

Whither Head Movement?

Boris Harizanov and Vera Gribanova

Received: date / Accepted: date

Abstract We argue that head movement, as an operation that builds head-adjunction structures in the syntax, has been used to model two empirically distinct classes of phenomena. One class has to do with displacement of heads (fully formed morphological words) to higher syntactic positions, and includes phenomena like verb second and verb initial-ity. The other class has to do with the construction of complex morphological words and is involved in various types of word formation. Based on the very different clusters of properties associated with these two classes of phenomena, we argue that they each should be accounted for by distinct grammatical operations, applying in distinct modules of the grammar, rather than by the one traditional syntactic head movement operation. We propose that the operation responsible for upward displacement of heads is genuine syntactic movement (Internal Merge) and has the properties of syntactic phrasal movement, including the ability to affect word order, the potential to give rise to interpretive effects, and the locality associated with Internal Merge. On the other hand, word formation is the result of postsyn- tactic amalgamation, realized as either *Lowering* (Embick and Noyer, 2001) or its upward counterpart, *Raising*. This operation, we argue, has properties that are not associated with narrow syntax: it is morphologically driven, it results in word formation, it does not exhibit interpretive effects, and it has stricter locality conditions (the Head Movement Constraint). The result is a view of head movement that not only accounts for the empirical differences between the two classes of head movement phenomena, but also lays to rest numerous perennial theoretical problems that have heretofore been associated with the syntactic head adjunction view of head movement. In addition, the framework developed here yields interesting new predictions with respect to the expected typology of head movement patterns.

1 Introduction

Head movement, conceived of as an operation that builds head-adjunction structures in the syntax, has been exceed- ingly useful over the last few decades in deepening our understanding of a range of empirical phenomena, including verb second effects, verb-initial clause structures, and lexical decomposition, among others. Despite these fruitful re- sults, with the advent of the Minimalist turn, head movement became difficult to integrate theoretically, at exactly the time when (phrasal) movement and base generation were being productively unified through the Merge conception of structure building. The result is a tension between descriptive usefulness on the one hand and insightful, elegant theoretical integration on the other. This tension has put head movement at the center of controversy, not least because its status with respect to the overall architecture of the grammar has become less clear over time (Chomsky, 2000).

*Harizanov and Gribanova are co-first authors of this paper.

The goal of this paper is to resolve this tension by addressing the theoretical issues with the traditional view of head movement while preserving the descriptive successes of that view. This now seems possible because, as case studies of head movement phenomena have accumulated, it has become ever clearer that they can be successfully factored into two categories—those that should be handled by syntactic operations and those that should be handled by postsyntactic operations. Our goal is to distinguish, both theoretically and empirically, between cases of true head movement, in which a fully formed word is moved to another structural position (via movement in the narrow syntax), and cases of word formation, in which syntactically independent pieces are unified morphologically (via postsyntactic operations). This move has beneficial consequences for our understanding of the diverse properties of phenomena we associate with head movement. By making the right cut between syntactic and postsyntactic operations, we make clear predictions about which head movement phenomena should exhibit the properties of syntactic movement (Internal Merge), and which should exhibit the properties of postsyntactic operations. The result, we argue, is a more granular and refined model of syntax-morphology interactions which opens up new and productive avenues of investigation into the nature of word formation and its interaction with genuinely syntactic processes.

Our discussion owes an intellectual debt to two existing proposals in the head movement literature, which opened the doors for the present inquiries in significant ways. The first was the insight of Rizzi and Roberts (1989), who suggested that head adjunction and substitution involved significantly different representations and operations. Head adjunction involved a kind of morphological selection by the host of the incorporation, resulting in the projection of an additional “slot” as the sister of the selecting head in D-structure. Substitution, by contrast, involved no such slot, resulting in the familiar lexical blocking effects: T-to-C movement is blocked in French if C is lexically realized (Rizzi and Roberts 1989: 17-18). Even though the current proposal is significantly different in its implementation and in its architectural claims, we share with Rizzi and Roberts the core idea that there are two types of head movement and that they should be distinguished theoretically. With the advent of the Minimalist motivation to understand better what operations are genuinely syntactic and which are interface operations, we now have the formal language to describe this distinction in terms of interacting grammatical modules (rather than as instances of different syntactic operations, as in Rizzi and Roberts 1989).

The second proposal, going chronologically, is that of Matushansky (2006), who was the first (to our knowledge) to propose that head movement configurations may involve two operations: in her terms, syntactic movement to a specifier position and a morphological m-merger process. We share with this proposal the idea that understanding of head movement phenomena requires operations in two different grammatical modules. Matushansky’s (2006) claim, however, is that all canonical cases head of movement—including, for example, V-to-T in French—involve application of *both* operations. Our proposal, by contrast, separates the empirical properties of head movement types into two groups, one of which has the properties of syntactic movement, and one of which has the properties of what is essentially a morphological process. Thus, for example, French V-to-T is an instantiation of the latter type of head movement and, as described in later sections, involves no syntactic movement whatsoever.¹

The basic idea behind head movement—that a head can move to another head position—first emerged in a unified way in the works of Travis (1984), Koopman (1984), and Baker (1985, 1988) (though see Aissen 1974 for an earlier precursor). Taken together, these works generalized head movement as a type of head-adjunction operation that could be seen as applying in numerous empirical domains. This included so-called “roll-up” head movement, which builds complex morphological words (including grammatical function-changing morphology, as in causatives, reflexives, applicatives, etc.); incorporation, which involves the formation of various kinds of compounds; and substitution, in which one head moves to another position, yielding changes in word order but not in morphological structure. Despite their diverse empirical properties, such constructions were taken to have certain general properties in common—including adherence to locality constraints specific to the movement of heads—which justified the unified approach to their theoretical understanding.

This unified vision has gradually unraveled, for both theoretical and empirical reasons. On the theoretical side, certain properties of the conception of structure building based on Merge, which subsumes base generation and movement, forced a change in the way head movement is viewed. This resulted in the well known objections first voiced by Chomsky (2001) and later discussed by Matushansky (2006) and Roberts (2010), among others. Throughout, the

¹ Although Matushansky (2006) presents arguments for the separability of the two operations (m-merger without syntactic movement and syntactic movement without m-merger), the cases in which this holds do not align with the distinction we advocate here—namely, the separation between head displacement and word formation, each with its own set of characteristic properties.

comparison has been with the fairly well established properties of phrasal movement. Briefly summarized, some of the major concerns with head movement are as follows. First, there are issues relating to the landing site of head movement: (i) head movement qua head-to-head adjunction is countercyclic in that it violates the Extension Condition, which requires any instance of Merge to apply to the root;² and (ii) head movement is not “upward” in that the landing site of the head-moved element does not c-command its origin, violating a basic principle of (phrasal) movement. In addition, there are issues relating to the nature of the movement itself: (iii) the ban on excorporation in head movement results in roll-up head movement (i.e., obligatory pied piping of all head-adjoined material). This differs crucially from the successive cyclic movement we find in the context of phrasal movement, where no further pied piping takes place (see Roberts 1991 for an alternative view of the ban on excorporation). (iv) The locality conditions on head movement are also thought to be more restrictive when compared to those on phrasal movement, as reflected by the standard definition of the Head Movement Constraint (Travis, 1984).³ Furthermore, (v) few, if any, cases of head movement have convincingly discernible interpretive effects, as do certain types of phrasal movement (though see Roberts 2010 for further discussion and potential examples of interpretive effects of head movement). Finally, (vi) what the featural trigger for head movement is within the traditional understanding seems unclear.

These objections have yielded numerous disparate attempts to reduce head movement to other, arguably independently needed mechanisms. These include, but are not limited to, remnant movement (Koopman and Szabolcsi, 2000), PF movement (Chomsky, 2001; Schoorlemmer and Temmerman, 2012; Platzack, 2013), reprojection (Georgi and Müller, 2010), and phrasal movement plus morphological rebracketing (Matushansky, 2006; Harizanov, 2014). This theorizing, in turn, drove empirical investigations into phenomena typically accounted for by head movement, which have revealed that these phenomena are often characterized by quite diverse properties. For example, as Harizanov (2014) demonstrates, Bulgarian pronominal clitics, though morphologically attached to the verbal complex, behave like phrasal units as far as the syntax is concerned, e.g., for the purposes of binding and quantifier stranding. This suggests that at least some phenomena previously associated with head movement analyses involve syntactic units that are simultaneously minimal and maximal, in the sense of Bare Phrase Structure (Chomsky 1995b). Yet, Roberts (2010) treats clitic clusters in Romance as arising from agreement mechanisms between the heads themselves, eschewing a phrasal approach. Studies of long head movement, in which a head appears to move past at least one other head, appear to indicate that the Head Movement Constraint is violable at least in some instances of head movement (see Roberts 1994 and, on Slavic in particular, Rivero 1994, Wilder and Cavar 1994, Embick and Izvorski 1997a, Harizanov 2016, among others). Likewise, studies of predicate fronting involving the movement of just heads (rather than phrases) seem to suggest that certain head movements are possible across clause boundaries (Landau, 2006; Vicente, 2007; Harizanov, 2016). Here, we aim to explain the diverse range of patterns associated with head movement by arguing that there are two mechanisms at play in deriving them.

In what follows, we provide a collection of arguments in favor of a theoretical split between a postsyntactic word-formation operation we call *amalgamation*, which can be instantiated as either raising or lowering of a head, and movement of a head in the narrow syntax (§2). These arguments come from observations about clusterings of properties that seem to be characteristic of syntactic movement vs. postsyntactic operations. We follow this up in §3 with an implementation of the postsyntactic amalgamation operation, proposing a *Raising* counterpart to the Distributed Morphology operation *Lowering*. We then discuss how the genuinely syntactic type of head movement is predicted to interact with postsyntactic amalgamation (§4). In particular, we note that syntactic head movement is expected to derivationally precede postsyntactic amalgamation, and explore the intricate consequences of this view, demonstrating that various predictions of the proposed system are borne out. As part of this effort, we advance an alternative understanding of phenomena that have traditionally been understood as involving word formation feeding syntactic head movement (such as verb second in German and Danish).

The main focus of the paper concerns two connected ideas: first, that various properties systematically cluster together in two distinct groups, and second, that one cluster of properties is not the kind that is typically associated with genuinely syntactic movement, but rather with postsyntactic operations (see also footnote 5). We take this to justify a split between postsyntactic and syntactic operations that have been heretofore treated as identical. Although

² This can be viewed as a violation of the No Tampering Condition (Chomsky, 2008), according to which Merge cannot change the objects it applies to, neither breaking them up nor adding features to them.

³ Although there have been attempts to reduce the Head Movement Constraint to general constraints on (e.g., phrasal) movement (Rizzi, 1990), the empirical effect of the Head Movement Constraint is nonetheless limited to heads. See section 2.3 for further discussion.

we provide an implementation of both operations in the latter sections of the paper, there is a clear sense in which working out various extensions of this alternative theoretical landscape raises crucial questions which we cannot hope to resolve here, and must leave for future work. In addition, the claims introduced here are not meant to preclude the existence of other postsyntactic operations which also participate in word formation or morpheme order manipulations (e.g., Merger Under Adjacency (Bobaljik, 2002), Local Dislocation (Embick and Noyer, 2001), etc.). Working out the complete array of predictions regarding the interaction between the mechanisms proposed here and other independently motivated Distributed Morphology operations is likewise an important task, but not the focus of the present work.

2 Dissecting head movement

The main claim we pursue in this paper is that the phenomena traditionally accounted for via head movement, as traditionally understood, should be divided in two classes, explained by two distinct operations—one is syntactic movement (Internal Merge), while the other is postsyntactic (morphological) amalgamation, involving no movement in the narrow syntax.⁴ We elaborate on each these distinct operations, as well as on their interaction, in §3 and §4. In these sections, we implement amalgamation in terms of two postsyntactic operations: *Lowering* (Embick and Noyer, 2001) and its upward counterpart, *Raising*. We take genuinely syntactic head movement, by contrast, to involve Internal Merge of a head with a phrase, which according to the Extension Condition must be the root (Fukui and Takano, 1998; Toyoshima, 2001; Matushansky, 2006; Koenenman, 2000; Fanselow, 2003; Donati, 2006).

Our core hypothesis is that the effects associated with manipulations of word order and with word formation are logically independent of each other, both in the sense that they have distinct empirical properties, and in the sense that analytical tools of different types (and in different modules of grammar) are at play. This hypothesis leads us to expect an empirical landscape in which (i) genuine syntactic movement (with all its trademark properties) of heads may be identified independently of morphological amalgamation, and (ii) amalgamation is attested in the absence of the characteristic properties of genuine syntactic movement (though possibly with the surface appearance of displacement). If this hypothesis is on the right track, it will enable us to understand better the sometimes contradictory mix of empirical and theoretical properties we see attributed to head movement in recent discussions.⁵

The goal of this section is to identify two sets of properties associated with head movement as it is traditionally construed. These properties systematically cluster in two distinct groups, suggesting a postsyntactic and a syntactic source for each of the groups (see Table 1). Amalgamation, being a postsyntactic operation, involves the formation of words out of independent morphosyntactic terminals. As such, it has properties that differentiate it from genuine syntactic movement. First, while amalgamation results in the formation of a morphologically complex unit, it does not necessarily lead to upward displacement of a head, as the traditional head adjunction view of head movement would predict (§2.1). Second, as a postsyntactic operation, amalgamation is predicted not to have any interpretive effects (§2.2). Finally, we expect that locality conditions on amalgamation may differ from (and be more stringent than) the locality conditions that genuine syntactic movement is subject to (§2.3).

By contrast, we expect genuinely syntactic head movement to be associated with properties typical of syntactic movement in general, including phrasal movement. For example, in contrast with amalgamation, genuinely syntactic head movement generally results in the pronunciation of the exponent of a morphosyntactic formative in a position higher than its initial merge position, and involves no morphological word-building (§2.1). Second, syntactic head movement may exhibit interpretive effects. Our discussion focuses on case studies which claim that head movement

⁴ We refer here to the PF branch of the Y-model, which in Distributed Morphology (Embick and Noyer, 2001) involves a number of possible postsyntactic operations, the application of which ultimately leads to a phonological output. In this way of understanding the interface between syntactic structure and phonological form, the part of the PF branch where most of the action occurs in our proposal is between the output of syntax and the insertion of phonological exponents (Vocabulary Insertion).

⁵ As discussed in section 1, our primary goal is to provide empirical and theoretical arguments for a separation of the two head movement types. We make the choice here to implement these two types in terms of syntactic movement and postsyntactic amalgamation within a Distributed Morphology model of the PF interface. It should be noted, however, that the broader arguments are in principle consistent with an alternative view in which the two types of head movement involve, for example, composition in a generative lexicon and syntactic movement. We leave it to future work to tease out any differing predictions made by such alternative implementations. Thanks to David Pesetsky and Paul Kiparsky for discussion of this point.

has interpretive effects, pointing out that, to the extent the relevant arguments go through, such cases are reported only for movements which we would classify as genuinely syntactic on independent grounds (§2.2). Finally, syntactic head movement should be subject to the same locality restrictions that apply to syntactic phrasal movement (§2.3).

	postsyntactic amalgamation	syntactic head movement
produces head-adjunction structures (which map to words)	yes	no
driven by morphological properties of heads	yes	no
obeys the Head Movement Constraint	yes	no
obeys constraints on phrasal movement	no	yes
potential for interpretive effects	no	yes

Table 1 Properties of syntactic head movement and postsyntactic amalgamation

2.1 Linearization and morphosyntactic complexity

In the traditional implementation of head movement (Baker 1988 and many others), the operation involves two components. There is a movement component, whereby a head is displaced upward, subject to the Head Movement Constraint (Travis, 1984). There is also an adjunction component, resulting in the adjunction of the displaced head to another head. When applied iteratively, the two components result in the formation of complex heads. Under the fairly standard assumption that morphemes represent separate syntactic heads, this view leads us to expect two interconnected outcomes when head movement applies. First, we expect the morphosyntactic complex that results from head movement to be higher in the syntactic structure than the initial Merge position of any of the complex's subparts that undergo head movement. Second, because there is adjunction, we expect the resulting complex to grow in size with each successive instance of head movement—this is the basic idea associated with so-called “roll-up” movement. Taken together, what we expect is a situation in which the morphosyntactic complex grows in size as head movement takes place up the tree (an idea elaborated in Abels 2003a; Brody 2000; Harley 2013b; Adger et al. 2009 and Hall 2015):

(1) The size-height correlation (SHC):

Under the syntactic head-adjunction view of head movement, if a head has undergone head movement, it should be a proper subpart of a head adjunction structure (i.e., a complex head).

Indeed, this correlation is often observed, famously so in Pollock's (1989) analysis of the French verbal system. In this canonical case study, the formation of a complex head containing both the verbal root and finite tense and inflection is accompanied by raising of the verbal root from its base position to a higher syntactic position, above negation, adverbs and quantifiers:⁶

(2) Finite verb w.r.t. adverb

- a. Astérix **mangeait** souvent du sanglier.
Asterix eat.3.IMPF often of boar
'Asterix often ate boar.'

(Harley 2013a:113, (2))

- b. *Astérix souvent **mangeait** du sanglier.
Asterix often eat.3.IMPF of boar

(3) Finite verb w.r.t. negation

- a. Jean ne **parlait** pas français.
Jean NE speak.3.IMPF not French
'Jean was not speaking French.'

(Harley 2013b:46, (1))

⁶ Our use of the terms *raising* and *lowering* at this stage in the discussion is descriptive: *raising* or *lowering* of X means the pronunciation of X in a higher or lower position than X's base position in the syntax.

- b. * Jean ne pas **parlait** français.
Jean NE not speak.3.IMPF French
- (4) Finite verb w.r.t. quantifier
- a. Mes amis **aiment** tous Marie.
my friends love.3PL.PRES all Marie
'My friends all love Marie.' (Pollock 1989:367, (5b))
- b. * Mes amis tous **aiment** Marie.
my friends all love.3PL.PRES Marie
(Pollock 1989:367, (5d))

By contrast, in nonfinite clauses, the verb does not have the same morphological requirements and therefore does not move, appearing instead to the right of the VP adjuncts:

- (5) a. Nonfinite verb w.r.t. adverb
à peine **parler** l'italien ...
hardly speak Italian
'to hardly speak Italian' (adapted from Pollock 1989:377, (24a))
- b. Nonfinite verb w.r.t. negation
ne pas **regarder** la télévision ...
NE not to watch the television
'not to watch television' (adapted from Pollock 1989:374, (16e))
- c. Nonfinite verb w.r.t. quantifier
... tous **sortir** en même temps de la salle
all leave at same time from the room
'all to leave the room at the same time' (adapted from Pollock 1989:377, (25c))

French represents the situation that the SHC leads us to expect: morphosyntactic complexes bearing the expression of T and Agr unify by head movement to that higher position, yielding concomitant word order differences between finite and nonfinite clauses—in the latter, the verbal complex bears no inflection and stays low.

What we demonstrate here, however, is that the expectations associated with the SHC are not borne out consistently (see also Brody 2000; Abels 2003a; Adger et al. 2009; Harley 2013b). First, movements to higher positions do not always yield more morphosyntactically complex structures. A canonical case of this is movement of T to C in English or German matrix clauses (den Besten, 1983; Vikner, 1995), which results in displacement of a head to a demonstrably higher position, but does not result in the expected growth of the morphosyntactic complex (in the case below, that of the auxiliary):

- (6) a. Paul has arrived on campus. (auxiliary in T)
b. Has Paul arrived on campus? (auxiliary in C)
- (7) a. Ich glaube, dass Fritz Dieses Auto in München geklaut hat.
I believe that Fritz this car in Munich stolen has.
'I believe that Fritz stole this car in Munich.' (auxiliary in T)
- b. Dieses Auto hat Fritz in München geklaut.
this car has Fritz in Munich stolen
'Fritz stole this car in Munich.' (auxiliary in C)

If adjunction is the mechanism involved in all types of head movement, then the theory as it stands is forced into the uncomfortable position of positing a null C in all cases of this type crosslinguistically, without any explanation for the source of this uniformity (see also Harley 2013b).

The second way in which the SHC is not borne out is that large morphosyntactic complexes are fairly regularly realized quite low in the structure. On the opposite end of the spectrum from French is a language like Danish, which exhibits distinct verb placement patterns in root and embedded clauses. It is a verb-second language in root clauses, a fact which we take to follow from the standard analysis of verb second orders in Germanic and Scandinavian—finite main verbs move to C and a phrasal constituent moves to the specifier of C (Vikner, 1995). As with German, there is no movement of the verb to C in embedded clauses.

- (8) a. Peter **drikker** often kaffe om morgonen.
Peter drinks often coffee in morning.DEF
'Peter often drinks coffee.' (Vikner 1995:47, (33c))
- b. Vi ved at Peter ofte **drikker** kaffe om morgenen.
we know that Peter often drinks coffee in morning.DEF
'We know that Peter often drinks coffee in the morning.' (Vikner 1995:47, (33f))

Most important for the purposes of our discussion is the observation that the order of verbs with respect to adverbs and negation indicates that the verb does not raise even to T in embedded clauses:

- (9) a. Jeg spurgte hvorfor Peter ofte **havde** læst den.
I asked why Peter often had read it.
'I asked why Peter had often read it.' (Vikner 1995:145, (32b))
- b. *Jeg spurgte hvorfor Peter **havde** ofte læst den.
I asked why Peter had often read it.
(Vikner 1995:145, (32a))
- (10) a. Jeg spurgte hvorfor Peter ikke **havde** læst den.
I asked why Peter not had read it.
'I asked why Peter had often read it.' (Vikner 1995:145, (31b))
- b. *Jeg spurgte hvorfor Peter **havde** ikke læst den.
I asked why Peter had not read it.
(Vikner 1995:145, (31a))

This is a case in which a morphosyntactically complex verb stays relatively low (or even *in situ*) despite bearing exponents that realize the featural makeup of T (tense and inflection). Although the focus here is on Danish, the pattern is not uncommon and occurs in a typologically diverse set of languages: a cursory review of the literature on head movement yields similar patterns in Faroese (Rohrbacher, 1994), Swedish (Vikner, 1995), Norwegian (Vikner, 1995), Kiowa (Adger et al., 2009), Mixtec (Ostrove, 2016), Chibemba (Julien, 2002), Ndebele (Pietraszko and Arregi, 2018), in the Spanish nominal domain (Lipták and Saab, 2014), and in at least one variety of Korean (Han et al., 2007). The evidence then points to a double dissociation: there is pronunciation in a higher position without morphological unification, and there is morphological unification without pronunciation in a higher position.

It is not the case that the head-adjunction view of head movement is completely incompatible with the facts presented here. Our point is, rather, that the head adjunction view must stipulate two things: (i) whenever displacement without morphological growth is attested, there is adjunction to a null head; and (ii) whenever morphological growth without upward displacement is attested, some other process (e.g., *Lowering* in Distributed Morphology) is involved. To the extent that the processes involved in (i) and (ii) pattern differently with respect to some set of properties—locality and interpretive effects, for example—the head adjunction account misses an important generalization.

The contrast between French and Danish with respect to the locus of the pronunciation of each language's verbal complex indicates that morphological amalgamation of distinct syntactic terminals need not involve actual upward movement. If we think of the individual syntactic heads in the extended projection of the clause as requiring morphological unification, then what we have in French and Danish are points of pronunciation at the highest and (potentially)

the lowest loci in the relevant range of projections.⁷ There is, in addition, evidence that languages may realize the amalgamated complex at other points along the extended projection. This differentiates the amalgamation cases from the traditional head adjunction approach, in the following respect. Although both systems in principle permit pronunciation at intermediate positions, the head adjunction view (in the absence of additional mechanism, e.g., *Lowering*) fails to predict cases in which the complex is pronounced at an intermediate point in the extended projection, but bears evidence of morphological amalgamation higher than the point of pronunciation. An example might be a case in which a T-Asp-*v*-V complex is pronounced in Asp. We are aware of two such case studies: Russian and Irish. We discuss Russian here to illustrate the point, and refer the reader to McCloskey 1996; Ostrove 2015 and McCloskey 2017 for elaborate discussions of Irish, in which the verbal complex includes a complementizer but is pronounced in a position demonstrably lower than C.⁸

Russian verbal complexes are comprised of morphemes that can be broken down into the morphosyntactic reflexes of (at least) roots, Asp (aspectual prefixes), Neg (negation), and T (tense and ϕ -agreement).⁹

- (11) ne za-bol'-e-va-la
 NEG PFX-hurt-THEME-2IMPF-PST.SG.F
 'she was not falling ill' (IMPERFECTIVE)

The classic observation about Russian verbs is that they do not pass the canonical tests for verb movement (Bailyn, 1995a,b): neither low adverbs (12) nor floating quantifiers (13) intervene between the verb and object, indicating that the finite verb is positioned relatively low, probably below T.¹⁰

- (12) Ivan často ubiraet (*často) komnatu.
 Ivan.NOM often cleans.3SG (*often) room.ACC
 'Ivan often cleans his room.' (Gribanova 2017a:1095, (32))

- (13) My vse čitaem (*vse) gazetu.
 we.NOM all read.1PL (*all) newspaper.ACC
 'We all read the newspaper.' (Gribanova 2017a:1095, (33))

However, there is evidence that the verb is not pronounced in its *in-situ* position either; its locus of pronunciation is at an intermediate point along the clausal spine, in neither the lowest nor the highest possible position. The evidence comes from examples like (14), in which an auxiliary occupies T and the verb has undergone across-the-board movement out of the VP/*v*P domain (Bailyn, 1995b).¹¹

- (14) Petja budet priglašat' Mašu v muzej segodnja, a Dinu v kino zavtra.
 Petja.NOM will.3SG invite.INF Maša.ACC in museum today CONJ Dina.ACC in movie tomorrow
 'Peter will invite Masha to the museum today, and Dina to a movie tomorrow.' (Gribanova 2013:96, (8))

If, as standardly assumed, Russian auxiliaries are housed in T (or lower than T—see Harves 2002), then the verb in such constructions is pronounced in some position between T and *v*P (15) (e.g., Neg in Gribanova 2017a).¹²

⁷ In the case of Danish, although the evidence indicates that the locus of pronunciation is below T, there is some question as to whether *v* or V is the target. As Gribanova and Mikkelsen (2017) point out, making the locus *v* will make the correct predictions with respect to word order in double object constructions, assuming that both objects are introduced inside VP. Here, we maintain, primarily for expository purposes, that the locus of pronunciation is V. If, however, it turns out that there is strong evidence for it being *v*, this does not affect our overall proposal. What we expect is that there should exist languages in which V is indeed the lowest point of pronunciation; it is possible that Kiowa or one of the other languages in our short list will make a good candidate.

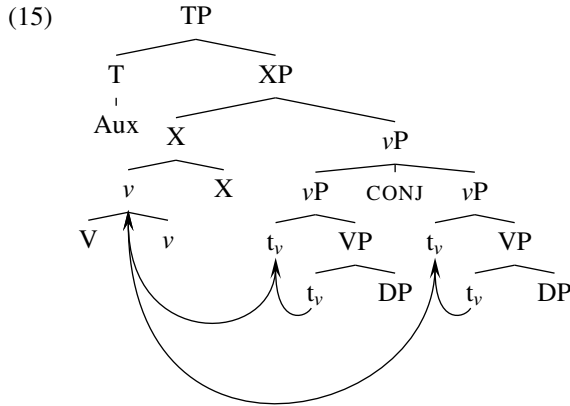
⁸ It is important to point out that neither genuine syntactic head movement nor postsyntactic amalgamation are linked to particular domains in an extended projection. The aforementioned Irish case is an example of morphological amalgamation in the C domain, while participle fronting in Bulgarian, discussed in §2.3, is an example of genuinely syntactic head movement that targets the T domain (see Harizanov 2016 for relevant arguments). We thank a reviewer for raising the important issue of proper double dissociation between the type of head movement observed and the type of landing site it targets.

⁹ Negation is written as a separate word orthographically but is proclitic on the verb.

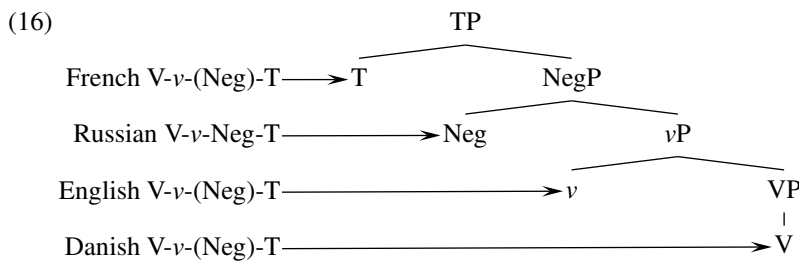
¹⁰ The negation test is irrelevant for Russian, since negation is always proclitic on the tensed auxiliary or verb.

¹¹ An alternative analysis for (14) would involve no movement of the verb, instead applying gapping (that is, deletion of an identical verb in the second conjunct of a coordinate structure). For arguments that exclude this possibility, see Bailyn (1995a) and Bowers (1993).

¹² Bailyn 1995 takes the coordination to be at the VP level, but this is because that paper predates the adoption of *v*P in the functional layer of the clause. Following Svenonius (2004), we assume that *v* hosts the verb's theme vowel, which determines numerous properties of the verb (among them argument structure, allomorphic selection, etc.). If this is the case, then coordination here takes place by hypothesis at the *v*P level.



This evidence points suggests that the verb is pronounced in a position below T, but nonetheless bears the exponents of T (tense and inflectional marking)—the locus of pronunciation is neither at the lowest nor at the highest possible point in the clausal spine. What this means is that morphologically complex verbal forms bearing the exponents of fairly high projections (e.g., T) may have variable points of pronunciation, as schematically exemplified in the tree in (16) (modulo the very real possibility that the series of projections across these languages may differ).¹³



Taken together with the evidence from English and German, we have an empirical and theoretical landscape in which amalgamation may take place in the absence of any evidence of syntactic head movement (e.g., Danish), and upward head movement may take place in the absence of any evidence of amalgamation (e.g., T-to-C in English and German). This pattern suggests that the morphological amalgamation and syntactic movement components of the operation we think of as head movement are logically independent of each other.

Our interpretation of these facts is that they indicate a different division of labor among the grammatical modules than has been previously assumed. On this view, genuine syntactic movement of heads can have effects on word order, but does not result in word formation. Amalgamation of heads into complex heads is a postsyntactic process, involving either raising or lowering of the relevant heads. Of course, these three types may co-occur: in Danish, for example, syntactic head movement (of V to C) co-occurs with amalgamation involving lowering of T to *v* and of *v* to V. And amalgamation involving raising can co-occur with amalgamation of the lowering kind, resulting in the Russian or Irish pattern, wherein the resulting complex is pronounced at the position associated with neither the highest nor the lowest head in the complex. We discuss such interactions in in §4.

2.2 Interpretive effects

Discussions about the architectural status of head movement often hinge on the question of whether head movement ever has interpretive effects. The Y-model of grammar leads us to expect that movement in the narrow syntax may in

¹³ We assume, along with Larson 1988, among others, that the English verb moves to *v*; nothing in our proposal hinges on this. On Danish, see footnote 7 and the discussion above.

principle have such interpretive effects. This expectation is largely borne out for phrasal movement, especially in the domain of \bar{A} -movement. The evidence in favor of interpretive effects of head movement is far less common, and the arguments are much more subtle. Although our discussion will necessarily make reference to some of these arguments, covering all of them here is, for space reasons, impossible. We therefore will refer to, but do not summarize, existing counter-arguments and rebuttals. We focus instead on the observation that the theoretical position outlined here leads us to an expectation about the potential interpretive effects of head movement. If there exist two subtypes of head movement, only one of which takes place in the narrow syntax, we expect that only the syntactic type will ever be able to be associated with interpretive effects. Our strategy will be to examine the existing arguments in favor of interpretive effects of head movement, and to demonstrate that for every case, the type of head movement involved is demonstrably the type that we would expect to be syntactic: for example, it may disobey the Head Movement Constraint (see §2.3), and it does not result in word formation.¹⁴

Before we begin, it is worth noting that there are putative interpretive effects of head movement which we do not consider here, because they boil down to effects that can be associated with features on the heads responsible for movement, rather than the movement itself. Examples of this type include the illocutionary force associated with verb second in Germanic languages (Wechsler, 1991; Wiklund, 2010); the interrogative force associated with verb-initiality in certain Slavic languages, when accompanied by an interrogative particle (King, 1995); the discourse effect associated with long head movement (Lambova, 2004); and the polarity focus effect associated with verb-initial orders in e.g., Russian (Gribanova, 2017a). In all of these cases, discourse effect or illocutionary force come about in conjunction with head movement, usually to the C domain. It is therefore possible to consider them as cases wherein semantic or discourse effects come about as the result of head movement. But analyses of these effects invariably involve some kind of feature (again, usually on a head in the C domain) that is associated with or triggers the head movement. It is therefore possible to analyze the relevant effects as arising not from the head movement itself, but from the featural encoding of the head that triggers the movement. We therefore do not consider such cases in the discussion that follows.¹⁵

The case studies we focus on in this section instantiate what Lechner (2007) calls *Semantically Active Head Movement* (SAHM) and can be divided into two classes. The first class involves arguments directly in favor of some kind of interpretive effect that results from the head having moved to a higher position. These typically involve evidence from scope-taking of modals or arguments based on c-command of some element as relevant for the licensing of another. A second class of arguments constitutes indirect evidence for the movement of a head to a higher position. Here, the structure of the argument is such that an explanation of some phenomenon holds *if* the head movement at issue is taken to be syntactic, e.g., since this enables variable binding. We discuss both types of argument in the subsections that follow, demonstrating in each case either (i) that the type of head movement at issue is not the amalgamation type, or (ii) that the argument does not go through. For putative counter-examples to the claim that only the syntactic type of head movement can have interpretive effects, and our responses, see the Appendix.

2.2.1 Heads taking scope

Recent investigations of the status of head movement—chief among them Matushansky 2006; Roberts 2010 and Hall 2015—focus broadly on the question of whether head movement can expand scope-taking possibilities. In such cases the canonical setup includes head movement of modals, which commonly have quantificational properties and could therefore be expected to yield a distinct interpretation if raised in the narrow syntax above another element relevant to scopal interpretation.

A first argument for the existence of interpretive consequences of head movement comes from Lechner (2007), who presents a subtle argument for the scope-expanding possibilities associated with head movement of modals with respect to negation. The argument is complex and based on a number of ancillary assumptions. It revolves chiefly around the so-called “split-scope” reading of examples like (17) (example (59) from McCloskey 2012:17):

¹⁴ A good portion of the material in this section grew out of discussions with Jim McCloskey. See McCloskey 2016 for further discussion and elaboration.

¹⁵ It is, however, telling that in all of the cases just mentioned above, the movement in question does not involve morphological amalgamation, but rather effects on word order. This is consistent with our hypothesis, in that, if a movement is syntactic, it is natural for it to be triggered by a featurally encoded head; this encoding may, in some cases, include discourse or illocutionary force features.

- (17) a. Not every lottery ticket can be drawn.
 b. SPLIT-SCOPE READING: there are no circumstances, in which all lottery tickets can be drawn.
 c. $[\neg \diamond \forall x [\dots x \dots]]$

Three crucial assumptions, defended by Lechner, are required in order to explain the availability of this interpretation. First, the modal is generated low (somewhere below T) and moves to its surface position; it is crucial to the account that this surface position is above T. Second, there is a high silent NEG operator which c-commands both the modal and the subject; this operator must c-command the negative subject in order for the negative subject to be licensed (following Penka 2002, Penka and von Stechow 2001 and von Stechow 1993). And finally, one must assume that these kinds of (strong) quantifiers cannot reconstruct below T.¹⁶

Taking these assumptions on, one arrives at the following analytical picture: the surface order in (17) does not reflect the split-scope reading, so this reading must come about as the result of interactions between the structural positions of the subject and the structural positions of the modal. The quantified subject originates low (presumably in Spec,vP) and raises through Spec,TP to an even higher surface position; for Lechner, this is Spec,AgrSP. Its possible positions of interpretation are its surface position—which would give the incorrect scopal interpretation for the split reading—and Spec,TP. It cannot reconstruct into its base position because of the assumption that reconstruction of strong quantifiers below T is not possible. The modal, too, starts low and moves to its surface position, somewhere above T, presumably. The split-scope interpretation is derivable if the modal is interpreted in its derived position and the subject is interpreted in a lower position (Spec,TP), even though the subject is pronounced higher.

The argument is complex, and there are slightly differing interpretations of it across the literature. Whatever the details may be, the key point for the purposes of the present argument is that the scope-expanding possibilities of head movement here are associated with a modal: there is no word-building component to the movement, only effects on word order. This is what our proposal predicts: syntactic head movement is the only kind of head movement we expect to have any interpretive effects.

A second case study comes from the interaction between aspectual verb fronting and the focus operator *only* in Shupamem, a Grassfield Bantu language (Szabolcsi, 2011). Szabolcsi's starting observation is that in English examples like (18), the focus operator *only* and the aspectual raising verb *begin* may interact to yield either one of two readings (example (62) from McCloskey 2012:18):

- (18) a. In May only Mary began to get good roles.
 b. Reading one: Only Mary is such that she began to get good roles.
 c. Reading two: It began to be the case that only Mary got good roles.

The two readings (18b,c) associated with (18a) are, in certain languages that have verb movement, disambiguated by word order. The Shupamem cases involve a lower and a higher pronunciation point for the verb, with an apparent difference in interpretation:

- (19) a. Ndúú Maria ká jéǰǰ jìngět ndáá lí?
 only Maria PST begin have.INF good roles
 'Only Mary is such that she began to get good roles.' (Szabolcsi 2011:10, (20))
 b. Á ká jéǰǰ ndúú Maria jìngět ndáá lí?
 it PST begin only Maria have.INF good roles
 'It began to be the case that only Mary got good roles.' (Szabolcsi 2011:11, (24))

Under the assumption that the word order in (19b) is derived via movement of the verb to its higher position, this is a type of scope-expanding head movement. Notice, once again, that this movement is also of the syntac kind (with effects on word order) and not morphological amalgamation.¹⁷

¹⁶ For evidence to the contrary, see Iatridou and Sichel 2011, and discussion in Hall 2015 and McCloskey 2016.

¹⁷ As a reviewer points out, the difference in interpretation could, in principle, be instead the result of the subject raising in (19a) but not (19b), with the verb *begin* staying put in both examples. If this is the right analysis, it would cast doubt on the overall claim that the verb movement is doing the semantic work here and the Shupamem examples would not be an instance of semantically active head movement. For discussion of this alternative approach, see Szabolcsi 2011, section 3.2.1.

To summarize, the head movement in both of the case studies presented here is taken to expand scope-taking possibilities, and it is also the type of head movement that affects word order. Crucially, it is *not* the type of head movement that results in morphological word-building. We suggest that this is not accidental but a systematic correlation that results from the syntactic status of the movement operation in question.

2.2.2 Indirect arguments for SAHM

A second set of arguments for interpretive effects of head movement is composed of analyses which have, as their consequence, the requirement that head movement feed the LF interface (e.g., participate in variable binding under the assumption that variable binding is computed at the interface). The primary set of arguments comes from cases in which head movement interacts with ellipsis licensing. The relevant observation is that in such cases head movement appears to behave like phrasal movement in leaving behind a trace that is relevant for variable binding. To the extent that we can explain constraints on ellipsis by using the logic associated with variable binding in the case of head movement, this necessitates the conclusion that the movement at issue takes place in the narrow syntax. Important for our purposes is the observation, to be expanded upon below, that the cases in question all involve head movement of the syntactic kind—this is the kind of head movement that re-arranges words but does not result in word formation.

The crucial arguments revolve around an ellipsis constraint, *MaxElide*, which requires that the larger of two potential ellipsis sites (e.g., TP vs. *v*P) be targeted for ellipsis if WH-movement has taken place out of the ellipsis site (creating a binding configuration). Thus, while both large and small ellipses are possible in the absence of WH-movement out of the ellipsis site (20), only sluicing (assumed to involve TP ellipsis) is an option if a WH-phrase has moved out of the ellipsis site (21).

- (20) a. John's leaving. Do you know when (he is)?
 b. Mary was trying to kiss someone, but I don't know when (she was).
- (21) a. They studied a Balkan language, but I don't know which. (Merchant 2008:139 (27b); see also Lasnik 2001)
 b. */? They studied a Balkan language, but I don't know which they did. (Merchant 2008:139 (27c))

We refer the reader to Takahashi and Fox 2005, Merchant 2008, and Hartman 2011 for in-depth descriptions of the workings of *MaxElide*. For our purposes, it suffices to say that there is a way of understanding the patterns in (21), based on an argument discussed in Takahashi and Fox 2005, wherein *MaxElide* requires the choice of the largest possible ellipsis domain within a parallelism domain.

- (22) PARALLELISM: For ellipsis of EC [elided constituent] to be licensed, there must exist a constituent—the parallelism domain—which reflexively dominates EC, which is semantically identical to an antecedent constituent, modulo focus-marked constituents. (modified slightly from Hartman 2011)

If a phrase moves out of a potential ellipsis site, the parallelism domain must be large enough to include the binder and the bindee; otherwise, the variable inside the ellipsis domain would be free, and nonidentical to its antecedent counterpart.¹⁸ This forces the parallelism domain (underlined in (23)) to be larger than it otherwise would be, and consequently limits *MaxElide* to the choice of a larger ellipsis site than would be required if no movement had occurred.

- (23) a. He might talk to someone, but I don't know who (*he might). (Gribanova 2017a:1113, (75))
 b. someone λy . he might talk to y but I don't know who λx . [_{TP} he might [_{VP} talk to x]]

There are two studies in which it has been argued that this same logic extends to cases in which head movement has applied out of the ellipsis site. The logic is as follows: head movement occurs out of an ellipsis site and forces a larger parallelism domain; this in turn forces *MaxElide* to target the larger of two potential ellipsis sites within that larger parallelism domain. For this logic to go through, it must crucially be the case that the head movement is syntactic in nature: recall that in cases of phrasal movement, the parallelism domain must be larger because of the necessity

¹⁸ This holds as long as we also assume Heim's (1997) ban on meaningless co-indexation, which ensures that the free variable in the antecedent and elided constituent are not accidentally co-indexed.

for both the binder and the bindee to be within one domain. In order for this to be true of the head movement cases, the movement in question must be of a syntactic nature (it must leave a variable for there to be binding). Thus, the logic of *MaxElide*, when extended to head movement, constitutes an indirect argument for the syntactic nature of that movement.

In the first case study, Hartman (2011) demonstrates for English that the application of subject auxiliary inversion in root, but not embedded, questions has consequences for the (un)availability of different sizes of ellipsis. In embedded contexts, where subject-auxiliary inversion is absent, both *vP* and *TP* ellipsis are permitted (24).¹⁹ This contrasts with root environments, where application of subject-auxiliary inversion reduces the choice of ellipsis site to just the larger one—*TP* (25).

- (24) a. Mary wants to leave, but I don't know when (she will).
 b. when $\lambda x. [_{TP} x [_{TP} she \lambda y. will [y \text{ leave}]]]$
- (25) a. Mary wants to leave. Really? When (*will she)?
 b. when $\lambda x. will \lambda z. [_{TP} x [_{TP} she \lambda y. z [y \text{ leave}]]]$

Of interest to us is the observation that the head movement in question involves, on the standard analysis, movement of *T* to *C*—no morphological amalgamation is involved, only word order rearrangement. If one accepts Hartman's argument,²⁰ the interpretive effects we find here—in the form of head movement giving rise to variable binding—are expected, because the head movement is of the type that we would classify as syntactic.

A second case study, demonstrating an analogous kind of *MaxElide* effect in Russian verb movement from *Neg* to *Pol*, is discussed in Gribanova 2017a. Movement from *Neg* to *Pol* derives discourse-special *VSO* orders and does not contribute to word-building, since the verbal complex is morphologically no different in verb-initial positions than it would be in *SVO* orders. It also gives rise to a *MaxElide* effect, wherein *TP* ellipsis, but not *vP* ellipsis, is licensed when the verb moves from *Neg* to *Pol*. This clustering of properties is not an accident on our account, but rather a consequence of the observation that head movement may be either syntactic or postsyntactic, with correspondingly differing properties associated with either the syntactic or the postsyntactic module.

A final argument in this group comes from Keine and Bhatt's (2016) analysis of German long passives, in which certain restructuring verbs can be passivized so that the restructuring verb and the embedded verb form a unit with respect to a battery of syntactic tests (Keine and Bhatt 2016:1446, (1a-c)):

- (26) a. Gestern hat Fritz den Traktor zu reparieren versucht.
 yesterday has Fritz.NOM the tractor.ACC to repair tried
 'Yesterday Fritz tried to repair the tractor.' ACTIVE
- b. Gestern wurde den Traktor zu reparieren versucht.
 yesterday was the tractor.ACC to repair tried
 'Yesterday it was tried to repair the tractor.' LOCAL PASSIVE
- c. Gestern wurde der Traktor zu reparieren versucht.
 yesterday was the tractor.NOM to repair tried
 'Yesterday it was tried to repair the tractor.' LONG PASSIVE

The authors' core insight is that certain effects render it necessary to adopt a hybrid account in which a subset of cases of the local passive is derived via head movement of the lower verb to the position of the higher verb (verb cluster formation). This happens just in case two conditions hold: first, the complement of the restructuring verb is defective (i.e., small, a *VP* rather than a *vP*); second, the lower *VP* remains *in situ*.

- (27) a. $[_{VP} [_{VP} DP V] V] \rightarrow [_{VP} [_{VP} DP t] [_v V V]]$ verb cluster formation (=26b)
- b. $[_{VP} DP V] \dots [_{VP} t V]$ no cluster formation (=26c)
-

¹⁹ The presence of a *WH*-phrase in such examples is irrelevant for the calculation of parallelism domains, since by assumption this *WH*-phrase is an adjunct. See Hartman 2011:378, (31) and Hartman 2011:372, (12) for examples analogous to (24,25).

²⁰ See Hall 2015, Messick and Thoms 2016 for discussions that question the validity of Hartman's result.

The relevance of this analysis for us is that it depends on the head movement in question being syntactic in nature. An argument for the syntactic status of the head movement is that it yields certain scopal effects—namely, the higher verb can reconstruct to a lower position if the verb cluster formation has happened.

Taking these two pieces of reasoning at face value, we conclude that the putative syntactic nature of the head movement in question is consistent with our proposal (although our technical implementation might be different from that of Keine and Bhatt). If the movement is syntactic, we expect it to potentially yield scope-expanding effects; and we also expect it to be subject to the same kinds of linearization constraints that apply to phrasal movement. Finally, in the cases where verb cluster formation occurs, the resulting object is not a morphological word, but rather two words, with independent word-internal phonologies.²¹

2.3 Locality

A central property of (and a long-standing puzzle about) head movement, as traditionally understood, is that it is subject to locality constraints that appear to be stricter than the locality constraints associated with phrasal movement. For example, the locality of Internal Merge is robustly governed by relativized minimality, whereby syntactic movement involves movement of a goal that is closest to the probe in terms of c-command and that matches the feature(s) the probe is searching for (Rizzi 1990, Chomsky 1995b). This conception of the locality of Internal Merge allows the goal to “skip” an element when moving as long as that element does not itself count as an eligible goal. If head movement is uniformly syntactic and thus involves Internal Merge, all else equal, it would be expected to obey the same locality constraints as phrasal movement. Yet, a subset of head movements is constrained by the Head Movement Constraint, which leads to apparently far more local movements:

- (28) Head Movement Constraint (HMC) (Travis, 1984, 131):
An X^0 may only move into the Y^0 which properly governs it.

This constraint was formulated to account for the observation that many instances of head movement involve movement of a head only to the next head up and do not involve “skipping” of heads. This differs from what we find for phrasal (syntactic) movement, where a goal can be moved across one or more intervening elements (including heads) as long as the intervening heads do not match the features that the probe is searching for.

One response to this state of affairs is to maintain that all head movement is syntactic, as in the traditional understanding, and attempt to derive the HMC as a special case of relativized minimality (Rizzi 1990, Roberts 1991). Another response, and the one we explore here, is that there are in fact two different operations—postsyntactic amalgamation and genuinely syntactic movement of heads—subject to distinct locality constraints. Genuinely syntactic head movement, being an instance of Internal Merge, obeys the same constraints as all other syntactic movement (e.g., stemming from relativized minimality) and does not obey the HMC (although it can result in movements that appear to comply with the HMC; see footnote 23). Postsyntactic amalgamation, on the other hand, involves morphological word building and, correspondingly, is an operation that is subject to a stricter locality constraint, which subsumes the effects of the HMC.

According to our proposal, the HMC embodies a generalization about the locality of a postsyntactic operation. In this context, the HMC is therefore properly viewed as a locality condition on postsyntactic amalgamation, which requires the affected heads to be structurally adjacent. This locality condition holds equally regardless of whether the resulting amalgam is pronounced in the position of the higher head (in instances of what we analyze as postsyntactic head raising) or the lower head (in instances of what we analyze as postsyntactic head lowering). In agreement with the structural conditions on the Distributed Morphology operation *Lowering* (Embick and Noyer, 2001), according to which a head can lower only to the head of its complement,²² we generalize a version of the HMC to postsyntactic amalgamation applying both upwards and downwards (see section 3).

In addition, if our discussion in previous sections is on the right track, postsyntactic amalgamation is expected to give rise to HMC effects, to be accompanied by evidence of word formation and to lack interpretive effects. In

²¹ There is evidence that the verb clustering long passives result in the elimination of an intonational break that would otherwise be there (Wurmbrand, 2001, 295ff). This, however, is not the same as being part of the same morphological word. See Harizanov 2018 for discussion of morphological vs. phonological wordhood.

²² See §3 for a slight modification of this definition, which nevertheless retains the restriction that results in HMC effects.

contrast, the locality of genuinely syntactic head movement should go hand in hand with absence of evidence for the construction of complex heads and potential to be associated with interpretive effects. Next, we focus on three such cases, though many others have been documented: V-to-C movement in Danish root clauses, long head movement in Bulgarian, and predicate clefting in Hebrew. In each of these cases, we point out that beyond instantiating violations of the HMC—in fact, long head movement in Bulgarian and Hebrew can cross finite clause boundaries—these paradigms do not involve word formation, but rather have effects on word order. They also involve various discourse effects, including types of (contrastive) focus and illocutionary force. This is exactly what we expect if such cases are instances of genuinely syntactic movement.

2.3.1 V-to-C without V-to-T

As discussed above, the Danish finite verb occupies what appears to be the C position in root clauses, yielding a verb second configuration, but stays in its low, thematic position in embedded clauses. Vikner (1995) argues that Faroese, Norwegian, and Swedish behave similarly. Thus, as Vikner (1995, 143–147) points out, while there is evidence that the verb in these languages moves to C, there is no independent evidence that the verb moves to an intermediate position (I, T, or INFL) (contra Holmberg 1986, 90, Lightfoot 1991, 46, Lightfoot 1993, 212, fn. 5). To make the movement of the verb to C comply with the HMC, it is usually proposed that V in these languages moves through I, T, or INFL only if it is on its way to C (i.e., in verb second clauses). However, in the absence of evidence for this intermediate movement step, we contend that it is consistent with the facts to treat syntactic verb movement as targeting C directly, crossing any intervening heads. According to this treatment, the derivation of verb second clauses in languages without independent V-to-T movement involves genuinely syntactic head movement of V to C that is not subject to the HMC.²³

Our understanding of this pattern leads to a situation in which there is postsyntactic amalgamation (lowering) of T to *v* and of *v* to V. This co-occurs with a genuine case of syntactic movement of V to C.²⁴ A discussion of how these operations will interact follows in §4. Following the theme of this section, we want to point out here that the movement of V to C in Danish yields only structural displacement but not word formation effects.

2.3.2 Long head movement

In Bulgarian, a participle canonically follows a finite auxiliary and precedes its own arguments. For example, in (29a,30a), the participle *pročel* ‘read’ follows the past perfect auxiliary *bjah* and precedes the object *knigata* ‘the book’. In addition to this order, Bulgarian also allows the participle-auxiliary order, shown in (29b,30b). Here, the participle appears displaced from its thematic VP-internal position to the clause-initial pre-auxiliary position (Harizanov 2016:1, (3)-(4)).

- (29) a. Bjah pročel knigata.
 had read the.book
 ‘I had read the book’
 b. Pročel bjah knigata.
 read had the.book

²³ It is possible that verb-second clauses in other Germanic languages also involve genuinely syntactic head movement, but that it has the appearance of a more local, HMC-obeying movement. For example, T-to-C movement in English questions appears to obey the HMC, but that might be because C happens to attract T, which is the head of C’s complement. Thus, it is possible to maintain that T-to-C in English is of the genuinely syntactic type of head movement and that it is accidental that the interacting heads are in the local relation prescribed by the HMC. A similar question arises for Danish (discussed below and in section 4.1.2), where it is the highest verbal element that undergoes movement to the C domain in root clauses. As a reviewer points out, in traditional approaches to head movement, this is accounted for by the HMC. Within the present approach, such cases must be the result of a relativized minimality constraint on the movement of the verbal element that is highest in the clausal structure in terms of c-command (this issue is further discussed in section 4.1.2). This is accommodated by our proposal, which allows but does not require syntactic head movement to skip heads.

²⁴ As mentioned in footnote 7, it may be the case that the locus of pronunciation in Danish is *v*, not V. Nothing in our proposal hinges on this being the case; to adopt this alternative analysis, the system we propose will require postsyntactic lowering of T into *v*, postsyntactic raising of V into *v*, and syntactic movement of *v* to C in V2 environments.

- (30) a. Razbrah če e pročel knjigata.
understood.1.SG that is read the.book
'I understood that he has read the book'
- b. Razbrah če pročel e knjigata.
understood.1.SG that read is the.book

Since it violates the HMC, this kind of participle fronting has been dubbed "long" head movement (Lema and Rivero 1990). The HMC is clearly violated by participle fronting when it crosses more than one auxiliary, as in (31) where the fronted participle moves over both *šte* and *sām*.

- (31) a. Šte sām pročel knjigata.
will have read the.book
'I will have read the book'
- b. Pročel šte sām knjigata.
read will have the.book

(Embick and Izvorski 1997a, (30))

Furthermore, participle fronting can escape both nontensed and tensed clauses (32), and is island-sensitive (33), suggesting a clear parallel with syntactic phrasal movement as far as the locality constraints it is subject to are concerned (Harizanov 2016:7, (23)-(24)):

- (32) a. **zagazil** može [da e]
gotten.in.trouble might to be
'he might've gotten in trouble'
- b. **zaspali** si pomislih [če bjaha decata veče]
fallen.asleep REFL I.thought that were the.children already
'I thought the children had already fallen asleep'
- (33) a. ***polučila** si trägna [predi da e podaräka si]
received REFL left.3.SG before to is the.gift REFL
'she left before she received her gift'
- b. ***pročel** [ne beše novata kniga]
read not was the.new book
'he had not read the new book'

Numerous superficially similar phenomena across languages have been labeled "long" head movement (Lema and Rivero 1990, Rivero 1991, 1992, 1994, Roberts 1994, Wilder and Cavar 1994, Schafer 1995, Borsley and Kathol 2000, among others). Many of these involve a nonfinite verb in clause-initial position which appears to have moved from its thematic position across an auxiliary (Slovak examples (39a-c) from Rivero 1991):

- (34) a. Ja som napísal list.
I have written letter
'I have written a letter.'
- b. Napísal som list.
- c. * Som napísal list.

An intuition that runs through most of the literature on long head movement in these languages is that fronting of the nonfinite verb is a way of avoiding an outcome where a phonologically weak element appears in clause-initial position. This kind of analysis seems plausible, given that the intervening auxiliaries in many of the languages under consideration are phonologically weak. For these languages, it is difficult to exclude an alternative analysis in which the nonfinite verb does not undergo syntactic movement, but the phonologically weak auxiliary undergoes prosodic

inversion to the right. However, a prosodic inversion analysis is much less likely in a language like Bulgarian, in which (i) a participle appears before an auxiliary that is not an enclitic (29), and (ii) a participle appears before an enclitic auxiliary that is already supported (30). We conclude, therefore, that at least in such cases, genuine syntactic movement is involved.

Additionally, participle fronting in Bulgarian involves heads and not phrases, as demonstrated by Rivero (1991), Lambova (2004), and Harizanov (2016) (see Broekhuis and Migdalski 2003 for an opposing view). Thus, participle fronting involves movement of a head that potentially crosses other heads as well as clause boundaries, in violation of the HMC. As expected, this kind of long head movement also does not lead to word formation and has discourse effects (Lambova 2004). This is the expected state of affairs if the nonfinite verbs in Bulgarian participle fronting undergo genuinely syntactic movement via Internal Merge, no different from the syntactic movement of phrases.

2.3.3 Predicate clefting

Certain types of predicate clefting patterns in a range of languages (e.g., Spanish, Hebrew) have been analyzed as involving a long-distance head movement to some clause-initial position. In Hebrew, which is the case we concentrate on here, Landau's (2006) finding is that this construction involves spell-out of the verb in two positions—one at the level of T and one inside a fronted VP—with one copy being fully inflected, while the other is in an infinitive form.²⁵ The internal argument of the verb in such cases can be fronted ('phrasal-infinitive fronting') or left behind ('bare-infinitive fronting'):

- (35) a. Liknot, hi kanta et ha-praxim.
to-buy she bought ACC the-flowers
'As for buying, she bought the flowers.' (Landau 2006:37, (8b))
- b. Le'ha'amin, hu ma'amin be-nisim.
to-believe he believes in-miracles
'As for believing, he believes in miracles.' (Landau 2006:37, (9b))
- c. Likro, hu kara et ha-sefer.
to-read he read ACC the-book
'As for reading, he read the book.' (Landau 2006:50, (37b))

Landau's analysis is that in cases where only the verb is fronted (and the internal argument is stranded), this results from long head movement to a clause-initial position. He demonstrates that both the VP-movement and the long V-movement cases are formed by \bar{A} -movement that can cross finite clause boundaries and are island sensitive (Landau, 2006, 42-43).

The relevance for our discussion is that Landau's (2006) analysis of Hebrew predicate clefts represents another case in which what is demonstrably movement of a head takes place over a long distance, violating the HMC. This same movement also has fairly clear discourse consequences (hence the analysis in which the head hosting the moved verb is a Top[ic] head), and involves no word formation—the fronted copy is actually less morphologically rich than its lower double. All this is, once again, expected on an account in which this kind of head movement is taken to be genuinely syntactic (as opposed to postsyntactic).

2.4 Summary

Summarizing the discussion in this section in broad strokes, our contention is that there are two distinct mechanisms behind the empirical effects which have, traditionally, been associated with head movement: a syntactic operation and a postsyntactic operation. This view is supported by the observations we have made in this section: each mechanism has its own clustering of properties, which in turn follow from the syntactic or postsyntactic status of the operation at issue:

²⁵ In section 4.1.4, we offer an implementation of Landau's (2006) analysis of verb doubling in Hebrew within the system proposed here.

- | | |
|---|--|
| <ul style="list-style-type: none"> – Syntactic head movement – does not form words – can “skip” heads – can have interpretive effects | <ul style="list-style-type: none"> – Postsyntactic amalgamation – forms words – affects structurally adjacent heads – does not have interpretive effects |
|---|--|

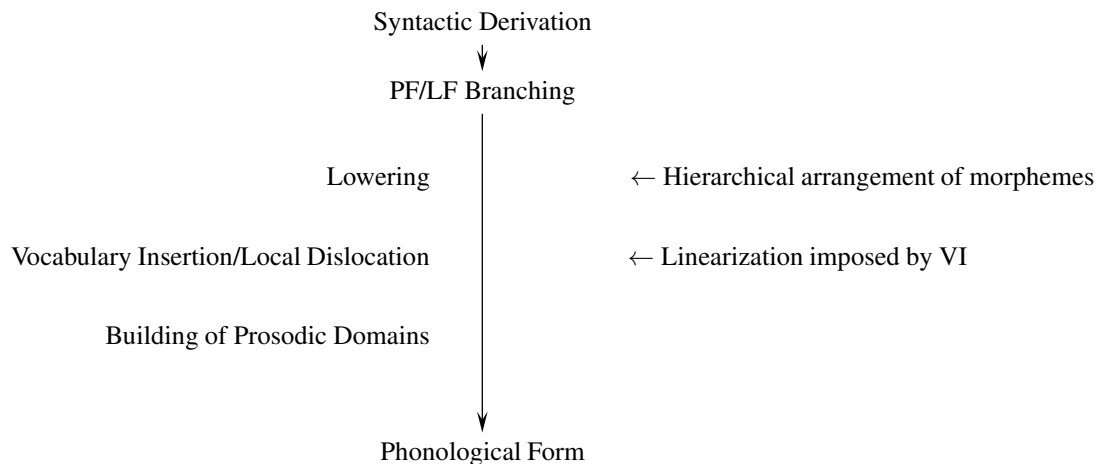
Therefore, if the displacement of a head does not obey the HMC and/or gives rise to interpretive effects, it can be concluded that genuinely syntactic head movement is involved. By contrast, if head displacement involves evidence of word formation, it can be concluded that amalgamation is involved. With this much in place, we are now in the position to discuss, in following sections, the concrete nature of these two operations, as well as predictions about their interaction.

3 Morphological amalgamation at PF

With the respective properties of amalgamation and syntactic head movement established, we can now characterize each of these operations in more depth. Positing that the morphological amalgamation of separate heads into a complex head takes place postsyntactically allows us to capture key characteristics of what is traditionally called “roll-up” head movement. Given this much, if we also assume, as is standard, that the output of syntax is mapped to LF and PF independently, we expect not to find any interpretive effects of amalgamation. We also expect that the locality constraints that govern amalgamation, which give rise, in part, to HMC effects, may be distinct from the locality constraints that syntactic movement is subject to. Finally, taking amalgamation to be the result of a postsyntactic process permits us to dissociate the morphosyntactic complexity of the resulting complex head and where it is pronounced relative to other elements in the extended projection. In other words, systematic violations of the size-height correlation (SHC) are predicted to be possible on this view.

We propose to build upon existing postsyntactic mechanisms within Distributed Morphology (DM) which capitalize on the availability of a derivationally structured component of grammar that interprets the output of syntax and ultimately maps it to PF.

(36) Operations in PF (based on Embick and Noyer 2001, *Fig. 1*)



Lowering is a postsyntactic operation that takes the output of syntax as its input and forms complex head structures by adjoining a head to the next head down in the extended projection. The operation is restricted to operate between a head and the closest head that it c-commands:

- (37) Postsyntactic head *Lowering* (adapted from Embick and Noyer 2001):
 $[_{XP} \dots X \dots [_{YP} \dots Y [_{ZP} \dots]]] \rightarrow [_{XP} \dots [_{Y} Y X] [_{ZP} \dots]]]$
 (where Y and X are heads, X c-commands Y, and there is no head Z that c-commands Y and is c-commanded by X)

Our proposal is that *Lowering* has an upward counterpart at PF, *Raising*. *Raising* also takes as its input the output of syntax, and forms a complex head by adjoining a head to the next head up. This is the operation at play in the derivation of, e.g., French finite verb forms, where T amalgamates with V and the resulting complex is pronounced in the position of T. *Raising* is likewise restricted to operate between two structurally adjacent heads:

- (38) Postsyntactic head *Raising*:
 $[_{XP} \dots X \dots [_{YP} \dots Y [_{ZP} \dots]]] \rightarrow [_{XP} \dots [_{X} Y X] [_{YP} \dots [_{ZP} \dots]]]$
 (where Y and X are heads, X c-commands Y, and there is no head Z that c-commands Y and is c-commanded by X)

Generalizing over the two types (*Raising* and *Lowering*), we have a postsyntactic operation which takes two structurally adjacent heads and outputs a structure in which one is adjoined to the other. The *Raising* and *Lowering* operations work together to form morphological units out of separate syntactic heads.²⁶

Our implementation of this basic idea involves a few concrete assumptions. First, we take it that amalgamation of a head is driven by a binary morphological selection feature M—specified with either a [+] value or a [–] value—which dictates that the head will have to amalgamate postsyntactically. Thus, a head which will ultimately undergo amalgamation must be specified with some value of the feature M.²⁷ Second, the positive specification of M on a given lexical item triggers *Raising* of that head to the next head up; a negative specification triggers *Lowering* of that head to the next head down. Absence of an M feature on a head will yield no change to that head at PF. Third, as illustrated in (37) and (38), amalgamation results in the creation of a head-adjunction structure, with the moved (i.e., raised or lowered) head Chomsky-adjoined postsyntactically to the target head. The resulting head-adjunction structure is thus associated with the position of the target head. Given this much, crosslinguistic differences in where along the extended projection an amalgamated unit is pronounced boil down to differences in how the relevant lexical items are specified for the [M] feature.²⁸ Fourth, upon amalgamation, the [M] feature of the moved head is erased/becomes inactive while the [M] feature of the target head (if present) is projected to the next higher segment of the target head (i.e., it becomes a property of the newly formed complex head). Fifth, the [M] feature is a property of the lexical item in question—i.e., it is one of the features in the feature bundle that constitutes the lexical item and, consequently, a lexical item comes specified with a value for [M] from the lexicon. Sixth, in line with what is already thought to be true of other postsyntactic processes (e.g., Vocabulary Insertion), amalgamation proceeds bottom up. That is, heads are considered for amalgamation as the morphosyntactic structure is traversed from the bottom up. Seventh, amalgamation is subject to a relative locality constraint, so that the target of amalgamation for a given head H must be a head that is structurally adjacent to H, as captured by the definitions in (37) and (38).²⁹ Finally, since amalgamation does not involve Internal Merge, it does not create multiple occurrences of the affected head—i.e., amalgamation does not leave

²⁶ We leave to future work the possibility that this state of affairs could be streamlined further—i.e., that there is a single operation, *Amalgamate*, which does all of the relevant work. Existing implementations of this general idea include Brody's (2000) Mirror Theory, and Svenonius' recent application thereof in combination with the concept of spans (Svenonius, 2016, 2018).

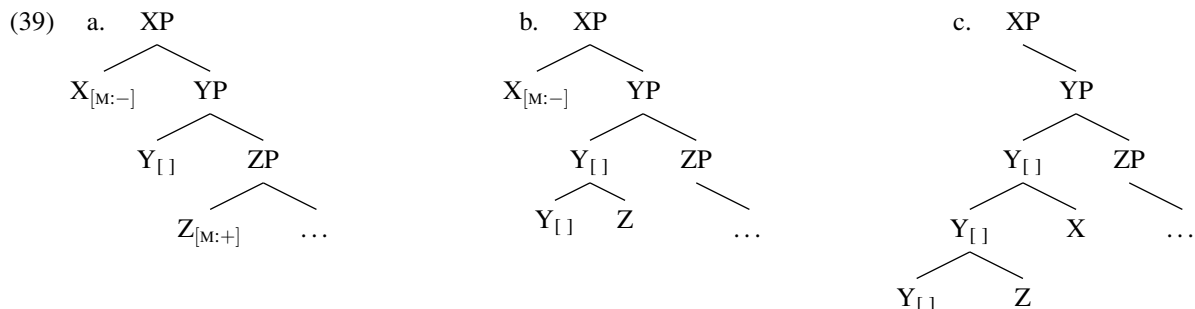
²⁷ Similar ideas can be found in Roberts 2010 and Rizzi and Roberts 1989, though implementations differ. We leave it to future work to sort out differences between our implementation and those of others, which may potentially have similar empirical coverage (Hale and Keyser, 2002; Brody, 2000; Adger, 2013; Hall, 2015; Svenonius, 2016, among others). The crucial point we make here, which is not shared by many of these approaches (except potentially Adger 2016), is that there exists a principled distinction between the mechanism responsible for morphological unification (amalgamation, or "roll-up" head movement) and the mechanism responsible for displacement in the syntax (Internal Merge). There are other differences too, however. For example, unlike amalgamation, Hale and Keyser's (2002) "conflation" operation manipulates phonological features rather than morphosyntactic terminals and, in turn, predicts that the units formed by conflation do not have complex internal morphosyntactic structure. To the extent that there is robust evidence for word-internal morphosyntactic structure, this appears to be an undesirable prediction.

²⁸ The observation that the complex head is pronounced at variable points in the extended projection is also incorporated in the accounts of Brody (2000) and Svenonius (2016, 2018), who specify the locus of pronunciation via features on heads, as we do. We leave to future work the question of whether implementational differences among these accounts will lead to distinct empirical predictions.

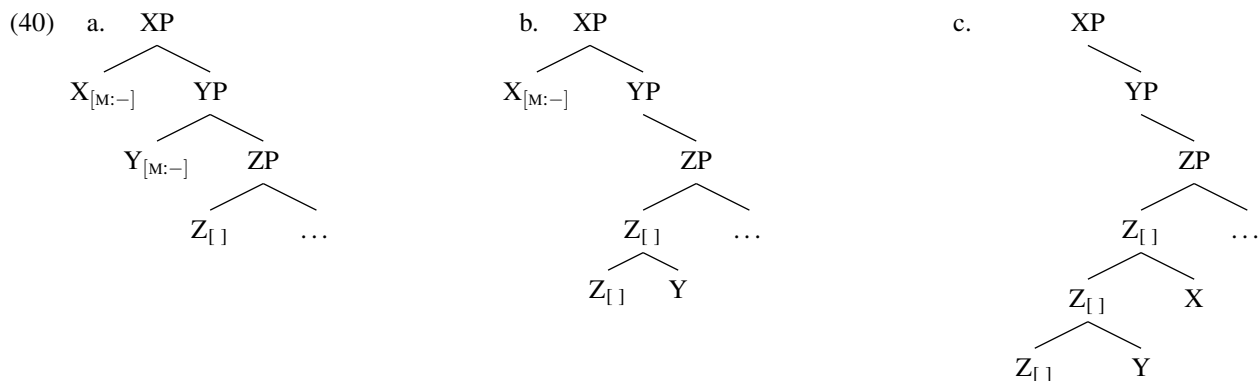
²⁹ We consider it likely that amalgamation is also subject to an absolute locality constraint to the effect that amalgamation cannot apply to heads across the boundary of an extended projection. See section 4.2 for discussion.

copies (within the Copy Theory of movement) or traces (within the Trace Theory of movement). This last assumption has consequences for our implementation of amalgamation which are discussed below in connection with (40a)–(40c).

A straightforward case of amalgamation is illustrated in (39a)–(39c). Traversing the output of syntax in (39a) bottom-up, we consider the heads Z, Y, and X in this order. Z is specified as $[M:+]$ and is adjoined to Y; Y is specified as $[\]$ so we keep going; X is specified as $[M:-]$ and is adjoined to Y.



Given the formulation of Lowering and Raising above, it is common for a head and the head of its complement to undergo amalgamation. However, as a reviewer points out, there are configurations that, given the definition of these operations—in combination with the assumptions that (i) amalgamation proceeds bottom-up and (ii) an amalgamated head leaves no trace/copy—do not involve amalgamation of a head and the head of its complement. In particular, while the Embick and Noyer’s (2001) definition of *Lowering* stipulates that a head *Lowers* to the head of its complement, this definition is too restrictive in the present theoretical context. Instead, the locality relation that needs to be established between the two heads participating in amalgamation is structural adjacency, as pointed out above.³⁰ This yields configurations like (40a)–(40c), in which the change between (40b) and (40c) involves postsyntactic *Lowering* of a head X to the closest head that X c-commands (namely, Z), even though ZP is not X’s complement.

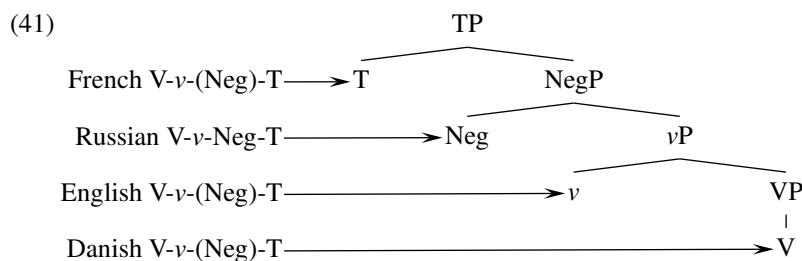


We assume that, if the configuration is such that *Raising* or *Lowering* fails to apply (i.e., their input conditions are not met) when a head is specified with a value for M, the postsyntactic derivation does not converge (at PF). For example, a nonconvergent derivation might result in a configuration where a head specified with $[M:+]$ is not c-commanded by a (sufficiently local) head that it can raise to or where a head specified with $[M:-]$ does not c-command a (sufficiently local) head that it can lower to. As an example of the latter scenario, consider the configuration in (39a) but suppose that Y is specified with $[M:-]$. This will lead to a nonconvergent derivation since, once Z raises to Y, there will be no head that Y can subsequently lower to (assuming the complement of Z is null or otherwise opaque, i.e., a phase boundary or the start of a new extended projection). In general, we assume that the result is a nonconvergent

³⁰ We take it that *Lowering* and *Raising* into structurally complex specifiers and adjuncts is prohibited by the island status of specifiers and adjuncts.

derivation whenever *Raising/Lowering* are required to apply by the presence of the M feature but the configuration is such that it prevents them from applying.³¹

The implementation of amalgamation proposed here leads to four beneficial outcomes. First, we are able to capture the observation that the resulting complex may in principle be pronounced at any point along the extended projection, apparently parametrized by language:



The postsyntactic derivation in (39a)–(39c) is essentially an abstract representation of what, in our view, happens in Russian: the verbal complex unifies V, v, Asp, Neg and T, but is realized at an intermediate point, above V but below T (namely, at Neg). Replacing the heads in (39a)–(39c) for the contentful ones in (42), and revising the structure slightly according to Gribanova (2017a), we get (43)–(48).³²

- (42) ne po-plav-a-la
 NEG PFX-swim-THEME-2IMPF-PST.SG.F
 ‘she did not swim a little bit’ (PERFECTIVE)

³¹ We assume this for completeness, but leave open the possibility that certain last-resort mechanisms may come into effect in such circumstances, subject to crosslinguistic parameterization.

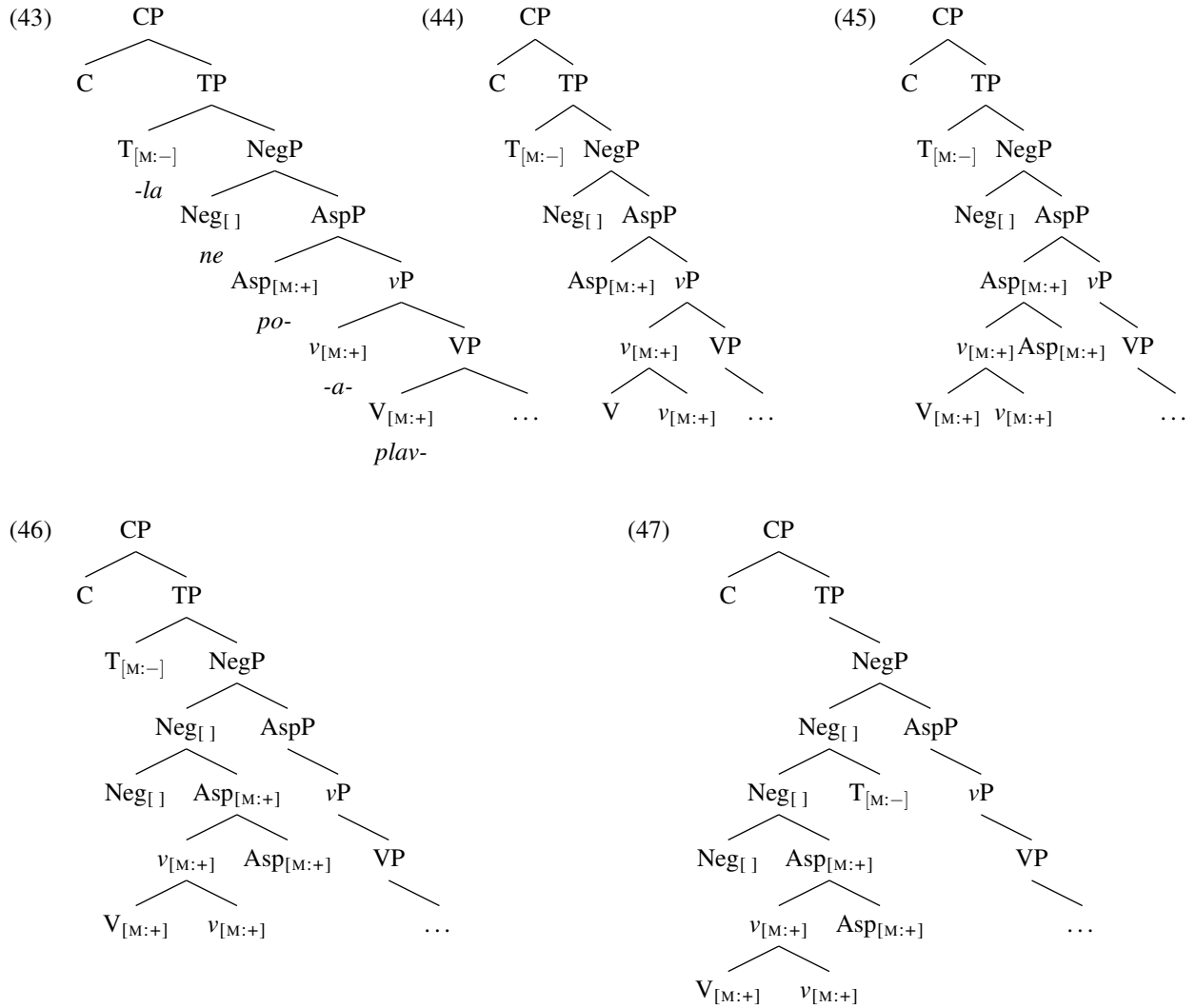
³² It might be tempting to analyze the interaction between T and V in English, whereby a bound tense affix appears to lower to the main verb (i.e., “Affix Hopping”), as an instance of amalgamation. In particular, T can be viewed as lowering into the head of its VP complement:

- (i) They enter+ed the room

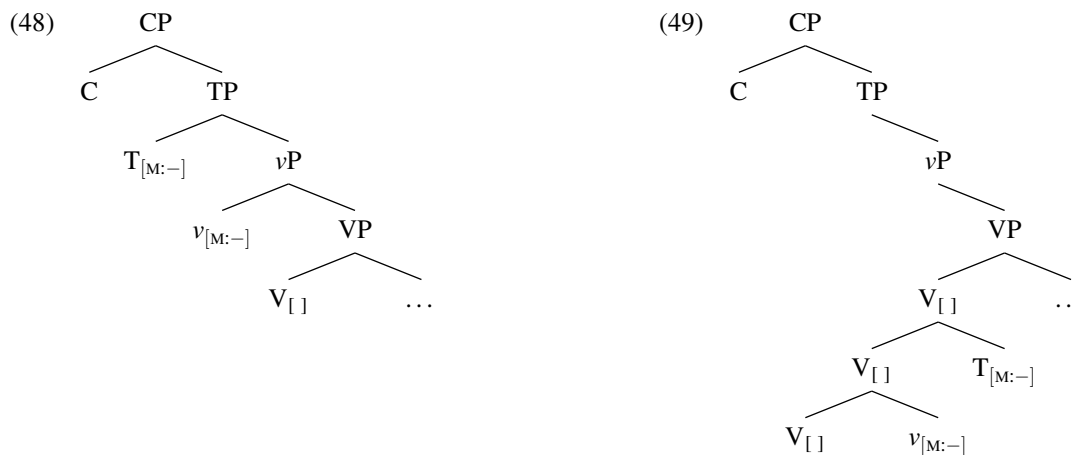
There appears to be evidence against an amalgamation analysis of Affix Hopping, however. Affix Hopping and amalgamation are subject to distinct locality constraints: Affix Hopping is known to be sensitive to the presence of overt specifiers, while amalgamation is not. For example, Affix Hopping fails to unify a main verb and a T that has undergone movement to C (in which case ‘do’-support emerges), as demonstrated in (32). In (32c), the subject intervenes between T (which occupies a derived position higher than the subject) and V, but there is no intervening head.

- (ii) a. *entered they the room?
 b. *they entered the room?
 c. did they enter the room?

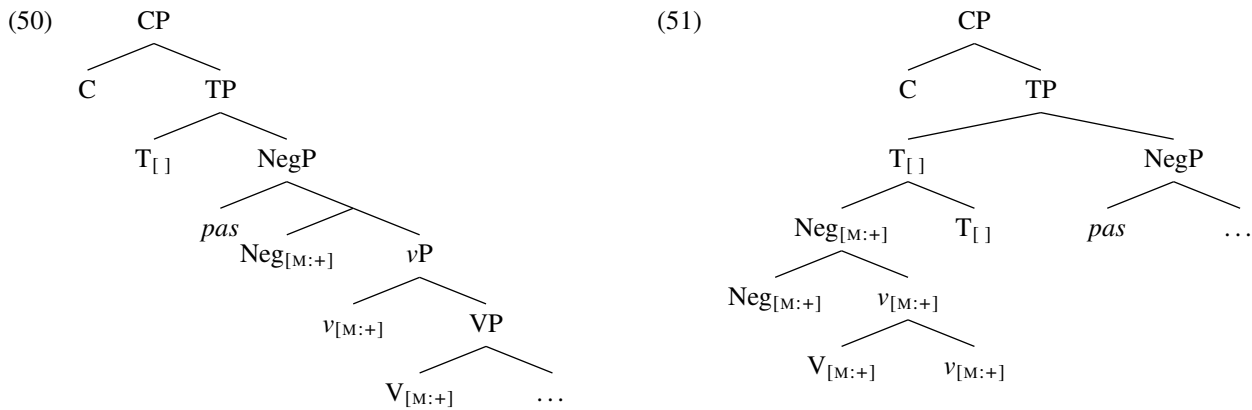
Amalgamation, on the other hand, is only sensitive to intervening heads but not any intervening specifiers (or adjuncts). This suggests that whatever mechanism is responsible for unifying T and main verbs in English (e.g., Merger Under Adjacency; see Bobaljik 2002 and Harley 2013b) is distinct from amalgamation.



The same set of tools will also produce the Danish-like pattern, in which the verb stays low relative to T while bearing the morphophonological reflex of T, and the French-like pattern, in which the verb is pronounced in T in finite contexts. Assuming a structure for the Danish clausal spine along the lines of Mikkelsen 2015 (though simplified), in the absence of auxiliaries we have amalgamation of V, v and T, where the resulting complex head is pronounced in V:



The opposite pattern—as in French, in which the Neg-V-v-T amalgam is pronounced in T—comes about as the relevant heads postsyntactically raise to T. For concreteness, we use approximately the structure from Pollock 1989, replacing the Agr projection with *v*:



Within the present approach, variation in where along the extended projection an amalgam is pronounced results from variation in which of the relevant lexical items are specified with which values of the [M] feature.³³

Another beneficial consequence of the present approach is that the bottom-up, cyclic nature of the postsyntactic derivation yields as a default the observation that complex heads created by amalgamation tend to obey the “mirror generalization”—the order of syntactic composition is mirrored in the affix ordering of the resulting morphological complex (Baker, 1988). A top-down application of similar operations would predict regular occurrences of anti-scopal affix order—exactly the opposite of the prevailing tendency crosslinguistically. In addition, the Head Movement Constraint is embedded in the definition of both *Lowering* and *Raising*.³⁴ Finally, the definition of both *Lowering* and *Raising*—in particular, that the input to both *Lowering* and *Raising* involves two structural adjacent heads—yields the ban on excorporation (Baker, 1988), the observation that heads cannot be moved out of complex heads.

Summing up, we have implemented a method for uniting disparate syntactic heads morphologically, via a mechanism which takes the output of syntax and creates head-adjunction structures. This mechanism applies bottom-up and is driven by the binary morphological feature [M]: if a head undergoes amalgamation, it must be specified with a value for [M] and the value determines the direction of amalgamation (*Raising* vs. *Lowering*). A natural question to ask is how this mechanism interacts with genuinely syntactic head movement of the T-to-C type; this question is addressed in the following section.

4 The interaction between postsyntactic amalgamation and syntactic head movement

As discussed in connection with verb-second clauses in Danish (section 2), postsyntactic amalgamation can co-occur with genuinely syntactic head movement. Here, we explore the nature of the interactions between amalgamation and syntactic head movement, and demonstrate how our theory models these interactions. The model dictates that all such interactions are limited to cases where the output of syntactic head movement serves as the input to amalgamation. The reason for this is that the former takes place in the syntactic component of grammar, the latter takes place at PF, and within the present framework of assumptions, the output of syntax constitutes the input to PF (and the output of PF is not fed back as the input to syntax).

As far as the nature of postsyntactic amalgamation is concerned, we assume the mechanisms described above in section 3. As far as the nature of syntactic head movement is concerned, we assume that it involves Internal Merge in

³³ As discussed in the main text above, the distribution of the values of the [M] feature is a matter of lexical specification and is, thus, language-specific. In other words, a lexical item can be specified as either containing [M] with a particular value or not. It follows from this that only specifications that lead to convergence would be actually attested.

³⁴ We leave for future work the possibility of deriving this part of the definition of *Raising* and *Lowering* and, thus, the relative locality constraint that amalgamation is subject to, from some deeper property of the PF component of grammar. See also footnote 26.

the narrow syntactic component of grammar. Without imposing any stipulations on what kinds of elements Internal Merge can apply to, we expect it to apply to heads and phrases equally. This is especially the case given that, in Bare Phrase Structure, the distinction between X^0 and XP is reduced to contextual relations (Chomsky, 1995a). Furthermore, the Extension Condition dictates that all instances of Internal Merge, including Internal Merge of strictly minimal items, involve merger with the root. Given this much, and assuming that the target of movement (i.e., the root) projects, syntactic head movement would be essentially head movement to a specifier position (cf. Kayne 1991, Fukui and Takano 1998, Toyoshima 2001, Matushansky 2006, Vicente 2007, Harizanov 2016).³⁵ It may, in addition, be possible for the moved head itself—rather than the target of movement—to project. In that case, syntactic head movement would result in what has been dubbed “reprojective” head movement (Iatridou et al. 1999, Izvorski 2000, Koenenman 2000, Fanselow 2003, 2009, Surányi 2005, Donati 2006, Donati and Cecchetto 2011).³⁶ Other options may be possible as well, but we do not pursue the issue here, since it does not affect the main argument for the bifurcation of head movement phenomena.

As far as the trigger of syntactic head movement is concerned, we would like to maintain that Internal Merge, whether of a nonminimal element or a strictly minimal one, is triggered in the same way (cf. Pesetsky and Torrego 2001, Matushansky 2006). This can be implemented in terms of feature valuation, as is standard: a syntactic object moves only when attracted by a feature with the EPP property (the probe)—in essence, a type of selectional feature that can impose certain requirements on (a subset of) the features of the goal (such as category or other features). However, other implementations are possible and consistent with our main point: whatever triggering options are available for syntactic phrasal movement should also be available for genuinely syntactic head movement.³⁷

If all movement is triggered by features with the EPP property, as suggested above, the question arises of why syntactic head movement is possible in the first place, given that (i) heads and the phrases they head are standardly assumed to be featurally identical, and (ii) the phrasal node is uniformly closer to the probe than the head of the phrase.³⁸ Rather than stipulating that a probe can search not just for particular features on a goal but also for a goal in a particular structural context (i.e., whether the goal is strictly minimal or maximal), we believe that, in the current theoretical context, a number of more principled routes to head movement appear to be available. It is furthermore worth noting that they are not mutually exclusive and could, in principle, be instantiated simultaneously.

First, it may be the case that head movement becomes available in a given configuration if phrasal movement is ruled out for an independent reason in that configuration. One possible candidate for the relevant kind of independent reason might be a constraint like Anti-locality (Abels 2003b, Grohmann 2003). Suppose that X is probing for a feature on its complement but that movement of the complement of X to Spec,XP is excluded by Anti-locality. It is then just in this case that the head of X’s complement, would be attracted to Spec,XP instead (for discussion of similar proposals, see Harizanov 2016 and Preminger to appear). More generally, whether XP- or X-movement takes place might be determined by general conditions that force, allow, or prohibit pied piping in a given configuration (Heck 2008, 2009, Cable 2012, Landau 2006).

A second way in which attraction may result in head movement involves relaxation of assumption (i) above. In the case of long head movement in Bulgarian, for example (discussed in section 2.3 and in Harizanov 2016), the standard interpretation of the facts is that the verb, rather than the VP, is fronted; this fronting is not discourse neutral (Embick and Izvorski 1997b, Lambova 2004, Franks 2008). Given this, the attraction may be contingent on the presence of a discourse relevant feature on V, and it must be the case that V but not VP bears the relevant discourse feature. This, in turn, must mean that while certain features (such as category features) do percolate to the phrasal node, not all features

³⁵ As a reviewer points out, the proposal that syntactic movement of a head can target a specifier position is incompatible with Chain Uniformity: “A chain is uniform with regard to phrase structure status” (Chomsky 1995b:253). For further discussion, see Fukui and Takano 1998, Toyoshima 2001, Matushansky 2006, Vicente 2007, Harizanov 2016. We take it that removing the Chain Uniformity stipulation is, in general, desirable but recognize that the empirical effects that necessitated it in the first place need to be explained in different and hopefully more basic and principled ways.

³⁶ Harizanov (2017) explores and empirically motivates an approach to syntactic head movement whereby both options—head movement to a specifier position and reprojective head movement—are instantiated.

³⁷ Further questions arise about the triggering mechanisms involved in reprojective head movement, which are discussed in detail in the rich literature on reprojection (e.g., Koenenman 2000, Fanselow 2003, 2009, Surányi 2005, Donati 2006, Donati and Cecchetto 2011). We thank a reviewer for raising these important questions about the triggering mechanisms involved in syntactic movement.

³⁸ Since this scenario involves attraction of a head by a probe, it only concerns head movement to a specifier position rather than reprojective head movement.

(must) do so. In fact, it is possible that discourse related features in general do not percolate since whether a phrase or its head bears the relevant feature has clear interpretive consequences: there is a need to formally distinguish, for example, a focused VP from a focused verb.

Finally, Preminger (to appear) relates the availability of movement of a head Y to a probe X in a particular configuration to whether X has entered into a prior syntactic relationship with YP. In particular, if no prior syntactic relationship between X and YP has been established, it is YP that undergoes movement to Spec,XP because YP is closer to the probe than its head Y. If, on the other hand, X has entered into a prior syntactic relationship with YP, only Y can be attracted to X. For Preminger, this follows from the interaction of two principles: the Principle of Minimal Compliance and Minimal Remege.

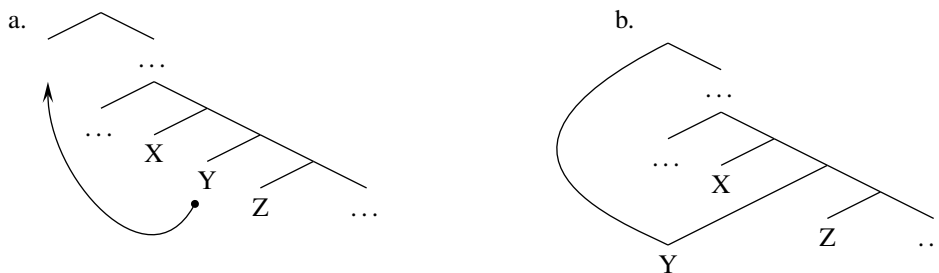
- (52) Principle of Minimal Compliance (Preminger to appear; see also Richards 1998):
Once a probe P has successfully targeted a goal G, any other goal G' that meets the same featural search criterion, and is dominated or c-commanded by G (=dominated by the mother of G), is accessible to subsequent probing by P irrespective of locality conditions.
- (53) Minimal Remege (Preminger to appear):
If X^0/X_{min} is movable, move only X^0/X_{min} .

Given these principles, once a syntactic relationship between X and YP has been established, the head Y of YP becomes accessible to probing by X due to the Principle of Minimal Compliance. In turn, Minimal Remege dictates that, if movement of Y can apply, it must. Together, these two principles then force Y (as opposed to YP) movement to X just in case X has entered into a prior syntactic relationship with the phrase that Y is the head of. In cases of local head movement, Preminger argues that the relevant prior syntactic relationship is (c-)selection; in cases of nonlocal head movement, it is agreement.

4.1 Possible derivations and concrete instantiations

In this section, we delineate the space of possible derivations involving both syntactic movement and amalgamation, discuss some of those possibilities in detail to illustrate the workings of the system, and provide a discussion of certain expectations that arise from the formulation of these mechanisms and their interaction. Consider first a configuration in which a head Y (which is a strictly minimal element; i.e., the head of a branching phrase) undergoes syntactic head movement, as in (54). This syntactic head movement generates a structure that contains two occurrences of the head Y, in its base generated (first-merge) position and one in its derived (second-merge) position:³⁹

- (54) Syntactic movement of Y

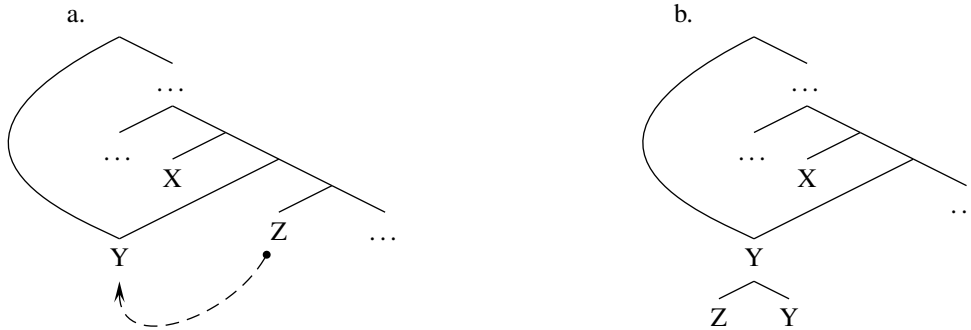


In this structure, X and Z are heads distinct from Y and from one another. Their relationship with Y is such that Y is the head of X's complement and Z is the head of Y's complement. Given this configuration, at PF, it is possible for

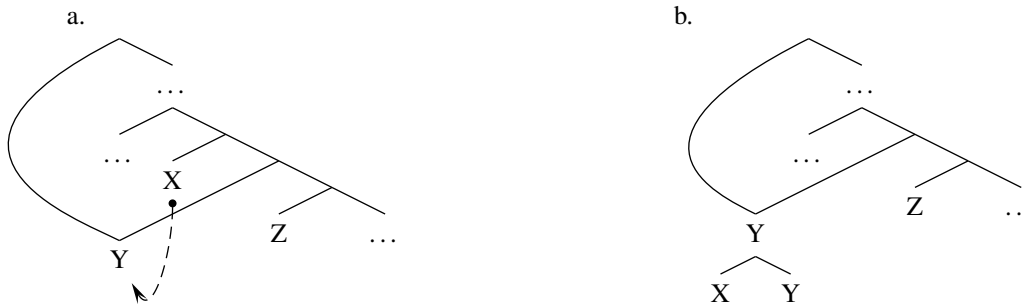
³⁹ We limit the discussion here to one-step movements (i.e., movement chains with two links only). Everything that we say can be extended trivially to multi-step movement and movement chains composed of more than one link. We also represent distinct occurrences of a single syntactic object in the structure using multidominance. This is an expository choice that emphasizes the fact that we are dealing with a single syntactic object occupying distinct structural positions, but other implementations are possible as well. Finally, we remain agnostic here as to whether the movement of Y is an instance of movement to a specifier position or reprojective movement (see the main text above for discussion).

either X or Z to amalgamate with Y. This would involve postsyntactic *Raising* of Z into Y or postsyntactic *Lowering* of X into Y:

(55) Post-syntactic Raising of Z into the low occurrence of Y

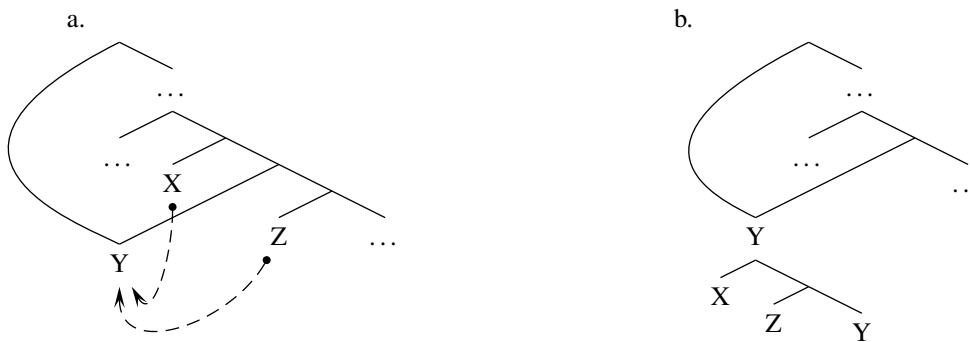


(56) Post-syntactic Lowering of X into the low occurrence of Y



It is also possible for both X and Z to amalgamate into Y. Since heads are considered for amalgamation from the bottom up, this would involve postsyntactic *Raising* of Z into Y, followed by postsyntactic *Lowering* of X into Y. The resulting complex head, of course, mirrors the order of syntactic composition of the individual heads:

(57) Postsyntactic *Raising* of Z, followed by *Lowering* of X into the low occurrence of Y



The assumption that Internal Merge creates structures in which a single syntactic object (here, Y) occupies distinct structural positions now becomes relevant. Amalgamation of X and/or Z into Y affects the single syntactic object Y,

and the effect of amalgamation—i.e., the resulting complex head—will now be associated with the same structural positions that Y was associated with prior to amalgamation. In other words, the results of postsyntactic amalgamation into a syntactic object are reflected in all of that object’s occurrences (see also Selkirk 1995; McPherson 2014; Ahn 2015; Harizanov 2018).⁴⁰ Once the relevant postsyntactic structures are generated (the (b) structures above), Chain Reduction determines which occurrences of the object Y are to be pronounced in accordance with a language specific parameter.⁴¹ For example, if the highest occurrence is pronounced in general in some language, we expect the highest occurrence of Y to be pronounced above in that language as well. On the surface, this creates the appearance of Y dragging X and/or Z along to its high position. In fact, it is cases like this that have often been assumed to suggest that amalgamation-type head movements must feed syntactic movement. However, we have presented an approach to the data—as we will see in more detail below—that explains this type of interaction without having to posit that amalgamation feeds syntactic movement.

A reviewer points out that (i) if amalgamation of a head into another syntactically moved head affects multiple occurrences of the syntactically moved head, and (ii) if multiple occurrences of the syntactically moved head are pronounced (for some independent reason), we would expect the multiple occurrences to be morphologically identical (i.e., all occurrences should reflect the result of amalgamation).⁴² Two cases of verb doubling, in which the fronted verb is an infinitive and the lower verbal occurrence is inflected, may constitute potential counterexamples to this expectation. While we agree that the system does make this prediction, neither of these two cases involves amalgamation of a head into another syntactically moved head (as per (i) above). One of the relevant cases is verb doubling in Hebrew bare infinitive fronting. However, as demonstrated in section 4.1.4 below, the derivation of Hebrew bare infinitive fronting does not involve amalgamation of a head into another syntactically moved head; instead, it involves amalgamation of a syntactically moved head (V) into another, unmoved, head (T). The other relevant case is verb doubling in Russian verb fronting configurations. However, there is evidence that the multiple verbal occurrences in verb fronting configurations in Russian do not form a simple movement chain; instead, they are the result of movement of a phrasal constituent that properly contains the verb. Consequently, Chain Reduction does not “see” the multiple verbal occurrences as occurrences of a single syntactic object and this is the reason they end up with distinct morphophonology (for details, see Harizanov and Gribanova 2017).

We next present case studies that instantiate each of the three types of interactions between postsyntactic amalgamation and syntactic head movement that are presented here. Specifically, we show that the derivation of German verb second clauses can be viewed as an instance of (55), where the finite verb postsyntactically raises into the low occurrence of T, and that the derivation of Danish verb-second clauses can be viewed as an instance of (56), where T postsyntactically lowers into the low occurrence of V. Finally, the view of Russian clause structure presented in Gribanova 2017a can be interpreted as an instantiation of (57).⁴³

⁴⁰ Not all postsyntactic operations need behave in the same way as amalgamation. For example, a reviewer points out that van Urk’s (2015; to appear) treatment of DP doubling by a pronoun in Dinka is the result of partial deletion of one occurrence of a DP accompanied by nondeletion of another occurrence of the same DP. This is then a case where Chain Reduction, which is responsible for deletion of occurrences to varying degrees, must not affect all occurrences equally and/or in the same way. This state of affairs is consistent with a view of the PF branch of grammar where some operations (such as amalgamation) manipulate the syntactic object associated with multiple structural positions directly, while others (such as linearization and Chain Reduction) manipulate the actual positions that a particular syntactic object is associated with or the phonological content associated with those positions. It is conceivable that the former type of operation applies at an earlier stage of PF than the latter type of operation, but we leave this and other intriguing related questions open at this point.

⁴¹ Note that *Lowering* and *Raising* must derivationally precede Chain Reduction. There is independent evidence on the basis of verb doubling in Russian that this is the relative timing of amalgamation and Chain Reduction (Harizanov and Gribanova 2017).

⁴² Under the circumstances described above, what we expect is that multiple occurrences should be identical as far as the part of PF between the output of syntax and Vocabulary Insertion is concerned. This is separate from the mechanisms that govern how those occurrences will be realized phonologically, if there is doubling. There is ample evidence that phrasal movement may yield pronunciation of more than one copy, and it is well known that copies under these circumstances may or may not be segmentally identical. Given this, all we are committed to here is that the processes governing spellout of multiple copies in the case of phrasal movement should be in force in the cases we are interested in here. As we elaborate here and in section 4.1.4, however, the relevant kinds of configurations do not arise either in Russian or Modern Hebrew predicate clefting.

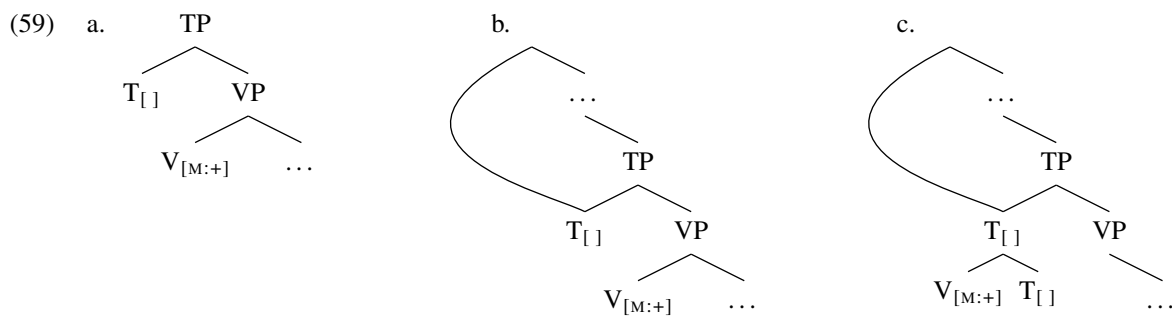
⁴³ A reviewer suggests an alternative treatment of head movement that does not rely on *Raising/Lowering*, according to which head movement is uniformly syntactic. The idea is that head movement applies in the traditional way, involving head-to-head adjunction and obeying the Head Movement Constraint (HMC), and that it creates multiple copies of the moved head(s). Given this much, different copies can be prioritized for interpretation at each of the interfaces, PF and LF. For example, it is possible that a high occurrence is interpreted at PF (i.e., pronounced) while a low occurrence is interpreted at LF and, in fact, such a derivation yields head movement with no interpretive effects. However, we think that this alternative approach misses a generalization, since it does not necessitate a connection between locality and the absence of interpretive effects—

4.1.1 The derivation of German verb second

Consider first the derivation of root clauses in German, in which the finite verb occupies the “second” position within the clause:

- (58) Dieses Auto hat Fritz in München geklaut.
 this car has Fritz in Munich stolen
 ‘Fritz stole this car in Munich.’

As discussed in section 2.1, in German root clauses, T undergoes syntactic movement, yielding structure (59b) on the basis of structure (59a). These structures do not represent precedence relations and specifiers/adjuncts are not included. In addition, while we acknowledge that T moves to the C-domain, we remain agnostic as to the exact position it comes to occupy as the result of this movement.⁴⁴ Structure (59b) serves as the input to the postsyntactic derivation (PF). At PF, we take it that V is specified as [M:+], and it undergoes postsyntactic *Raising* to (i.e., amalgamates with) the low occurrence of T, yielding structure (59c).



Finally, the complex T is pronounced only in the highest position in which it occurs (i.e., in the C-domain) in accordance with the general principles that govern Chain Reduction in German.

If the clause contains multiple verbal elements, such as auxiliaries, in addition to the main verb, it is the highest verbal element that amalgamates with T in German. In that case, the [M:+] feature, which triggers raising to T, must be on the highest auxiliary rather than the main verb. What ensures that the [M:+] feature will be associated with the highest verbal element in German and not with any of the lower verbal elements? Under the present approach, the solution depends on whether it can be maintained that V raises to T in German. Suppose first that V does not actually amalgamate with T by raising to T in German. On this approach, German would be just like Danish in that T instead lowers to the closest verbal element (see section 4.1.2). This analysis of the amalgamation between T and V in German is consistent with the absence of evidence for V-to-T raising in German and receives direct support from positive evidence that V does not, in fact, raise to T (Haider 2010).⁴⁵

Alternatively, German may be (and has traditionally been) viewed as being unlike Danish in that (the highest) V does indeed raise to T—this is the view adopted in (59c). One analysis of this state of affairs, afforded by the present approach to amalgamation, involves allowing the [M:+] feature to either appear on a verb or not. This analysis, however, overgenerates by incorrectly predicting that main verbs will amalgamate with higher, overt auxiliaries. An

namely, why should the HMC-compliant head movement (for us, word formation) consistently lack interpretive effects while long distance head movement can have interpretive effects?

⁴⁴ Two reviewers ask the very interesting question of how our theory can account for the oft observed complementary distribution between verb movement to the C domain and the presence of an overt complementizer. The nature of the account crucially depends on the proper treatment of the genuinely syntactic head movement in the present theoretical context. As suggested in the introduction to section 4, two possibilities—which may in fact coexist—are head movement to a specifier position and reprojective head movement. If V (or T) movement to the C domain is an instance of head movement to a specifier position, no immediate explanation is available of the complementarity between V (or T) movement and the presence of an overt C. On the other hand, a treatment of V (or T) movement to the C domain as reprojective head movement allows for an understanding of this complementarity: in the derivation of a clause in which V (or T) moves and reprojects, there is no independent C head that is ever merged (Fanselow 2003, 2009, Müller 2011). Note, finally, that under the canonical syntactic head adjunction approach, these facts find no explanation, since there is no inherent reason that head movement should be triggered only by a null C but not by an overt one.

⁴⁵ If it is T that lowers to V in German, it must be V that is targeted by the movement to the C domain, as in Danish (see section 4.1.2).

alternative is to assume that a verbal element can be associated with either $[M: +]$ or $[M: -]$ in German and that certain Ts (like the past and present T) select upon External Merge for the $[M: +]$ version while other Ts (like modals) and auxiliaries select for the $[M: -]$ version. This approach captures the intuition that some Ts, in a certain sense, “attract” the verbal head of their complement.⁴⁶ Another way to capture this intuition directly is to assume that a morphological selectional feature on a head H can cause another head to amalgamate into H in a deviation from the apparatus proposed in section 3. We leave open the question of whether this latter, less restrictive conception of the relevant morphological selectional features is empirically warranted.

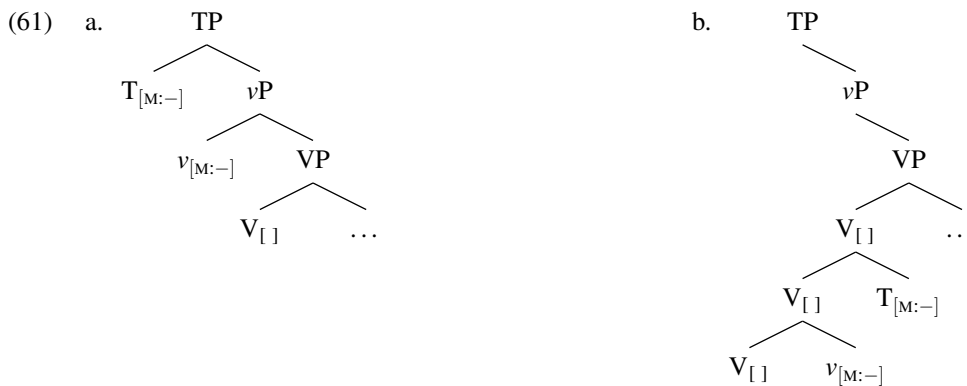
4.1.2 The derivation of Danish verb second

Consider now the question of how verb-second orders are derived in Danish, a language for which there is no evidence for verb movement to T.

- (60) Peter **drikker** often kaffe om morgonen.
 Peter drinks often coffee in morning.DEF
 ‘Peter often drinks coffee.’

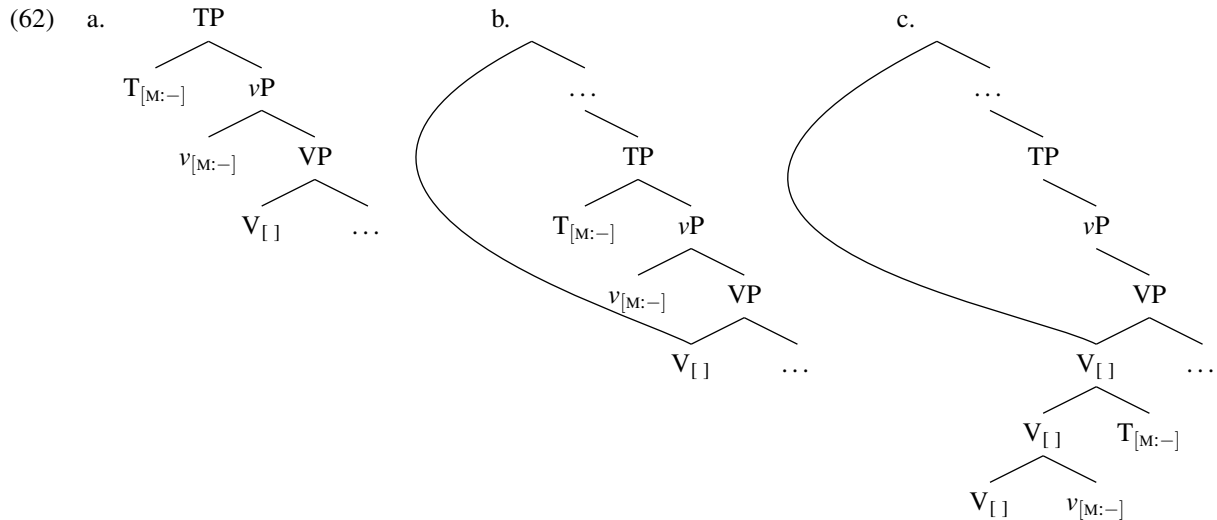
(Vikner 1995:47, (33c))

In Danish embedded clauses, the verb remains low, presumably in its VP-internal position, while finite T lowers to it, as discussed in section 2.1. This indicates that finite T is specified as $[M: -]$, which triggers postsyntactic *Lowering* to v ; there is also postsyntactic *Lowering* of v into V, triggered by the $[M: -]$ feature on v :



Danish root clauses, on the other hand, exhibit verb-second order. Within our framework of assumptions, the derivation of such clauses involves, first, syntactic movement of V to the C-domain (again, we remain agnostic as to the exact position targeted by this movement). The second step in the derivation involves the same two *Lowering* operations (T to v and v to V) observed in embedded clauses. That is, the syntactic movement of V is followed by postsyntactic *Lowering* of finite T and of v to the low occurrence of V:

⁴⁶ A reviewer suggests that such an account arguably amounts to a conspiracy. Be that as it may, it is worth pointing out in this connection that the traditional account in which V-to-T movement involves movement of V to v and then of v to T, likewise appears to involve a conspiracy. Specifically, as pointed out by Müller (2011) in a slightly different context, according to this account T attracts v (not V) and only v attracts V, so that the existence of full V-to-T movement emerges as a “fortunate coincidence” (ibid.).



The resulting complex V is pronounced only in the highest position in which it occurs (i.e., in the C-domain) in accordance with the general principles that govern Chain Reduction in Danish.⁴⁷

This treatment is based on an assumed distinction between the position of verbs in non-V2 clauses in Danish vs. German: consistent with the discussion in §4.1.1, we take the position of the verb in German non-V2 clauses to be T, whereas its position in Danish verb-second clauses is V.⁴⁸ To the extent that such a differentiation is justified, it is reflected in our analysis via differences in the values of two features which are subject to crosslinguistic variation. The first concerns the head that undergoes syntactic movement (T in German, V in Danish) and the second concerns the head that undergoes postsyntactic amalgamation (V undergoes Raising to T in German, T undergoes Lowering to *v* and *v* undergoes Lowering to V in Danish).

4.1.3 The derivation of Russian VSO and SVO clauses

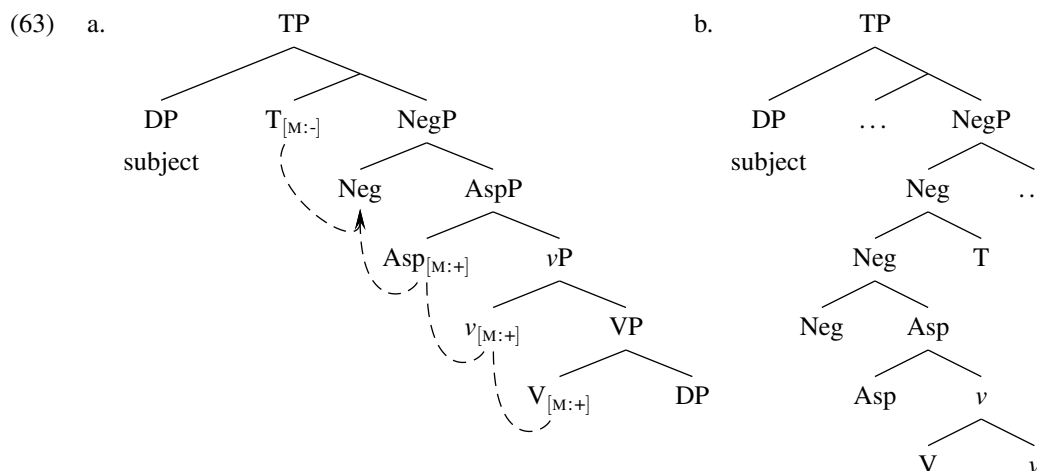
There are also cases in which there is both postsyntactic *Lowering* and *Raising* into the lower occurrence of a moved element (where the higher occurrence is pronounced). These are cases like (57), which have the appearance of postsyntactic amalgamation “feeding” syntactic head movement. In the current proposal such effects are achieved via careful derivational timing—crucially, the *Lowering* and *Raising* operations must derivationally precede Chain Reduction (see footnote 41 and surrounding discussion).

The discussion of Russian in section 3 already brought up the case of Russian SVO clauses, which on the view entertained here will be derived via postsyntactic applications of both *Lowering* and *Raising*. While T, which is associated with tense and agreement information, is morphologically dependent, we have already seen in section 3 that the verbal complex does not raise as far as T, although it raises away from its base position (“short verb movement”). So T is joined with the rest of the complex, but there is evidence against the idea that the verbal complex is pronounced in that position; this is one of the pieces of evidence we used in section 3 to argue for a dissociation between syntactic movement and morphological amalgamation. We can now understand this effect as a combination of *Lowering*

⁴⁷ If the clause contains multiple verbal elements, such as auxiliaries, in addition to the main verb, it is the highest verbal element that undergoes syntactic movement to the V2 position in Danish. If this movement is of the reprojective kind, as suggested in footnote 44 and by Fanselow (2003, 2009) and (Müller, 2011), among others, the fact that the highest verbal element moves can be derived as follows. Assume with Müller 2011 that a verb can be optionally endowed with the feature that triggers reprojective movement and that, in this case, this feature triggers Internal Merge of a verb with TP. Given this, it is the closest verb to TP in terms of c-command (i.e., the highest one) that gets to undergo this movement. If the lower of two verbs is specified for the relevant feature but the higher one is not, the higher verb still acts as a defective intervener in that it blocks the movement of the lower verb but it cannot itself move, resulting in a crash (since the movement triggering feature on the lower verb cannot be satisfied in such a configuration).

⁴⁸ The evidence in favor of this claim may be quite subtle, and we do not elaborate on the distinction here. If the distinction between German and Danish non-V2 clauses collapses empirically, then it is the case that both languages should receive roughly the same analysis, namely the one we present here for Danish.

and *Raising*—represented by dashed lines below—using the clausal structure independently argued for in Gribanova 2017a:



Such a configuration yields *SVO* clauses, assuming that the subject occupies *Spec,TP* on the surface (see Gribanova 2017a for justification). Russian also permits *VSO* orders under polarity focus:⁴⁹

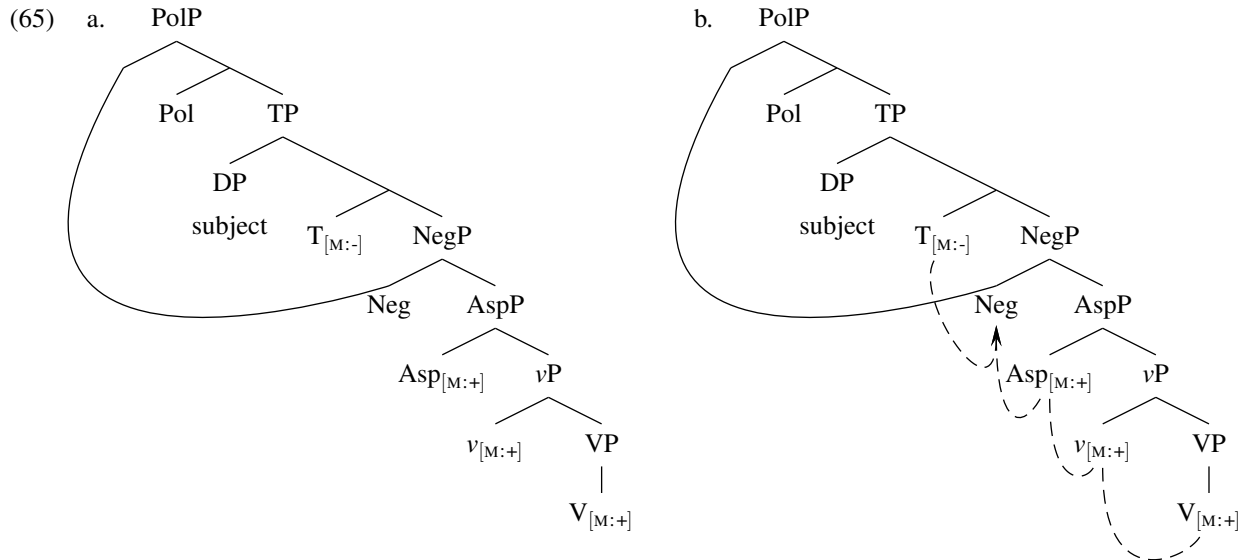
- (64) a. Ja v vojnu pil tože kakoj-to. V Germanii. Klopami paxnet.
 I in war drank also some-kind in Germany bedbugs.INSTR.PL smells.3SG
 ‘I also drank some kind during the war. In Germany. It smells of bedbugs.’ (Gribanova 2017a:1091, (23))
- b. Da ne paxnet on klopami!
 PRT NEG smell.3SG it.NOM bedbugs.INSTR
 ‘No, it doesn’t smell of bedbugs!’ (‘it’ = cognac)⁵⁰

For such cases, Gribanova 2017a demonstrates that there is an additional, syntactic movement step from *Neg* to a *Pol* head just above *T*.⁵¹ For Gribanova, there is an independently motivated *AGREE* relation between the *Pol* and *Neg* heads, and this syntactic dependency feeds syntactic movement of *Neg* to *Pol*. On our implementation, this pattern will be derived via syntactic movement of *Neg* to *Pol* (65a), followed by postsyntactic *Lowering* and *Raising* into *Neg* (65b).

⁴⁹ The unstressed, clause-initial *da* particle is a discourse particle, orthographically identical to (but semantically distinct from) the stressed particle that means “yes”.

⁵⁰ Vasilij Šukšin. *Kalina krasnaja*, 1973.

⁵¹ Evidence that the movement is, indeed, syntactic, comes from the observation that this movement step triggers *MaxElide* effects in Russian verb-stranding ellipsis (Gribanova, 2017a).



When the above-elaborated assumptions about Chain Reduction and copy pronunciation are adopted, this combination of operations will give rise to an effect wherein it appears that all of the amalgamated parts are “dragged along” to Pol. The result, in Russian, is a surface effect in which raising of the verb to a high position in the clause yields VSO orders and the related polarity focus effect (on which see Gribanova 2017a).

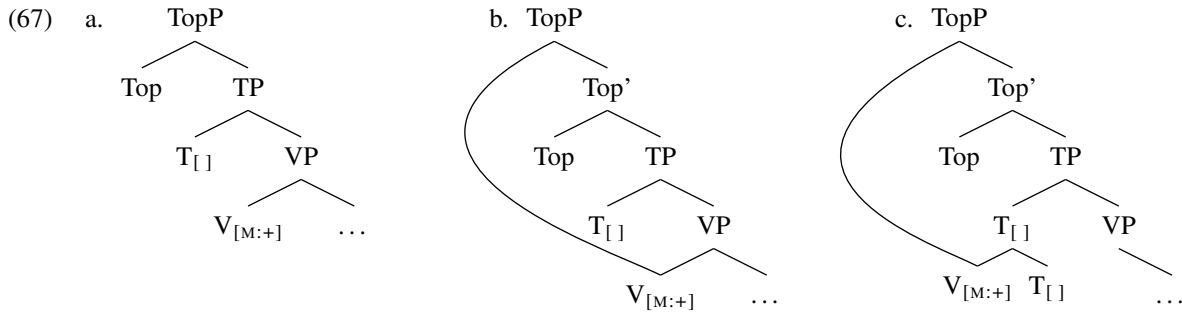
4.1.4 The derivation of bare infinitive fronting in Hebrew

The cases considered so far all involve a head (or heads) amalgamating into an occurrence of a head that has undergone syntactic head movement. However, it is also possible for a syntactically moved head to amalgamate into some other head. Consider the following example of bare infinitive fronting in Hebrew (Landau 2006). As discussed in section 2.3.3, bare infinitive fronting involves verb doubling: the low occurrence of V spells out tense and agreement features while the high occurrence of V appears as an infinitive.⁵²

- (66) Liknot, hi kanta et ha-praxim.
 to-buy she bought ACC the-flowers
 ‘As for buying, she bought the flowers.’ (Landau 2006:37, (8b))

We implement Landau’s (2006) analysis by assuming that the verb undergoes syntactic head movement, yielding structure (67b) on the basis of structure (67a) (specifiers are not included in these structures for expository reasons). Structure (67b) then serves as the input to the postsyntactic derivation (PF). At PF, since V is specified as [M:+] in Hebrew, it undergoes postsyntactic *Raising* to (i.e., amalgamates with) T, yielding structure (67c).

⁵² There is some variability across languages with respect to the degree to which the two verbal elements in a predicate cleft pattern must match. Harbour (1999) makes clear that the analogous pattern in Classical Hebrew makes possible a greater degree of mismatch than Modern Hebrew; this includes mismatches in the derivational (voice) morphology. Classical Hebrew is further of relevance to us for the way in which it apparently involves a nontrivial interaction between what appears to be syntactic and postsyntactic head movement (Harbour, 2007). A proper investigation of these very important patterns will have to be left for another time.



Finally, Chain Reduction applies to (67c). The occurrence of V in Spec,TopP is pronounced because it is the highest occurrence of V, and it is spelled out as an infinitive because it is an occurrence of just the verb, without T.⁵³ We follow Landau (2006) in assuming that the low occurrence of V, adjoined to T, is pronounced because it “is associated with the phonological requirement of T⁰—namely, the need to spell out tense and agreement features, or the ‘Stray Affix Filter’.” Thus, while there would normally be a pressure to delete a low occurrence, in this case, the affixal character of T overrides this pressure, resulting in double pronunciation of V.⁵⁴

It is worth noting that, while the reasons for verb doubling in Hebrew bare infinitive fronting are independent of syntactic head movement and amalgamation—i.e., they have to do with P-recoverability (i.e., the phonological requirement of T), as Landau (2006) suggests—syntactic head movement of V and its amalgamation into T in Hebrew create the configuration that is a prerequisite for doubling to occur. A further prediction of the proposed system about when multiple pronunciation of a given head is expected is discussed in the following section.

4.2 Further predictions

Assume first that some head Y undergoes syntactic movement, which can involve a single or multiple steps, yielding 2 or more occurrences of Y in the structure. When amalgamation applies, at PF, there are a number of options that are available as to how the derivation can proceed. First, amalgamation can apply to Y in either of its occurrences generated by syntactic movement. Second, amalgamation can be either inbound (i.e., some head amalgamates into the head Y) or outbound (i.e., the head Y itself amalgamates into some other head). Third, amalgamation can involve either lowering or raising. Finally, Chain Reduction marks some of the occurrences of Y for nonpronunciation according to language-specific properties. These options can be summarized as follows:

1. Options having to do with syntactic movement
 - a movement chain with 2 occurrences vs. a movement chain with 3 occurrences vs. ...
 - a movement chain with *n* occurrences
2. Options having to do with postsyntactic amalgamation
 - inbound amalgamation (i.e., of some head into the moved item) vs. outbound amalgamation (i.e., of the moved item into some other head)
 - lowering vs. raising

⁵³ Alternatively, Landau (2006) proposes that the occurrence of V in Spec,TopP is pronounced because it “is associated with a phonological requirement imposed by Top⁰, namely, the characteristic intonation of fronted VPs.”

⁵⁴ As pointed out to us by Daniel Harbour (p.c.), this same explanation cannot hold of predicate cleft patterns in Classical Hebrew, where present tense participles bear no tense morphology, but are nevertheless doubled (Harbour, 2007). We leave the very interesting question of what motivates doubling in this case for a separate investigation.

3. Options having to do with Chain Reduction

- pronounce the 1st occurrence (i.e., the lowest) vs.
pronounce the 2nd occurrence vs.
- ...
- pronounce the n^{th} occurrence (i.e., the highest)

In sum, there are total of 4 options, having to do with the syntactic movement step, with the amalgamation step, and with the Chain Reduction step.⁵⁵ It is also the case that the mechanisms proposed here interact with other, independently motivated Distributed Morphology mechanisms such as Merger Under Adjacency, among others.

In our estimation, computational modeling will be essential in identifying the complete range of typological predictions made by this system, especially once it is embedded in a broader theory of the PF component of grammar; we aim to take on such computational work next. Here, we instead discuss some specific patterns that we expect to find (or not) on the basis of our understanding of the interaction of amalgamation and syntactic head movement (either of the reprojective or movement-to-specifier variety). While each of these predictions deserves far more than the one paragraph we provide, we intend for this discussion to serve as a starting point for further work on these issues.

The first prediction concerns an absolute locality constraint that amalgamation is subject to but syntactic head movement, in general, is not. The standard assumption that PF computations are restricted to apply within cyclic domains defined by the syntax (i.e., phases), in conjunction with the assumption that amalgamation applies at PF, predicts that amalgamation cannot apply across phase boundaries. To the best of our knowledge, amalgamation—i.e., the kind of “head movement” that results in word formation—systematically applies within phases but not across phase boundaries (assuming that the head of a phase, e.g., C, and its complement are not separated by a phase boundary). It is indeed a long-established observation that this type of head movement does not cross extended projections; that is, it does not cross CP and DP boundaries (Grimshaw 1991). This observation can be derived from the amalgamation mechanism and its interaction with phase-based spellout (Chomsky, 2000).⁵⁶ In contrast, syntactic movement can (but does not have to) cross phase boundaries. Thus, genuinely syntactic head movement is predicted to be able to move from one phase onto another and, as demonstrated in section 2.3, cases of long head movement such as those in Bulgarian and Hebrew behave accordingly.⁵⁷ A further consequence of this view is that at least certain types of noun incorporation are necessarily syntactic: in particular, if N incorporates into V, stranding DP-internal material behind, the displacement of N crosses a phase boundary (see Jenks 2013 for another case of apparently cross-phasal N movement). Such cross-phasal displacement, on the view outlined here, can only be modeled as an instance of genuinely syntactic head movement.

A second prediction concerns the relative locality constraint that amalgamation is subject to: *Raising* and *Lowering* are defined to apply between structurally adjacent heads. As a result, under the standard assumption that specifiers and adjuncts are islands, (i) amalgamation of two heads can apply across intervening adjuncts and specifiers, and (ii) neither *Lowering* nor *Raising* can target material in an adjunct or a specifier. A similar point is made by Embick and Noyer (2001), who leverage this observation for *Lowering* in their case study on the realization of the Bulgarian definite marker, which is a suffix on the leftmost adjective or noun in a DP but not on intervening demonstratives or degree words. To the extent that we do find a head H forming a complex head with a head other than the head of H’s

⁵⁵ 2 and 3 above are language-specific parameters having to do with the featural content of particular lexical items (as far as amalgamation is concerned) and particular facts about how Chain Reduction operates in any given language. 1, on the other hand, simply concerns the particular derivation and how many movements have taken place. The four options above result in $2 \times 2 \times n^2$ derivations involving just 2 heads, one of which is the syntactically moved head. If there are multiple (i.e., >1) heads that amalgamate with the syntactically moved one, there will be additional possible outcomes. The particular formulation of amalgamation in section 3 predicts only a subset of all imaginable outcomes to be possible—see below for further discussion.

⁵⁶ Given that there are instances of amalgamation in Danish and Russian which appear to cross the vP boundary and that vP has been argued to be a phase at least in some languages, the question arises as to how to understand the relevant facts in Danish and Russian. One possibility, which is in principle empirically distinguishable from the alternative, is that vP is not a phase in these languages and that is why amalgamation is able to apply across the vP boundary. In fact, we know of no positive evidence to the contrary. If, on the other hand, vP is a phase in Danish and Russian, that might indicate that the relevant locality domain for amalgamation is the extended projection, rather than the phase. In turn, this would mean that amalgamation cannot cross DP and CP boundaries, as required by the empirical facts.

⁵⁷ Incidentally, there are many putatively syntactic movements (e.g., most A-movements) that are clause-bounded (i.e., they cannot cross finite clause boundaries). Therefore, the clause-boundedness of some instance of displacement cannot by itself lead to the conclusion that this displacement is not syntactic. As expected, then, some syntactic head movements (e.g., Danish V-to-C) are clause-bounded while others (Bulgarian and Hebrew long head movement) are unbounded.

complement, this must arise from a process other than amalgamation (see, for example, Harizanov 2018, Bennett et al. 2015 for such cases).

Third, the discussion about interpretive consequences of head movement in §2.2 has a PF counterpart. Our empirical claim in that section is that, to the extent interpretive effects of head movement are ever found, such effects consistently hold of the genuinely syntactic type of head movement, and not of amalgamation. On the PF side, it is again the genuinely syntactic type of head movement, and not amalgamation, that may result in pronunciation of multiple occurrences of the displaced head (e.g., verb doubling). In this connection, it is important to note that it is not syntactic head movement per se that yields doubling—presumably, doubling is the result of PF pressures on the pronunciation of a syntactic object’s occurrences—instead, syntactic head movement is one prerequisite for doubling because it creates multiple occurrences. In the proposed system, the reason amalgamation does not result in multiple pronunciation is that, since it is not Internal Merge, amalgamation does not create multiple occurrences (see section 3). This, in turn, ensures that a head that has undergone only amalgamation cannot be interpreted in multiple structural positions at PF or LF. Thus, if simultaneous (partial) pronunciation of multiple occurrences of a given head are observed, syntactic movement has to have been involved (e.g., as in the predicate clefting examples discussed in §2.3.3).

Fourth, all else being equal, Chain Reduction is expected to apply in the same manner to all syntactic movement in a language regardless of whether it involves movement of a branching phrase (a nonminimal syntactic object) or a head (a strictly minimal syntactic object). That is, whatever the language-specific conditions are on the pronunciation of occurrences of a moved syntactic object in a language, we expect those same conditions to hold both for phrasal movement and for syntactic head movement. Thus, all other things being equal, if the default is “highest copy pronunciation”, we expect that an instance of syntactic head movement (e.g., long head movement) in the language will not result in the pronunciation of a lower occurrence than the highest.

The fifth (and final) observation we wish to make here concerns the various clusterings of properties that we expect to see associated with amalgamation. Our claim has been that the effect of morphological growth results directly from the application of amalgamation, but not from the application of genuinely syntactic head movement. If we see morphological growth of the relevant kind, involving only heads in an extended projection, then we expect morphological growth to exhibit locality effects associated with the Head Movement Constraint (which follows from the definition of *Lowering* and *Raising*). Therefore, if we find morphological growth in combination with the skipping of heads, this can only be the result of both amalgamation and syntactic head movement, not amalgamation alone. Consider the configuration below, where H1 has two occurrences, one high (leftmost) and one low (rightmost) (i.e., it has undergone syntactic movement):

(68) H1 H2_[M:+] H3_[M:+] H4 H5 H6_[M:-] H1

This will result in a complex consisting of [[H1 H6] [H2 H3]], skipping H4 and H5, which are inert. This kind of configuration can only be derived if there is a higher and a lower occurrence of H1 to begin with.

5 Conclusion

In the discussion so far, we have (i) presented a series of empirical and theoretical arguments in favor of a distinction between genuine syntactic head movement and postsyntactic amalgamation, and (ii) provided an implementation of these two mechanisms and their interaction. Zooming out from the details of the proposal, we believe that this series of moves has two sets of additional beneficial consequences.

The first is theoretical: the current proposal provides a way to understand the difficulties that have been associated with head movement for quite some time now. Why is it that some properties of head movement are distinct from those of narrow syntactic phrasal movement? The theoretical split we have proposed provides a resolution to this troubling set of puzzles: many of the purported properties of head movement which are theoretically troubling to a purely syntactic view of the operation are completely compatible with a view in which those properties can instead be attributed to postsyntactic amalgamation.

For example, the observation that most, if not all, cases of head movement have no interpretive effect follows directly from our proposal that head movements associated with word formation result from postsyntactic amalgamation. In the Y-model assumed here, this follows from the assumption that the LF and PF interfaces, while both connected

to syntax, are not directly connected to each other. Other consequences follow as well, including the observation that locality conditions on phrasal and head movement only differ when we compare phrasal movement and the kind of head movement that involves word formation (i.e., postsyntactic amalgamation). Genuine syntactic head movement demonstrably obeys the locality conditions that apply to phrasal movement, including movement across clause boundaries and island sensitivity. Finally, the observation that head movement of the adjunction type violates conditions on structure-building (e.g., the Extension Condition) is problematic only if this type of head movement is thought to be genuinely syntactic; on our proposal, head movement resulting in adjunction is the morphologically-motivated kind—exactly the kind we expect to involve postsyntactic, rather than syntactic, operations. Genuine syntactic head movement seems instead to be completely consistent with the idea that syntactic structures are built by Merge with the root.

A second set of beneficial consequences is that the proposal potentially offers novel insights into long-standing puzzles about the way head movement interacts with other grammatical operations. This includes interactions with ellipsis, especially of the verb-stranding varieties, in which identity conditions on the stranded verb seem to vary in semi-systematic ways. For example, Gribanova 2017b proposes that different types of head movement may discriminate between different patterns with respect to ellipsis identity conditions on stranded verbs crosslinguistically. In addition, it is possible that postsyntactic amalgamation vs. syntactic movement interact with Chain Reduction in distinct ways, yielding different results with respect to copy pronunciation, in e.g., predicate fronting constructions (see Harizanov and Gribanova 2017 and LaCara 2016a,b).

There remain numerous aspects of this proposal that require further investigation, and which we believe will lead to independently exciting developments. We list some of the major ones here. An obvious question concerns the predicted typology of possible interactions between the two operations described here. This typology, which we believe can, in a fruitful way, be worked out computationally, should also lead to a set of configurations predicted to be manifested across languages. It is possible that some of these interactions will be ruled out for independent reasons, narrowing down the resulting predictions significantly. A second area for future investigation concerns the precise formulation of the operations involved in syntactic head movement, about which we have said little in this paper. Integrating syntactic head movement into a framework like Bare Phrase Structure is likely to have nontrivial consequences for the nature of structure-building and structure preservation (Harizanov, 2016, 2017). On the other hand, we have also provided a formulation of postsyntactic amalgamation, which will interact with other postsyntactic mechanisms—linearization, vocabulary insertion, phase-based spellout—in a way that likely has consequences for our understanding of empirical phenomena at PF: allomorphy, exponence, as well as word-internal phonology. All of these will have to be the subject of future investigation. Our primary focus here has been to demonstrate the fundamental utility, both empirical and theoretical, of attributing word-formation effects and effects on word order to distinct modules of the grammar.

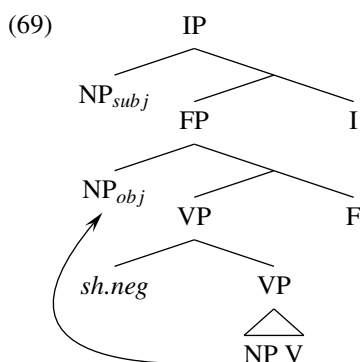
6 Appendix: Only syntactic head movement yields interpretive effects

In §2.2, we demonstrated that for the available case studies on any potential interpretive effects of head movement, the movement in question is the kind which our proposal would take to be syntactic. We know of three putative exceptions to this pattern: head movement in Korean (Han et al., 2007), English NPI licensing by head-moved negation (Ladusaw, 1997, 1980; McCloskey, 1996), and DP interpretation relative to verb position in Spanish (Benedicto, 1998). All three are putative exceptions to our generalization in §2.2, because the head movement in all three cases appears to be the inflection-driven type, and it results in the building of a morphophonological complex. This is the type of head movement that we propose is postsyntactic in nature, and therefore we predict that interpretive effects related to this kind of head movement should not exist. We cover the Korean and English cases in detail below, arguing that there is an alternative explanation for the attested patterns. Space considerations prevent us from implementing a detailed discussion of Benedicto’s proposal here; fortunately, we can refer to Hall 2015, 106–117 for a summary of Benedicto’s arguments and an in-depth counter-argument.

6.1 Korean verb raising and the scope of negation

The first case we consider in detail is Han et al.'s (2007) of Korean verb raising, in which it is demonstrated experimentally that within the population of Korean speakers there are actually two distinct grammars—one in which the verb is raised, and one in which it is not. The group that interests us is the one whose grammar makes use of verb raising; this is because the primary evidence for this raising comes from the scope interactions between negation (attached to the verb) and a quantified DP in its surface position (Korean is a surface scope language). If head movement of the so-called short negation particle (along with the rest of the verbal complex) results in a difference in scope-taking possibilities with respect to a quantified object, then this is an interpretive consequence that is associated with amalgamation-type head movement—a result our theory predicts should not exist.

The argument consists of several components. First, the authors take the stance, along with Hagstrom (2000, 2002), that objects move from their base positions into a higher position, outside of VP. They also assume that short negation (the type of negation that is of interest to us here) is adjoined to VP, so that the surface position of the object is higher than the initial position of the negative particle.⁵⁸



Given this kind of a structure, the reasoning is that if a verb moves up the tree to a position higher than the position that hosts object shift, the negation (which is pied-piped along) will be able to scope above the quantified object.⁵⁹ What Han et al. demonstrate is that a portion of the population they tested does indeed get the reading in which the negation scopes higher than the quantified object—this, for them, leads to the conclusion that the verb has undergone movement, pied-piping negation along, to a high position.⁶⁰

There are, however, at least two entirely different analyses which Han et al. do not consider, but which require no verb movement and still account for the relevant facts. On one of these analyses, the population for which negation can scope over the quantified object speaks a variant of Korean in which the negation itself may adjoin higher in the structure. The fact that children's Korean often involves a stage in which the short negation is pronounced separately from the verb seems to support this analysis: in such cases, the negation appears pre-verbally, suggesting that it is merged higher in the structure than the shifted object.

(70) Sentences produced by 2- and 3-year-old Korean-speaking children (Han et al. 2007:18, (44)):

- a. Na an pap mek-e.
 I NEG rice eat-DECL
 'I will not eat rice.'

(Cho and Hong, 1988)

⁵⁸ It is well known that Korean also has a type of negation called long negation, and the two forms may co-occur. Han et al. 2007 further make the case that short negation is most likely an adjunct, or (less likely) it is the head of a lower Neg projection (where long negation is housed in a higher Neg projection).

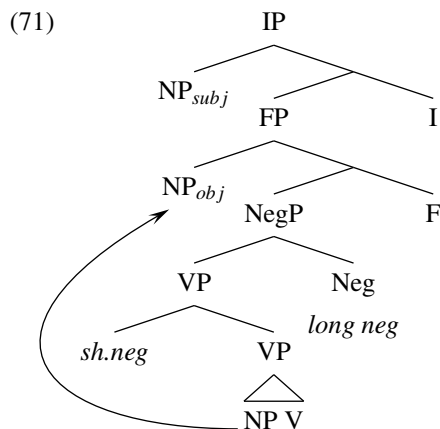
⁵⁹ We put aside, for the moment, the difficulty that short negation is hosted in an adjunct position, and the traditional understanding of head movement will not permit movement through an adjunct position.

⁶⁰ As with English, there is an open question as to how the negation can scope over the object since, with a head-adjunction approach, it will be too deeply embedded in the complex head. See the latter part of this section for discussion if this issue.

- b. An mak uwl-e.
 NEG much cry-DECL
 ‘(I) do not cry much.’ (Cho and Hong, 1988)
- c. An kyelan mek-e.
 NEG egg eat-DECL
 ‘(I) won’t eat eggs.’ (Hahn, 1981)
- d. An kkum kkwe-ese...
 NEG dream dream-because
 ‘Because (I) did not dream...’ (Kim, 1997)

That such a syntactic configuration is possible in a stage of acquisition supports the idea that the scope of negation with respect to the quantified objects may be a consequence of different possibilities with respect to the height of attachment of negation.⁶¹ If this is the case, we have an alternative understanding of why there may be two grammars among Korean speakers: one subset of the population adjoins the negation lower than the surface position of the object; one subset of the population adjoins it higher than the surface position of the object. If this kind of explanation is viable, the result is that the empirical picture sketched by Han et al. (2007) does not necessitate the conclusion that there is scope-expanding head movement that is of the amalgamation type.

A second alternative, suggested by David Adger (via an anonymous reviewer), is that the object in such configurations may reconstruct to the lower position for some speakers of Korean, yielding a different grammar and therefore a different interpretation for some subset of the population. Scope freezing is shown in Han et al. 2007 between subjects and objects; this means that subject reconstruction must be impossible across the board. But object reconstruction is still consistent with the scope freezing between subjects and objects, and therefore could in theory obtain. This approach has the advantage that it predicts a result from a follow-up study in Han et al. 2016, where the authors demonstrate that speakers treat long and short negation similarly: if they get short negation out-scoping the object, they also get long negation out-scoping the object. This follows from a structure like (71), where short and long negation co-occur as Han et al. (2007) would have them.



⁶¹ A reviewer expresses the concern that making available variable positions for negation across languages may invalidate numerous other arguments with respect to verb movement and the relative position of verbs with negation. The language-specific situation in Korean, though, speaks to the necessity of such a move: there are already two ways to expone different types of negation, long and short; Han et al.’s (2007) structure has short negation as an adjunct, making it more likely, at least in principle, that there might be different adjunction sites in the structure. The idea that all languages project negation in the same position in the extended projection is an interesting theoretical stance, but it seems to us that the jury is out about the degree to which it is consistent with existing typological surveys. What is important is that acquirers (in the default case) get quite a bit of convergent evidence about the position of negation in their language. The hypothesis we put forth here about Korean—that speakers may make different conclusions about the relative height of negation—is possible because Korean is a right-headed language with a short negation that is projected on the left, leaving little overt evidence for its position. The data from earlier stages of acquisition seems to confirm this. In essence, we are making the point that Han et al.’s argument about the absence of strong evidence with respect to the position of the verb may apply equally well to the position of short negation. The ambiguity is far less likely to arise in left-headed languages, i.e., French.

The two instances of negation in (71) are quite close to one another, with the result that if the object reconstructs for speakers with one of the possible grammars, it will reconstruct to a position that is beneath both instances of negation.⁶²

6.2 English NPI licensing

The second paradigm we consider here comes from English, and is an oft-cited case in which it appears that head movement of T to C, pied-piping clitic negation, licenses an otherwise unlicensed negative polarity item (NPI). We start with the basic facts, namely that one way of licensing an NPI in English is to position it so that it is c-commanded by something negative (McCloskey, 1996). We see this in the phrasal context in (72), where the NPI *anybody* is licensed only when *nobody* undergoes raising and c-commands it from the raised position.

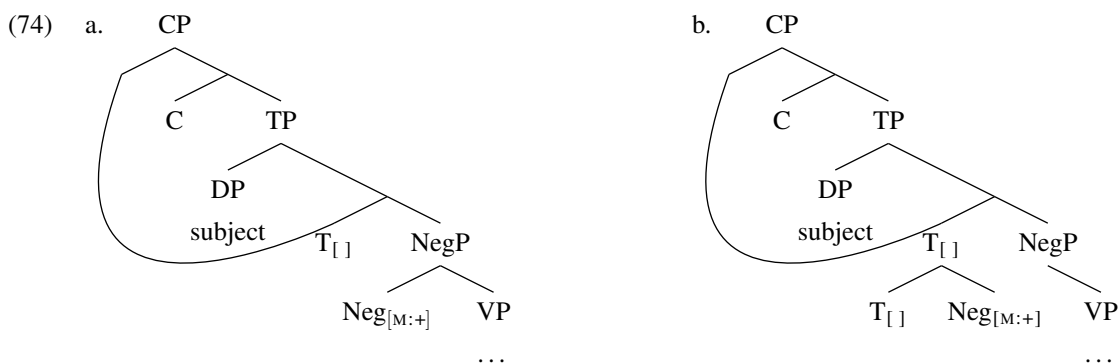
- (72) a. * After the meeting, it seemed to anybody that nobody was satisfied with the outcome.
 b. After the meeting, nobody seemed to anybody to be satisfied with the outcome.

Important for our discussion is that this pattern seems to extend to head movement of negation to a position above the NPI:

- (73) a. * Which of them does anybody like?
 b. * Which one of them does anybody not like?
 c. Which one of them doesn't anybody like?

(73) demonstrates that negation must be present to license a subject NPI in English (73a), but merely its presence in the clause is not enough (73b); the negation must be raised—via head movement on the traditional analysis—to a position that is higher in the clause than the NPI (73c). To understand why this paradigm bears on our overall argument, let us first consider one way in which the theory proposed here might be made to account for it. On our view, the head movement involved will be of two different types: the movement of T to C will be syntactic, as it involves no word formation; but the unification of the auxiliary or modal in T with negation *does* result in word formation, and may therefore, at least in theory, be construed as an instance of postsyntactic amalgamation.

Below, we will argue against such an approach; in order to do this, we first sketch how the account would work. As discussed in §4, the postsyntactic and syntactic types of head movement may interact; in this case, the syntactic component will involve movement of T to C; PF will amalgamate Neg into T, yielding a complex head containing T and Neg. This complex head will be realized in C, in accordance with the English tendency to realize the highest occurrence in a syntactic movement chain.



⁶² There is an additional question as to how the negation comes to be a part of the rest of the complex. Han et al. (2007) indicate that the composition is via head movement, at least in the case of short negation—but this will only work if short negation is the exponent of a head in the clausal spine, not an adjunct. Since the question does not bear directly on the two alternatives outlined here, in that regardless of what the relevant process is, it should not affect scope, we will leave this as an open question for the moment.

If what is crucial for the scope effect described just above is the presence of negation in the higher position that licenses the NPI, and if it is postsyntactic amalgamation of Neg into T that is partially responsible for the presence of Neg in the higher position, this instance of amalgamation would have interpretive effects. If this is the correct interpretation of what is going on in (73), then (73) is a counterexample, involving postsyntactic amalgamation that—counter to our expectations—feeds the LF interface.

We propose here that there is an alternative explanation for the English NPI licensing paradigm *that involves no postsyntactic amalgamation*. We think that this alternative is empirically superior, and adopting it will have the consequence that the only head movement yielding an interpretive effect is the movement of T to C—a head movement we would designate as taking place on the syntax. The alternative we wish to put forward here is that what matters for NPI licensing in (73) is the presence of a C head, valued with negative polarity features and c-commanding the relevant NPI. Head movement of T to C (or of Neg to T) is therefore orthogonal to this, and an indirect reflex of the fact that in Standard American English, the C that comes with unvalued polarity features is also the C that triggers T to C.

There are two ways to implement the idea that a C head with negative features is what licenses the subject NPI. One is to claim, with Biberauer and Roberts 2010, that English has both positive and negative auxiliaries, the latter a result of a T that comes with valued negative polarity features, as has been claimed for the Uralic languages, Latin, Old English, and Afrikaans. Evidence for this comes, among other things, from Zwicky and Pullum's (1983) observation that *n't* triggers stem allomorphy (e.g., **will'nt/won't*). On such an approach, the negative auxiliary in T will be raised to C, possibly in conjunction with an AGREE relation that values unvalued polarity features on C, licensing the NPI directly.⁶³ If this is the case, the only head movement involved is T to C, and to the extent that it is even responsible for the NPI licensing, the head movement is not of the kind that leads to word formation—as our proposal would predict.⁶⁴ A second possibility is the one proposed in Roberts 2010 and summarized in Hall 2015, in which both C and T are encoded with unvalued polarity features, while Neg has them valued. T agrees with Neg, and C subsequently agrees with T. Polarity feature valuation or sharing among the C-Pol-T complex receives quite a bit of support crosslinguistically; for analogous proposals, see McCloskey 2017 for Irish, Landau 2002 for Hebrew, and Gribanova 2017a for Russian. As Hall (2015) points out, it is then the presence of a C valued with negative polarity features that is responsible for the NPI licensing; head movement becomes orthogonal to the discussion.

Either approach described above seems empirically superior to an account that relies on head movement of Neg, insofar as it can unify the licensing of subject NPIs in a broader range of environments than can an account in which head movement of Neg is the key. If the unifying observation is that a negatively-valued C is what matters, then we have a way of understanding not just (73), but also two further observations. The first comes from Hiberno-English (Harris, 1984; Duffield, 1993), in which lower negation also seems to license subject NPIs in declaratives.

- (75) a. Although anybody don't seem to like to live in Russia....They're all trying to get out of it.
 b. You couldn't pick a daisy but it was a sin. Now, anything is no sin....
 (from Duffield 1993, citing Filppula's (1986,1991) corpus of Hiberno-English)

If we take the NPI licensing in Hiberno-English to be the result of valuation of C's unvalued polarity features by a lower negation, the pieces fall into place. The difference between Standard American English and Hiberno-English on this account is whether this kind of C triggers T to C movement. The valuation of C's polarity features remains the common link between them, and it is the availability of this negative C that licensed the subject NPI in both varieties, not the head movement.

The second empirical argument, originating in Ladusaw 1997, is the observation that NPIs can be licensed in the complement of certain inherently negative predicates, such as *unlikely*.

- (76) It is unlikely that anyone will be appointed.

⁶³ If T to C is an instance of syntactic head movement as reprojection, as we suggest in §4, then the T with negative features will also c-command the NPI directly from its raised position. This is an improvement over the standard head movement qua head adjunction account, since in that account—as pointed out by Hall (2015)—the head-moved negation will be too deeply embedded to c-command the NPI it is supposed to license.

⁶⁴ A reviewer points out that another argument for the semantic potency of T-to-C movement with contracted negation comes from Romero and Han 2004 and Romero 2005, who argue that the raising of negation to a higher position is what affects the interpretation of polar questions in English, yielding Ladd's ambiguity (Ladd, 1981).

(77) That anyone will be appointed is unlikely.

Laka (1990) accounts for this observation by proposing that certain predicates select for inherently negative complementizers, which in turn license NPIs inside their complement, even if it is moved.

To the extent that a negatively valued C is the common thread among these disparate environments in which NPIs are licensed, it seems that attributing the licensing of subject NPIs to head movement in (73) is too narrow an approach. This is broadly consistent with the vast literature on the licensing of NPIs in English and crosslinguistically, going back to Ladusaw 1997, 1980; McCloskey 1996; Linebarger 1980; Han and Siegel 1997, among others, which makes quite clear that the scope of negation is only one context where licensing may occur. Ladusaw's (1979) and Fauconnier's (1975, 1979) major contribution was the observation that NPIs are licensed in downward entailing environments, but a focal point of more recent discussion has been the observation that NPIs are also licensed in interrogative contexts that cannot, on the canonical view, be considered downward entailing (Han and Siegel, 1997; Progovac, 1993; Nicolae, 2013, 2015).

(78) NPIs licensed in polar questions:

- a. Has anyone attended the meeting?
- b. * Anyone has attended the meeting?
(ungrammatical as a rising declarative)
- c. Have the committee members attended any meetings recently?
- d. * The committee members have attended any meetings recently?
(ungrammatical as a rising declarative)

(79) NPIs licensed in (certain) WH-questions:

- a. Who attends any of these meetings?
- b. * Which meetings does anyone attend?

What these examples make clear is that the paradigm in (73) is exceptional, in the sense that NPIs are generally licensed in both polar and WH-interrogatives. The subject-object asymmetry observed in (79) is apparently linked to the observation (Han and Siegel, 1997) that the NPI must be c-commanded by the WH-phrase when both are in their base positions; see Nicolae 2013, 2015 for a series of proposals that propose a semantics for questions that relates them back to the idea of a downward entailing context being the key licensing factor, yielding the contrast in (79) as a direct consequence. Whatever the correct unifying analysis, our point here simply that there is no necessity to appeal to the head movement of negation to a high position as the direct cause of the NPI licensing. It is equally possible that the appearance of negation in a high position and the NPI licensing are both reflexes of a deeper property (e.g., compatibility with some kind of unifying semantic operator).

We conclude that the paradigm under discussion receives a better account without any appeal to head movement. This also has the consequence that this paradigm is no longer a threat to the predictions made by the overall proposal under discussion here.

Acknowledgements

We are grateful to the following colleagues for helpful comments and feedback on this work: Klaus Abels, David Adger, Karlos Arregi, Rajesh Bhatt, Lauren Eby Clemens, Sandra Chung, Güliz Güneş, Heidi Harley, Peter Jenks, Paul Kiparsky, Nicholas LaCara, Beth Levin, Paul Kiparsky, Anikó Lipták, David Pesetsky, Omer Preminger, Ian Roberts, Peter Svenonius, Lisa Travis, and audiences at FDSL 12, Leiden University, the McGill Word Structure research group, and the 91st annual LSA meeting in Austin, TX. Conversations with Jim McCloskey over the course of about two years have been crucial to this paper's development. We thank him especially for all of these exchanges, and for letting us steal the paper's title from an old handout of his. Finally, we are grateful to Daniel Harbour and three *NLLT* reviewers for detailed comments and constructive criticism. All errors are the authors' responsibility alone.

References

- Abels, Klaus. 2003a. Auxiliary adverb word order revisited. *University of Pennsylvania Working Papers in Linguistics* 9 (1).
- Abels, Klaus. 2003b. Successive cyclicity, anti-locality, and adposition stranding. PhD diss, University of Connecticut.
- Adger, David. 2013. *A syntax of substance*. Cambridge, MA: MIT Press.
- Adger, David, Daniel Harbour, and Laurel Watkins. 2009. *Mirrors and microparameters: Phrase structure beyond free word order*. Cambridge: Cambridge University Press.
- Ahn, Byron. 2015. Giving reflexivity a voice: Twin reflexives in English. PhD diss, University of California, Los Angeles.
- Aissen, Judith. 1974. Verb raising. *Linguistic Inquiry* 5 (3): 325–366.
- Bailyn, John Frederick. 1995a. A configurational approach to Russian ‘free’ word order. PhD diss, Cornell University.
- Bailyn, John Frederick. 1995b. Underlying phrase structure and ‘short’ verb movement in Russian. *Journal of Slavic Linguistics* 3 (1): 13–58.
- Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.
- Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Benedicto, Elena. 1998. Verb movement and its effects on determinerless plural subjects. In *Romance linguistics: Theoretical perspectives*, eds. A. Schwegler, B. Tranel, and M. Uribe-Etxebarria, 25–40. Amsterdam/Philadelphia: John Benjamins.
- Bennett, Ryan, Emily Elfner, and James McCloskey. 2015. Prosody, focus and ellipsis in Irish. Ms..
- Biberauer, Theresa, and Ian Roberts. 2010. Subjects, tense and verb-movement in Germanic and Romance. In *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press.
- Bobaljik, Jonathan. 2002. A-chains at the PF-interface: copies and ‘covert’ movement. *Natural Language and Linguistic Theory* 20 (2): 197–267.
- Borsley, Robert D., and Andreas Kathol. 2000. Breton as a V2 language. *Linguistics* 38 (4): 665–710.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24 (4): 591–656.
- Brody, Michael. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31 (1): 29–56.
- Broekhuis, Hans, and Krzysztof Migdalski. 2003. Participle fronting in Bulgarian. *Linguistics in the Netherlands* 20: 1–12.
- Cable, Seth. 2012. ?pied-piping: Introducing two recent approaches. *Language and Linguistics Compass* 6 (12): 816–832.
- Cho, Young-Mee Yu, and Ki-Sun Hong. 1988. Evidence for the VP constituent from child Korean. In *Papers and reports on child language development* 27, 31–38. Stanford University Department of Linguistics.
- Chomsky, Noam. 1995a. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Cambridge: Blackwell.
- Chomsky, Noam. 1995b. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, eds. R. Martin, D. Michaels, and J. Uriagereka. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory essays in honor of Jean-Roger Vergnaud*, eds. Robert Freidin, C. P. Otero, and Maria Luisa Zubizarreta. MIT Press.
- den Besten, Hans. 1983. On the interaction of root transformations and lexical deletion rules. In *On the formal syntax of the Westgermania*, ed. Werner Abraham, 47–131. Amsterdam: John Benjamins.
- Donati, Caterina. 2006. On wh-head movement. In *Wh-movement: Moving on*, eds. Lisa Lai-Shen Cheng and Norbert Corver, 21–46. Cambridge, MA: MIT Press.
- Donati, Caterina, and Carlo Cecchetto. 2011. Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry* 42 (4): 519–560.
- Duffield, Nigel. 1993. On case-checking and NPI licensing in Hiberno-English. *Rivista di Linguistica* 7: 215–244.
- Embick, David, and Roumyana Izvorski. 1997a. Participle-auxiliary word orders in Slavic. In *Annual Workshop on*

- Formal Approaches to Slavic Linguistics: the Cornell Meeting 1995*, eds. Natasha Kondrashova, Wayles Browne, Ewa Dornisch, and Draga Zec, Vol. 4, 210–239. Michigan Slavic Publications.
- Embick, David, and Roumyana Izvorski. 1997b. Participle-auxiliary word orders in Slavic. In *Annual Workshop on Formal Approaches to Slavic Linguistics: the Cornell Meeting 1995*, eds. Natasha Kondrashova, Wayles Browne, Ewa Dornisch, and Draga Zec, Vol. 4, 210–239. Michigan Slavic Publications.
- Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32: 555–595.
- Fanselow, Gisbert. 2003. Münchenhausen-style head movement and the analysis of verb second. *UCLA Working Papers in Linguistics* 13: 40–76.
- Fanselow, Gisbert. 2009. Bootstrapping verb movement and the causal architecture of German (and other languages). In *Advances in comparative Germanic syntax*, eds. Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, and Justin Nuger, 85–118. John Benjamins.
- Fauconnier, Gilles. 1975. Pragmatic scales and logical structure. *Linguistic Inquiry* 6: 353–376.
- Fauconnier, Gilles. 1979. Implication reversal in a natural language. In *Formal semantics and pragmatics for natural languages*, eds. F. Guenther and S. Schmidt, 289–301. Dordrecht: Reidel.
- Filppula, Markku. 1986. Some aspects of Hiberno-English in a Functional Sentence Perspective. *University of Joensuu Publications in the Humanities* 7.
- Filppula, Markku. 1991. From Anglo-Irish to Hiberno-English: divergence and convergence in the Irish dialects of English. In *Proceedings of the International Congress of Dialectologists, Bamberg*, ed. W. Vierich. Wiesbaden: Frank Steiner Verlag.
- Franks, Steven. 2008. Clitic placement, prosody, and the Bulgarian verbal complex. *Journal of Slavic Linguistics* 16 (1): 91–137.
- Fukui, Naoki, and Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 7 (1): 27–86.
- Georgi, Doreen, and Gereon Müller. 2010. Noun-phrase structure by Re-Projection. *Syntax* 13 (1): 1–36.
- Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language and Linguistic Theory* 31 (1): 91–136.
- Gribanova, Vera. 2017a. Head movement and ellipsis in the expression of Russian polarity focus. *Natural Language and Linguistic Theory* 35 (4): 1079–1121.
- Gribanova, Vera. 2017b. Head movement, ellipsis, and identity. Stanford University.
- Gribanova, Vera, and Line Mikkelsen. 2017. On the interaction of head movement and ellipsis in Danish. Ms..
- Grimshaw, Jane. 1991. Extended projection. Ms. Brandeis University.
- Grohmann, Kleantes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. John Benjamins Publishing Company.
- Hagstrom, Paul. 2000. Phrasal movement in Korean negation. In *Proceedings of the 9th student conference in linguistics (SCIL 9)*, eds. Ljuba Veselinova, Susan Robinson, and Lamont Antieau, 127–142. Cambridge, MA: MIT Working Papers in Linguistics.
- Hagstrom, Paul. 2002. Implications of child error for the syntax of negation in Korean. *Journal of East Asian Linguistics* 11: 211–242.
- Hahn, Kyung-Ja Park. 1981. The development of negation in one Korean child. PhD diss, University of Hawaii, Honolulu.
- Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.
- Hale, Ken, and Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.
- Hall, David. 2015. Spelling out the noun phrase: Interpretation, word order, and the problem of ‘meaningless movement’. PhD diss, Queen Mary University of London.
- Han, Chung-hye, and Laural Siegel. 1997. Syntactic and semantic conditions on NPI licensing in questions. In *Proceedings of the 15th West Coast Conference on Formal Linguistics*, eds. B. Agbayani and S. Tang. Stanford, CA: CSLI Publications.
- Han, Chung-hye, Jeffrey Lidz, and Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38 (1): 1–47.
- Han, Chung-hye, Jeffrey Lidz, and Julien Musolino. 2016. Endogenous sources of variation in language acquisition.

- PNAS 113 (4): 942–947.
- Harbour, Daniel. 1999. The two types of predicate clefts: Classical Hebrew and beyond. *MIT Working Papers in Linguistics*.
- Harbour, Daniel. 2007. Beyond PersonP. *Syntax* 3 (10): 223–242.
- Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* 4 (32): 1033–1088.
- Harizanov, Boris. 2016. Head movement to specifier positions in Bulgarian participle fronting. Presented at 90th Annual Meeting of the Linguistic Society of America (LSA 90). https://stanford.edu/~bharizan/pdfs/Harizanov_2016.LSA_handout.pdf.
- Harizanov, Boris. 2017. On the nature of syntactic head movement. Colloquium, Department of Linguistics, McGill University, Montréal, Québec.
- Harizanov, Boris. 2018. Word formation at the syntax-morphology interface: Denominal adjectives in Bulgarian. *Linguistic Inquiry* 49 (2): 283–333.
- Harizanov, Boris, and Vera Gribanova. 2017. Post-syntactic head movement in Russian predicate fronting. Presented at the 91st annual meeting of the Linguistic Society of America.
- Harley, Heidi. 2013a. Diagnosing head movement. In *Diagnosing syntax*, eds. Lisa Lai-Shen Cheng and Norbert Corver, 112–119. Oxford: Oxford University Press.
- Harley, Heidi. 2013b. Getting morphemes in order: Merger, affixation and head movement. In *Diagnosing syntax*, eds. Lisa Lai-Shen Cheng and Norbert Corver. Oxford University Press.
- Harris, John. 1984. Syntactic variation and dialect divergence. *Journal of Linguistics* 20: 307–327.
- Hartman, Jeremy. 2011. The semantic uniformity of traces: Evidence from ellipsis parallelism. *Linguistic Inquiry* 42 (3): 367–388.
- Harves, Stephanie. 2002. Genitive of negation and the syntax of scope. In *Proceedings of ConSOLE IX*, eds. M. van Koppen, E. Thrift, E. J. van der Torre, and M. Zimmerman, 96–110.
- Heck, Fabian. 2008. *A theory of pied-piping*. Mouton de Gruyter.
- Heck, Fabian. 2009. On certain properties of Pied-Piping. *Linguistic Inquiry* 40 (1): 75–111.
- Heim, Irene. 1997. Predicates or formulas? Evidence from ellipsis. In *Proceedings from Semantics and Linguistic Theory VII*, ed. Aaron Lawson, 197–221. Cornell, Ithaca: CLC Publications.
- Holmberg, Anders. 1986. Word order and syntactic features in the Scandinavian languages and English. PhD diss, University of Stockholm.
- Iatridou, Sabine, and Ivy Sichel. 2011. Negative DPs, A-movement, and scope diminishment. *Linguistic Inquiry* 42 (2): 595–629.
- Iatridou, Sabine, Elena Anagnostopoulou, and Roumyana Izvorski. 1999. Some observations about the form and meaning of the perfect. In *Festschrift for Ken Hale*, ed. Michael Kenstowicz. Cambridge, MA: MIT Press.
- Izvorski, R. 2000. Free relatives and related matters. PhD diss, University of Pennsylvania.
- Jenks, Peter. 2013. Head movement in Moro DPs: Evidence for a unified theory of movement. In *Proceedings of the 31st West Coast C on Formal Linguistics*, ed. Robert E. Santana-LaBarge, 248–257. Somerville, MA: Cascadilla Proceedings Project.
- Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.
- Kayne, Richard S. 1991. Romance clitics, Verb movement, and PRO. *Linguistic Inquiry* 22 (4): 647–686.
- Keine, Stefan, and Rajesh Bhatt. 2016. Interpreting verb clusters. *Natural Language & Linguistic Theory* 34 (4): 1445–1492.
- Kim, Young-Joo. 1997. The acquisition of Korean. In *The crosslinguistic study of language acquisition, vol. 4*, ed. Dan I. Slobin, 334–443. Hillsdale, N.J.: Lawrence Erlbaum.
- King, Tracy Holloway. 1995. *Configuring topic and focus in Russian*. Stanford: CSLI Publications.
- Koenenman, Olaf. 2000. *The flexible nature of verb movement*. Utrecht: LOT.
- Koopman, Hilda. 1984. *The syntax of verbs: From verb-movement rules in the Kru languages to universal grammar*. Dordrecht: Foris.
- Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
- LaCara, Nicholas. 2016a. A Germanic verb movement paradox. Presented at the Stanford Workshop on Head Movement. Stanford University.

- LaCara, Nicholas. 2016b. Verb phrase movement as a window into head movement. *Proceedings of the Linguistic Society of America* 1(17): 1–14.
- Ladd, Robert D. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the seventeenth regional meeting of the Chicago Linguistic Society*, 164–171. Chicago Linguistic Society.
- Ladusaw, William. 1980. On the notion ‘affective’ in the analysis of negative-polarity items. *Journal of Linguistic Research* 1: 1–16.
- Ladusaw, William. 1997. Polarity sensitivity as inherent scope relations. PhD diss, University of Texas at Austin. Published in the Distinguished Dissertations in Linguistics Series, Garland Press.
- Laka, Itziar. 1990. Negation in syntax: On the nature of functional projections in syntax. PhD diss, MIT.
- Lambova, Mariana. 2004. On triggers of movement and effects at the interfaces. In *Studies in generative grammar*, 75: *Triggers*, eds. Anne Breibarth and Henk C. van Reimsdijk, 231–258. Walter de Gruyter.
- Landau, Idan. 2002. (Un)interpretable Neg in Comp. *Linguistic Inquiry* 33: 465–492.
- Landau, Idan. 2006. Chain resolution in Hebrew (V)P-fronting. *Syntax* 9 (1): 32–66.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19 (3): 335–391.
- Lasnik, Howard. 2001. When can you save a structure by destroying it? In *Proceedings of the North East Linguistic Society 31*, eds. Min-Joo Kim and Uri Strauss, 301–320. Amherst, MA: GLSA.
- Lechner, Winifred. 2007. Interpretive effects of head movement. <http://ling.auf.net/lingBuzz/000178>.
- Lema, José, and María Luisa Rivero. 1990. Long head movement: HMC vs. ECP. In *Proceedings of the 20th annual meeting of the North East Linguistic Society*, ed. Juli A. Carter, 337–347. Amherst, MA: GLSA.
- Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
- Lightfoot, David. 1993. Why UG needs a learning theory: Triggering verb movement. In *Historical linguistics: Problems and perspectives*, ed. Charles Jones, 190–214. London: Longmans.
- Linebarger, Marcia C. 1980. The grammar of negative polarity. PhD diss, MIT.
- Lipták, Anikó, and Andrés Saab. 2014. No N-raising out of NPs in Spanish: ellipsis as a diagnostic of head movement. *Natural Language and Linguistic Theory* 32: 1247–1271.
- Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 1: 69–109.
- McCloskey, James. 1996. On the scope of verb raising in Irish. *Natural Language and Linguistic Theory* 14: 47–104.
- McCloskey, James. 2012. Polarity, ellipsis and the limits of identity in Irish. Workshop on Ellipsis, Nanzan University. <http://ohlone.ucsc.edu/jim/PDF/nanzan-handout.pdf>.
- McCloskey, James. 2016. Interpretation and the typology of head-movement: A re-assessment. Talk presented at the Workshop on the Status of Head Movement in Linguistic Theory, Stanford University.
- McCloskey, James. 2017. Ellipsis, polarity and the cartography of verb-initial orders in Irish. In *Elements of comparative syntax: Theory and description*, eds. Enoch Aboh, Eric Haeberli, Genoveva Puskás, and Manuela Schönenberger, 99–151. De Gruyter Mouton.
- McPherson, Laura. 2014. Replacive grammatical tone in the Dogon languages. PhD diss, University of California, Los Angeles.
- Merchant, Jason. 2008. Variable island repair under ellipsis. In *Topics in ellipsis*, ed. Kyle Johnson, 132–153. Cambridge: Cambridge University Press.
- Messick, Troy, and Gary Thoms. 2016. Ellipsis, economy and the (non)uniformity of traces. *Linguistic Inquiry* 47 (2): 306–332.
- Mikkelsen, Line. 2015. VP anaphora and verb-second order in Danish. *Journal of Linguistics* 51 (3): 595–643.
- Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. Philadelphia, PA: John Benjamins Publishing Company.
- Nicolae, Andreea C. 2013. Any questions? Polarity as a window into the structure of questions. PhD diss, Harvard University.
- Nicolae, Andreea C. 2015. Questions with NPIs. *Natural Language Semantics* 23: 21–76.
- Ostrove, Jason. 2015. Allomorphy and the Irish verbal complex. UC Santa Cruz. Available at: <http://people.ucsc.edu/~7Ejostrove/QP1.pdf>.
- Ostrove, Jason. 2016. Mixtec syntax and morphology. UC Santa Cruz.
- Penka, Doris. 2002. Zur Semantik der negativen Indefinita im Deutschen Universität Tübingen: Tübingen-Linguistik-Report Nr. 1.

- Penka, Doris, and Arnim von Stechow. 2001. Negative indefinita unter modalverben. In *Modalität und modalverben im Deutschen*, eds. Reimar Müller and Marga Reis, 263–286. Hamburg: Helmut Buske Verlag.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 355–426. MIT Press.
- Pietraszko, Asia, and Karlos Arregi. 2018. Generalized head movement. *Proceedings of the Linguistic Society of America*.
- Platzack, Christer. 2013. Head movement as a phonological operation. In *Diagnosing syntax*, eds. Lisa Lai-Shen Cheng and Norbert Corver, 21–43. Oxford: Oxford University Press.
- Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Preminger, Omer. to appear. What the PCC tells us about “abstract” agreement, head movement, and locality. *Glossa*.
- Progovac, Ljiljana. 1993. Negative polarity: Entailment and binding. *Linguistics and Philosophy* 16: 149–180.
- Richards, Norvin. 1998. The principle of minimal compliance. *Linguistic Inquiry* 29 (4): 599–629.
- Rivero, Maria-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review* 8: 319–351.
- Rivero, María Luisa. 1992. Patterns of V⁰-raising in Long Head Movement and Negation: Serbo-Croatian vs. Slovak. In *Syntactic theory and Basque syntax*, eds. Joseba Andoni Lakarra and Jon Ortiz de Urbina, 365–386. Gipuzkoako Foru Aldundia.
- Rivero, Maria-Luisa. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory* 12: 62–120.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi, and Ian Roberts. 1989. Complex inversion in French. *Probus* 1 (1): 1–30.
- Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22 (1): 209–218.
- Roberts, Ian. 1994. Two types of head movement in Romance. In *Verb movement*, eds. David Lightfoot and Norbert Hornstein, 207–242. Cambridge University Press.
- Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.
- Rohrbacher, Bernhard. 1994. The Germanic VO languages and the full paradigm: A theory of V to I raising. PhD diss, University of Massachusetts-Amherst.
- Romero, Maribel. 2005. Two approaches to biased yes/no questions. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 352–360. Somerville, MA: Cascadilla Proceedings Project.
- Romero, Maribel, and Chung-Hye Han. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27 (27): 609–658.
- Schafer, Robin. 1995. Negation and verb second in Breton. *Natural Language and Linguistic Theory* 13 (1): 135–172.
- Schoorlemmer, Erik, and Tanja Temmerman. 2012. Head movement as a PF-phenomenon: evidence from identity under ellipsis. In *Proceedings of the 29th West Coast Conference on Formal Linguistics*, eds. Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, and Alex Trueman, 232–240. Somerville, MA: Cascadilla Press.
- Selkirk, Elisabeth. 1995. Sentence prosody: intonation, stress and phrasing. In *The handbook of phonological theory*, ed. John A. Goldsmith. London: Blackwell Publishing.
- Surányi, Balázs. 2005. Head movement and reprojection. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. ELTE Tomus* 26: 313–342.
- Svenonius, Peter. 2004. Slavic prefixes and morphology: An introduction to the Nordlyd volume. *Nordlyd* 32 (2): 177–204.
- Svenonius, Peter. 2016. Spans and words. In *Morphological metatheory*, eds. Daniel Siddiqi and Heidi Harley, 201–222. John Benjamins Publishing Company.
- Svenonius, Peter. 2018. Delimiting the syntactic word. Presented at *Linguistics at Santa Cruz*.
- Szabolcsi, Anna. 2011. Certain verbs are syntactically explicit quantifiers. In *The Baltic international yearbook of cognition, logic and communication. volume 6, formal semantics and pragmatics: Discourse, context, and models*. <http://cognition.lu.lv/symp/6-call.html>.
- Takahashi, Shoichi, and Danny Fox. 2005. MaxElide and the re-binding problem. In *Proceedings of Semantics and Linguistic Theory XV*, eds. Effi Georgala and Jonathan Howell, 223–240. Ithaca, New York: CLC Publications.

- Toyoshima, Takashi. 2001. Head-to-spec movement. In *The minimalist parameter*, eds. Galina M. Alexandrova and Olga Arnaudova, 115–136. Amsterdam and Philadelphia: John Benjamins.
- Travis, Lisa. 1984. Parameters and effects of word order variation. PhD diss, MIT.
- van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A Dinka Bor case study. PhD diss, Massachusetts Institute of Technology.
- van Urk, Coppe. to appear. Pronoun copying in Dinka and the Copy Theory of Movement. *Natural Language and Linguistic Theory*.
- Vicente, Luis. 2007. The syntax of heads and phrases: A study of verb (phrase) fronting. PhD diss, Leiden University.
- Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.
- von Stechow, Arnim. 1993. Die aufgaben der syntax. In *Ein internationales handbuch zeitgenössischer forschung*, eds. Arnim von Stechow and Theo Vennemann, 1–88. Berlin: Mouton de Gruyter.
- Wechsler, Stephen. 1991. Verb second and illocutionary force. In *Views on phrase structure*, eds. Katherine Leffel and Denis Bouchard, 177–191. Netherlands: Springer.
- Wiklund, Anna-Lena. 2010. In search of the force of dependent verb second. *Nordic Journal of Linguistics* 33 (1): 81–91.
- Wilder, Chris, and Damir Cavar. 1994. Word order variation, verb movement, and economy principles. *Studia Linguistica* 48: 46–86.
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Mouton de Gruyter.
- Zwicky, Arnold, and Geoff Pullum. 1983. Cliticisation vs. inflection: English *n't*. *Language* 59: 502–513.