kayne. New York: Oxford University Press.
- KEINE, STEFAN. 2015. Selective opacity. Ms., University of Massachusetts Amherst.
- KLIMA, EDWARD S. 1964. Negation in English. *The Structure of Language: Readings in the Philosophy of Language*, ed. by Jerry A. Fodor and Jerrold J. Katz, 246–323. Englewood Cliffs, NJ: Prentice Hall.
- KOENEMAN, OLAF.; and HEDDE ZEIJLSTRA. 2014. The Rich Agreement Hypothesis Rehabilitated. *Linguistic Inquiry* 45.571–615.
- KOOPMAN, HILDA. 1996. The Spec Head configuration. *Syntax at Sunset, UCLA Working papers in syntax and semantics Volume 1*, ed. by Edward Garret and Felicia Lee, 37–64. Los Angeles: University of California, Los Angeles.
- LECHNER, WINFRIED. 2006. An interpretive effect of Head Movement. *Phases of Interpretation*, ed. by Mara Frascarelli, 45–71. Berlin: Mouton de Gruyter.
- LEDGEWAY, ADAM. 2010. Subject licensing in CP. *Mapping the Left Periphery*, ed. by Paola Benincà and Nicola Munaro, 257–296. Oxford: Oxford University Press.
- LEMA, JOSÉ.; and MARÍA LUISA RIVERO. 1989. Long head movement: ECP vs. HMC. *Proceedings of NELS 20*, 333–347. Amherst, MA: GLSA.
- MAHAJAN, ANOOP KUMAR. 2001. *Word order and remnant VP movement*.
- MANETTA, EMILY. 2011. *Peripheries in Kashmiri and Hindi-Urdu: The syntax of discourse-driven movement*. Amsterdam/Philadelphia: John Benjamins.
- MANOVA, STELA.; and MARK ARONOFF. 2010. Modeling affix order. *Morphology* 20.109–131.
- MARANTZ, ALEC. 2007. Phases and words. *Phases in the theory of grammar*, ed. by S. H. Choe, 191–222. Seoul: Dong-In Publishing Co.
- MATUSHANSKY, ORA. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37.69–109.
- MATYIKU, SABINA MARIA. 2017. Semantic effects of head movement : Evidence from negative auxiliary inversion. Yale University.
- MCCARTHY, JOHN J.; and ALAN PRINCE. 1993. Generalized alignment. *Yearbook of Morphology*. 79–153.
- MIGDALSKI, KRZYSZTOF MAREK. 2006. The syntax of compound tenses in Slavic. Tilburg University.
- MOHR, SABINE. 2009. V2 as a Single-Edge Phenomenon. *Selected Papers from the 2006 Cyprus syntaxfest*, ed. by Kleanthes K. Grohmann and Phoevos Panagiotidis, 141–159. Cambridge: Cambridge Scholars Publishing.
- MUNSHI, SADAF.; and RAJESH BHATT. 2009. Two locations for negation: Evidence from Kashmiri. *Linguistic Variation Yearbook* 9.205–205–239.
- OLTRA-MASSUET, ISABEL. 2013. Variability and allomorphy in the morphosyntax of Catalan past perfective. *Distributed Morphology Today: Morphemes for Morris Halle*, ed. by Ora Matushansky and Alec Marantz, 1–20. Cambridge, MA: MIT Press.
- OUHALLA, JAMAL. 1991. *Functional categories and parametric variation*. London: Routledge.



- PANAGIOTIDIS, PHOEVOS. 2014. *Categorial features: a generative theory of word class categories*. Cambridge: Cambridge University Press.
- PASTER, MARY. 2005. Pulaar verbal extensions and phonologically driven affix ordering. *Yearbook of Morphology* 2005.155–199.
- PASTER, MARY ELIZABETH. 2006. *Phonological Conditions on Affixation*. University of California, Berkeley.
- PESETSKY, DAVID.; and ESTHER TORREGO. 2001. T-to-C Movement : Causes and Consequences. *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 355–426. Cambridge, MA: MIT Press.
- PLATZACK, CHRISTER. 2013. Head Movement as a Phonological Operation. *Diagnosing Syntax*, ed. by Lisa Lai-Shen Cheng and Norbert Corver, 21–43. Oxford: Oxford University Press.
- POLETTI, CECILIA. 2000. *The Higher Functional Field*. New York: Oxford University Press, USA.
- POLETTI, CECILIA. 2002. The left-periphery of V2-Rhaetoromance dialects: a new view on V2 and V3. *Syntactic Microvariation*, ed. by Sjef Barbiers, Leonie Cornips, and Susanne van der Kleij, 214–242. Amsterdam: Meertens.
- POLETTI, CECILIA. 2014. *Word Order in Old Italian*. Oxford: Oxford University Press.
- POLLOCK, JEAN-YVES. 1989. Verb movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20.365–424.
- PYLKKÄNEN, LIINA. 2002. *Introducing arguments*. Ph.D dissertation, Massachusetts Institute of Technology.
- RICE, KEREN. 2011. Principles of affix ordering: An overview. *Word Structure* 4.169–200.
- RICHARDS, NORVIN. 2016. *Contiguity Theory*. Cambridge, MA: MIT Press.
- RIVERO, MARIA-LUISA. 1991. Long Head Movement and Negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review* 8.319–351.
- RIZZI, LUIGI. 1997. The fine structure of the left periphery. *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- ROBERTS, IAN G. 2005. *Principles and Parameters in a VSO Language: A Case Study in Welsh*. Oxford: Oxford University Press.
- ROBERTS, IAN G. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.
- SPEAS, MARGARET. 1991. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- SPORTICHE, DOMINIQUE. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19.425–449.
- STERIADE, DONCA. 1995. Underspecification and markedness. *The Handbook of Phonological Theory*, ed. by John Goldsmith, 114–174. Oxford: Blackwell.
- SURÁNYI, BALÁZS. 2005. Head movement and reprojection. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. Tomus XXVI*, 313–342. Budapest: ELTE.
- SVENONIUS, PETER. 1994. C-selection as feature checking. *Studia Linguistica* 48.133–155.
- THRÁINSSON, HÖSKULDUR. 1996. On the (non)-universality of functional projections.

- Minimal ideas*, ed. by Werner Abraham, Samuel David Epstein, Hoskuldur Thrainsson, and C. Jan-Wouter Zwart, 253–81. Amsterdam: John Benjamins.
- TRAVIS, LISA. 1984. Parameters and Effects of Word Order Variation. Ph.D dissertation, Massachusetts Institute of Technology.
- USSISHKIN, ADAM. 2007. Morpheme position. *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 457–472. Cambridge: Cambridge University Press.
- VAN RIEMSDIJK, HENK. 1988. The representation of syntactic categories. *Proceedings of the Conference on the Basque Language, Basque World Congress* 1.104–116.
- VAN RIEMSDIJK, HENK. 1998. Categorical feature magnetism: The endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2.1–48.
- VANGSNES, ØYSTEIN ALEXANDER. 1999. ‘Identification’ and the role of morphology in the Scandinavian noun phrase. 99–176.
- VIKNER, STEN. 1995. *Verb movement and expletive subjects in the Germanic languages*. New York and Oxford: Oxford University Press.
- WALKDEN, GEORGE. 2017. Language contact and V3 in Germanic varieties new and old. *Journal of Comparative Germanic Linguistics* 20. Springer Netherlands. 49–81.
- WALKER, RACHEL. 2011. *Vowel patterns in language*. New York: Cambridge University Press.
- WILDER, CHRIS.; and DAMIR ČAVAR. 1994a. Word Order Variation, Verb Movement, and Economy Principles. *Studia Linguistica* 48.46–86.
- WILDER, CHRIS.; and DAMIR ČAVAR. 1994b. Long Head Movement? Verb Movement and Cliticization in Croatian. *Lingua* 93.1–58.
- WILLIAMS, EDWIN. 1981. On the notions “Lexically related” and “Head of a word”. *Linguistic Inquiry* 12.245–274.
- WOLFE, SAM. On Microvariation in V2 Systems in Old Romance and Beyond. *Linguistic Variation*.
- ZWICKY, ARNOLD M. 1970. Auxiliary Reduction in English. *Linguistic Inquiry* 1.323–336.
- ZWICKY, ARNOLD M.; and GEOFFREY K. PULLUM. 1983. Cliticization vs. Inflection: English N’T. *Language* 59.502–513.

---

<sup>i</sup> This structural characterization of movement resembles the one posited in reprojection approaches to head movement (Fanselow 2003; Surányi 2005; Georgi and Müller 2010). The moved head then projects its label to the root node, thus satisfying the Uniformity Condition. In Section 5, I show that reprojection is obviated if Category Percolation (Keine 2015) is adopted.

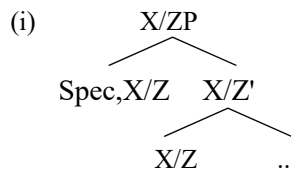
<sup>ii</sup> The observation here resembles the claim in other works that functional projections are only licensed if they have a filled specifier or an overt head (Koopman 1996; Vangsnes 1999; Giusti 2002; Julien 2005). However, in Feature Scattering these criteria do not affect whether functional category features are present

---

in extended projections (assumed to always carry the same features), but in how they are mapped to phrase structure.

iii Given this structural definition, which precludes Coalescence from applying to two heads with an intervening specifier, this environment often resembles that of linear adjacency. The two structural definitions would differ in a context in which a head-initial projection dominated a head-final one with no specifier,  $[_{\text{ZP}} Z [_{\text{YP}} [_{\text{XP}} \dots ] Y]$ , satisfying structural but not linear adjacency. While I know of no instances of bundling that take place in this configuration, this could be explained if true head-final structures are excluded from syntactic representations (Kayne 1994).

iv The output of Coalescence is identical to that of Matching Projections in the theory of Haider (1988). However, the contexts in which it applies are not identical. Haider proposes that functional projections whose heads have no phonetic realization are superimposed onto another projection within the same extended projection. Unlike Haider's proposal, I propose that projections with null heads are licensed if they have an overt specifier.



v I thank an anonymous reviewer for pointing out the relevance of such examples to the discussion.

vi In the formulation of Chomsky (2008), Merge of syntactic objects X and Y leaves both unchanged. Although Coalescence is not a variety of Merge, one might consider this as a broader restriction on syntactic operations.

vii The approach resembles the grammar of Chomsky (2001) in which the [EPP] feature is distinct from [uF] probes on a lexical item, and Agree only triggers movement if the probing item contains [EPP]. The present proposal differs from Chomsky in that [EPP] is not 'satisfied' or checked once a specifier is Merged.

viii A similar analysis appears to be necessary in the analysis of verb placement in standard dialects of Danish, Norwegian, and Swedish, in which root clauses show verb-second order derived by V-to-C

---

movement, while embedded contexts show an English-like pattern where verbs remain in VP and subjects occur Spec, TP (Vikner 1995). Tentatively, we can posit that in these languages, null Aux<sub>D</sub> is licensed uniquely in embedded clauses. The nature of this licensing restriction, however, remains to be explained.

<sup>ix</sup> For arguments that a low C-domain feature like Finiteness are associated with subject probes, see Poletto (2000) Aboh (2006), Ledgeway (2010).

<sup>x</sup> In a similar vein, Leu (2015) argues that the overt complementizer vs. verb-second alternation reflects two strategies used to activate the highest clausal projection in German. In brief, the projection can be activated either by an “-*ass*” head that attracts a lower complementizer head corresponding to “*d-*”, or filled by movement of a remnant vP.

<sup>xi</sup> The reduced forms of *will*, *have*, *are*, and *am* are substantially more restricted, appearing only after pronouns and *wh*-words.

<sup>xii</sup> I do not exclude the possibility that V° moves first through AspP before raising to TP. However, this is not critical for the purposes of this discussion, since the two possible derivations will result in the same bundled head structure.

<sup>xiii</sup> At a broader level, the pattern indicates that the availability of head movement can depend on properties of the moving head other than dominance vs. recession or the presence of [EPP]. Similarly, it has been proposed that the “movability” of a head can be predicted from its prosodic properties (Richards 2016) or aspects of its morphological exponence, as proposed in the Rich Agreement Hypothesis (c.f. Koenenman and Zeijlstra (2014) for a recent overview). At the moment, however, I will have to leave it an open question how these factors should be integrated in the syntactic representations. I thank an anonymous reviewer for bringing the question to attention.