Verb Phrase Movement as a Window into Head Movement

Nicholas LaCara · UMass Amherst   June 1, 2016

Abstract

Since Chomsky (2001) suggested that head movement might be a PF operation, there has been debate about the proper place of head movement in the grammar. The interaction of verb movement with verb phrase fronting can shed light on how and when head movement occurs. This paper looks at cases where verb phrase fronting generates two copies of the verb (as in Portuguese or Hebrew), one in the fronted vP and one in an inflectional position, showing how a PF approach to head movement can explain this pattern while addressing some potential problems in other languages.

Keywords

Contents

1   Introduction

Languages with V°-to-T° movement and verb phrase topicalization (henceforth VPT) exhibit VERB-DOUBLING VPT, or VVPT, where a copy of the verb (root) is pronounced in both the topicalized vP and in an inflectional position. Multiple languages exhibit VVPT, including Hebrew, Portuguese, Russian, and Spanish.[1] In the following examples, the topicalized vP is bracketed, while copies of the verb root are underlined:

(1)   Hebrew (Landau 2006)

[li<u>k</u>not   et   ha-praxim]   hi   ka<u>n</u>ta.
buy.INF   ACC   the-flowers   she   buy.PST

'As for buying the flowers, she bought (them).'

(2)   Portuguese (Bastos 2001)

[<u>lav</u>-ar   o   carro]   o   João   <u>lav</u>-ou.
wash-INF   the   car   the   João   wash-PST.3SG

'As for washing the car, João washed (it).'

(3)   Russian (Abels 2001)

[<u>Duma</u>-t'   o   ženit'be]   on   <u>duma</u>-et…
think-INF   about   marriage   he   think-PRES…

'He does think about marriage…'

(4)   Spanish (Vicente 2009)

[<u>Le</u>-er   el   libro   rápido]   Juan   lo=ha   <u>le</u>-ído.
read.INF   the   book   quickly,   Juan   it=has   read-PRF

'As for reading the book fast, Juan has read it fast.'

Although the details of each implementation vary by language, the general approach to VVPT is to treat it as a species of remnant movement under the copy theory of movement (Chomsky 1995). One copy of the verb root is generated by moving vP to the left periphery. At least one other copy of the verb is pronounced in T° (or some other inflectional position) for morphological reasons, and this copy is thought to move to this position via HEAD MOVEMENT. The syntax underlying (2) can thus be schematized as in (5)
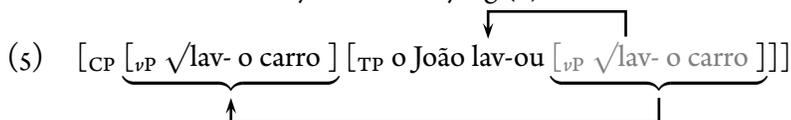
(5) $[_{CP}$ $[_{vP}$ $\sqrt{}$lav- o carro $]$ $[_{TP}$ o João lav-ou $[_{vP}$ $\sqrt{}$lav- o carro $]]]$

While this view is generally accepted (see the references cited in the examples), what remains unclear is *why* two copies of the verb get to be pronounced. The central empirical problem is that phrasal material moved out of the topicalized *v*P, unlike verbal material, cannot be pronounced in the fronted *v*P (Nunes 2004, Gärtner 1998):

(6)   a.   [$_{vP}$ Elected t_i], John$_i$ was.

      b.   *[$_{vP}$ Elected John], John was.

Even in cases where both phrasal and verbal material are moved out of the *v*P, only the verbal material may be pronounced twice. As shown in (7), a copy of the verb (*entregada*) is pronounced in the fronted *v*P, but a copy of the subject (*la medalla*) cannot be (Vicente 2007):

(7)   Spanish (based on Vicente 2009:171, (20)):

[Entreg-ad-a	(*la medalla)	al	ganador]$_k$,	la	medalla$_i$	ha	sido
awarded-PASS-FEM	the medal	to.the	winner,	the	medal	has	been

entreg-ad-a	t_k.
awarded-PASS-FEM	

'Awarded to the winner, the medal has been.'

This differential behavior poses a tricky problem for the copy theory of movement: Why should copies of phrasal material behave differently from copies of verbal material? The general intuition in the literature is that this has something to do with the kinds of movement involved: Whereas *v*P topicalization is phrasal movement, movement of the verb is head movement, and these are already known to have different properties (see Roberts 2011 for a summary).

One answer to the question comes from Nunes (2004). Nunes proposes that multiple copies of the same element cannot be pronounced because it would result in a linearization problem. In each of (6) and (7), copies of the subject cannot be linearized relative to each other, and so the copy in the topicalized *v*P must be eliminated. Copies in head chains are special, though, in that they can be morphologically reanalyzed, rendering an element inside a morphologically complex element invisible to the linearization process.

The problem with this is that there is no principled way to determine when morphological reanalysis will occur. Nunes simply invokes it as necessary. Other authors have provided alternative accounts, including Bastos (2001) and Landau (2006), but as I will discuss, these remain problematic.

Instead, I propose that if head movement is not modeled under the copy theory of movement as Nunes assumes, we end up with a straightforward way of explaining VVPT while retaining the central insights of his approach to copy reduction. The basic idea is as follows: Under a copy-theoretic view of head movement, we are forced to ignore some copies for the purpose of linearization. This is implemented with Nunes's MORPHOLOG-ICAL REANALYSIS, but there is no principled way of determining when this will apply. If, on the other hand, head movement is not derived by copying but instead involves some other mechanism, we reduce the number of copies of the verb generated to begin with,

thereby obviating the need to stipulate which copies are ignored for linearization. Under the approach discussed here – Hale & Keyser's (2002) and Harley's (2004) CONFLATION approach for concreteness – there is no need to stipulate anything like Morphological Re-analysis of certain copies – adopting a non-movement approach to head movement derives the correct results automatically.

1.1 Roadmap

I begin in section 2 by reviewing Nunes's (2004) CHAIN REDUCTION approach to linearizing multiple syntactic copies. In Section 3, I introduce CONFLATION, which I use to model head movement following Harley (2004, 2013). There, I show that this approach to head movement makes the correct predictions for VVPT without additional stipulations. Following this, in Section 4 I argue that Nunes' original approach relies on too many stipulations to account for the data, and other related approaches remain problematic. Section 5 notes a few remaining issues and concludes the paper.

2 Chain reduction and remnant movement

This section introduces Nunes's (2004) theory of chain reduction. It then describes the general approach to remnant movement, of which verb-doubling verb phrase topicalization is typically thought to be a special case. This approach to multiple syntactic copies underlies the forthcoming discussion regarding theoretical implementations of head movement.

2.1 Chain reduction

As noted above, it is generally assumed that the verb-doubling phenomenon is predicted by the copy theory of movement: Moving a vP generates two copies of the verb, and under the right circumstances it becomes possible to pronounce more than one of these copies. One of the more worked-out approaches to deciding which copies in a movement chain to pronounce comes from Nunes (2004). The basic idea behind Nunes's approach is that the need to pronounce a single copy of an element is the result of constraints on linearization (following Kayne's (1994) LINEAR CORRESPONDENCE ALGORITHM, or LCA): Individual links in a movement chain usually count as non-distinct, and non-distinct elements cannot be linearized relative to one another. Consequently, it is not usually possible to linearize such copies and superfluous links must be deleted. This is implemented by the operation CHAIN REDUCTION:

(8) CHAIN REDUCTION (Nunes 2004:27, (44)):
    Delete the minimal number of constituents of a nontrivial chain CH that suffices for CH to be mapped into a linear order in accordance with the LCA.

As the name implies, chain reduction relies critically on the copies of an element being part of the same chain. If movement is really copying, then there must be a reason that traces of movement (typically) remain unpronounced. Nunes proposes that this is because two or more copies of the same element cannot be linearized with respect to one another:

For the purposes of linearization, copies are non-distinct, and following the principle IR-REFLEXIVITY, an item may neither precede nor follow itself (if a precedes β, then $a \neq \beta$). Thus, in order to satisfy irreflexivity, certain copies must not be pronounced at PF. Nunes argues that the highest copy is typically preserved because it is assumed that that copy will have (more) uninterpretable features checked.

Chain reduction can be demonstrated straightforwardly with a passive subject , as in (9).[2] As is commonly assumed, the subject *John* originates as the complement of the verb *elect*. Movement of the subject from its base position to subject position generates a second copy. The two copies of *John$_i$* are treated as non-distinct. It is therefore not possible to pronounce both because *John$_i$* would have to both precede and follow itself.

(9) $[\, \text{John}_i^2 \, [\, \text{was} \, [\, \text{elected} \, \cancel{\text{John}_i^1} \,]\,]\,]$

Following chain reduction, *John^1* is deleted in order to avoid the violation of irreflexivity. *John^2* is preserved on the assumption that it checks more features than *John^1* (*e.g.* Case or the EPP). As a result, only one copy of the subject is pronounced.

2.2   Remnant movement and multiple chains

Remnant movement occurs when some element a is moved out of a phrase β, and β subsequently moves (see Müller 1998). This means that a trace copy of a is contained in β:

(10) $[_\beta \, ... \, t_a \, ... \,] \, ... \, [\, ... \, a \, ... \, [\, ... \, t_\beta \, ... \,]\,]$

This introduces a complication for Chain Reduction, since copies of the same element exist in multiple chains. Nunes argues, nonetheless, that Chain Reduction is capable of capturing remnant movement as long as the correct assumptions about copies are made, and with the further assumption that copy reduction applies 'blindly'. Following criticism by Gärtner (1998), Nunes (2004:52–54) notes that chain reduction does not appear to explain why *John* goes unpronounced in the vP-topicalization example in (11a). On its own, it predicts (11b).

(11)   a.   Elected, John was.

b.   *Elected John, John was.

As shown in (12), this is because the copy *John^3* in the topicalized vP does not form a chain with the copy of *John^2* in the subject position:

(12) $[_{v\text{P}} \, \text{Elected John}_i^3 \,]$, John_i^2 was $[_{v\text{P}} \, \text{elected John}_i^1 \,]$.

Chain 1 / Chain 2

Because of this, Gärtner (1998) claims that Chain Reduction cannot license the deletion of *John^3* in the topicalized vP.

---

2   Here and throughout, I label copies with a superscripted number as a means of referring to specific copies; this is for conveninence and is not represented theoretically; Copy Reduction cannot see these.

This is an apparent problem both empirically and theoretically. It is an empirical problem because Chain Reduction appears to incorrectly favor (11b) over (11a). This is a problem at theoretical level as well, because if there is no way to delete $John^3$, then it should not be possible to pronounce the structure schematized in (12) at all: It will always violate irreflexivity because it will never be possible to linearize $John^3_i$ relative to $John^2_i$.

As a response, Nunes proposes Chain Reduction is actually somewhat BLIND: An instruction to delete a single copy may delete more than one. In order to linearize $John^2$ and $John^1$, Chain Reduction receives the instruction to delete the copy of $John$ that is the sister of *elected*. The trick here is to assume that this instruction actually targets both $John^1$ and $John^3$: Each copy is the sister of a copy of *elected*. If this assumption is made, both copies will be deleted, resulting in (11a).[3]

This handles phrasal movement, as shown, but as we will see below it becomes problematic for the data discussed above. How cases of head movement are handled is contingent on the view of head movement that one adopts. As I discuss in the following section, if one adopts a non-movement approach to head movement phenomena, Nunes's approach to remnant movement works straightforwardly. In Section 4, I return to movement-based approaches and show that there is no non-stipulative way of accounting for the verb-doubling data. There we will see echoes of the empirical and theoretical problems discussed above, which do not arise under non-movement accounts.

3   Head movement as conflation

In this section, I introduce CONFLATION, a feature-based, syntactic mechanism that accounts for head-movement phenomena without recourse to syntactic movement.[4] I argue that this approach to head movement accounts for the phenomenon of verb doubling without further stipulation. This is significant because verb doubling is not one of the phenomena that conflation was originally meant to model, thus lending novel, independent support to this approach to non-movement approaches to head movement.

3.1   Non-movement approaches to head movement

Since Chomsky (2001:37) suggested that head movement may not be derived in the same way as phrasal movement, several authors have developed implementations of head movement that do not rely on movement, *per se*. The idea is that head movement can be modeled without appeal to the operation MOVE (that is, Copy + Merge). Chomsky himself proposes that head movement might be largely a PF operation, on the basis that it does not appear to show LF effects. Other recent implementations, including those proposed by Brody (2000), Hale & Keyser (2002), Harley (2004, 2013), Platzack (2013), and Zwart

---

3   One might reasonably worry that this is a fairly loose definition of '*the* sister of some element', given that definite descriptions usually carry some sort of uniqueness presupposition. Provided the assumed non-distinctness of copies, however, I do not think this is a completely unreasonable position to take – neither copy of the *v*P is distinct from the other.

4   Indeed, once one adopts a non-movement account of head movement phenomena, the term *head movement* becomes a bit of a misnomer. Nonetheless, the term *head movement* is the most broadly – if not the only – acknowledged term for the phenomenon in the Generative Linguistics literature, and so I will continue to use the term here.

(2001) are still properly syntactic, but do not rely on head movement being derived by the operation Move.[5]

3.2 Conflation

For the purposes of this paper, I adopt the CONFLATION approach to head movement (Hale & Keyser 2002, Harley 2004, 2013).[6] Under this view, morpho-phonological features are passed up the tree as structure is built; Hale and Keyser suggest that Conflation may be part of Merge. I adopt conflation primarily for concreteness, as other non-movement accounts of head movement should work just as well. Conflation is simply the most straightforward of the available approaches and requires the least modification to account for the data under discussion.

A common assumption in late-insertion models of the syntax–morphology interface is that individual morphemes are associated with individual syntactic heads (Embick & Noyer 2001:559).[7] Consequently, morphologically complex elements must be derived. In the verbal domain, it has long been assumed that this derivation is accomplished in the syntax by head movement (Travis 1984, Vikner 1995, Pollock 1989, amongst many others). Under the broadly assumed Minimalist implementation of this traditional view, head movement is effected by Move: Individual heads are copied from their base positions and adjoined (that is, merged) to the next highest head.

Under the conflation approach, morphologically complex elements are not formed by Move. Instead, the assumption is that the syntax underlying such elements is the same as if the heads had not moved. This means that an element with a PF surface form $Z°+Y°+X°$ must have the syntax $[_{XP} \ldots X° [_{YP} \ldots Y° [_{ZP} \ldots Z° \ldots]]]$.

Instead, morphologically complex elements are formed by sharing features up the tree as syntactic structure is built. The main assumption is that heads can come with a set of (morpho-) phonological features that can be shared with other heads. I will call these features π.[8] As Harley (2004) notes, we can think of these features as those which trigger lexical insertion in post-syntactic theories of morphology (*e.g.,* Vocabulary Insertion in Distributed Morphology); see also Platzack 2013. The main assumptions for sharing of these features are laid out in (13).

(13)   Key assumptions for Hale & Keyser's (2002) Conflation (based on Harley 2004):

   a.   The label of any constituent has ALL the features of the head, including some representation of a phonological matrix π.

   b.   Conflation occurs when a constituent a is merged with a sister head β whose set of features is 'defective'. The features π_a are merged into π_β.

   c.   For Economy reasons, the conflated set of features is only pronounced once, in its uppermost position.

---

5   Approaches that reduce verb movement to remnant vP movement, such as Müller 2004, also exist, but if head movement is reduced to phrasal movement, then the ideas in this poster will not work out.

6   See also Zwart 2001 for a similar though different idea.

7   This idea, in fact, antedates late insertion models considerably.

8   Harley refers to such features as the *p-sig* of a head. I find this label a bit to cumbersome, and so I adopt Platzack's (2013) convention instead.

This means that as the tree is built via Merge, the features π are passed up the tree, on the assumption that the label of a phrase shares all of the features of the head. Thus, as shown in the trees in (14), the features are associated with $Z°$, π_z are shared with $Y°$ when $Y°$ merges with $Z°$, and the features of $Y°$, $[\pi_z, \pi_y]$, are shared with $X°$ when $X°$ merges with YP. The result is that all of the phonological features $[\pi_x, \pi_y, \pi_z]$ wind up on a single head: $X°$. Assumption (13c) ensures that the features on $X°$ are pronounced to the exclusion of those on both $Y°$ and $Z°$.

(14)

$$
\begin{array}{c}
\text{YP} \\
[\pi_y, \pi_z]
\end{array}
\xrightarrow{\textit{Merge } X°}
\begin{array}{c}
\text{XP} \\
[\pi_x, \pi_y, \pi_z]
\end{array}
$$

YP $[\pi_y, \pi_z]$ — Y° $[\pi_y, \pi_z]$ Z° $[\pi_z]$

XP $[\pi_x, \pi_y, \pi_z]$ — X° $[\pi_x, \pi_y, \pi_z]$ YP $[\pi_y, \pi_z]$ — Y° $[\pi_y, \pi_z]$ Z° $[\pi_z]$
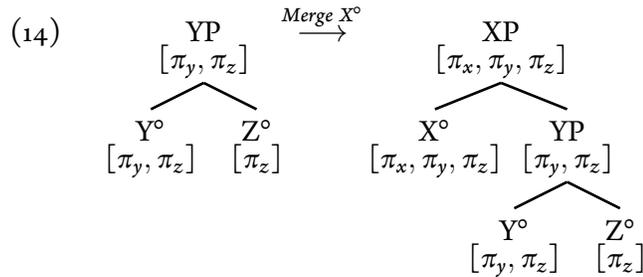
The upshot of all this is that the phonological features of one head come to be associated with another without appealing to the operation Move. This means that movement is not necessary for the features of one head to be displaced to another position. No additional copies of heads are generated

3.3 Verb doubling with conflation

If we assume head movement is derived via Conflation, then the derivation of verb phrase topicalization is rather straightforward. Verb phrase topicalization will copy vP and merge it in SpecCP. In so doing, it will generate an additional copy of the verb root $V°$ in the topicalized copy of the vP, but this will be the only additional copy of the verb generated:

(15) $[\,[\,v°\,[V°\,...\,]\,]^2_i\; C°\,[T°\,[\,v°\,[V°\,...\,]\,]^1_i\,]\,]$
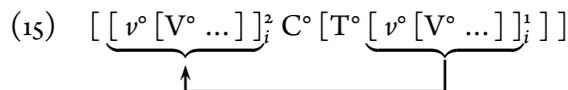
With conflation, there is only one chain containing a copy of the verb. Since there is only one chain, there are only two copies of the verb in the output of the narrow syntax: The copy of the verb in the base position, and the copy of the verb in the fronted vP. Chain Reduction predicts that the lower copy of the vP should be deleted, to the exclusion of $T°$. This leaves only the higher copy of vP (the copy in SpecCP) to be pronounced:

(16) $[\,[\,v°\,[V°\,...\,]\,]^2_i\; C°\,[\,T°\,\cancel{[\,v°\,[V°\,...\,]\,]^1_i}\,]\,]$
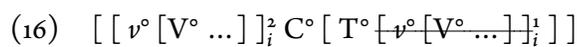
Crucially, the languages which display verb-doubling VPT are verb movement languages. These are languages in which verbs appear in an inflectional position outside of the verb phrase. Under Conflation, this means that these inflectional heads carry defective π features, following (13b), and thus require conflation with the verb's π features. Remember that this is not actual movement of the verb, but Conflation of the phonological features. This means that the $v°$ in the fronted vP will bear the features of $V°$, and that $T°$ will bear the features of $v°$ and $V°$, but the heads themselves remain *in situ*:

(17) $\left[\ \left[\ _{v^\circ \atop [\pi_V, \pi_v]}\ [\ V^\circ\ ...\ ]\ \right]_i^2 C^\circ \left[\ _{T^\circ \atop [\pi_T, \pi_V, \pi_v]}\ \left[\ _{v^\circ \atop [\pi_V, \pi_v]}\ [\ V^\circ\ ...\ ]\ \right]_i^1\ \right]\ \right]$

The fact that the verb's features conflate with T° explains how it is possible for the verb to be pronounced twice. When it comes time for Chain Reduction to occur, the lower copy of vP will be deleted as in (16) above. But as shown in (17), even though the lower copy of vP is deleted, features from v° and V° conflate onto T° before this happens. T° is not part of the moved material, so it is not targeted by Chain Reduction. This strands T° with the phonological features from v° and V°, leaving them to be pronounced:

(18) $\left[\ \left[\ _{v^\circ \atop [\pi_V, \pi_v]}\ [\ V^\circ\ ...\ ]\ \right]_i^2 C^\circ \left[\ _{T^\circ \atop [\pi_T, \pi_V, \pi_v]}\ \overline{\left[\ _{v^\circ \atop [\pi_V, \pi_v]}\ [\ V^\circ\ ...\ ]\ \right]_i^1}\ \right]\ \right]$

This means that there are now two heads with conflated phonological features: T° and the v° in the topicalized VP. Following (13c), lexical material is inserted into the uppermost head with conflated material. Assuming uppermost to be determined by c-command, these positions should be T° and v°.[9] When lexical insertion happens, the features on T° will trigger the insertion of the verb in T°, and the features on v° in the fronted vP will trigger the insertion of the infinitive.

3.4 Summary

Conflation explains straightforwardly how multiple copies of the verb are pronounced in vP topicalization. Movement of vP generates an extra copy of the verb root, as is typically assumed. Morpho-phonological features from the trace copy of V° conflate with those on T°. When Chain Reduction occurs, the trace copy of vP is deleted, but the features on T° survive since there is no instruction to delete T°. This leaves copies of the phonological features of V° in the topicalized vP and on T°, resulting in the verb's morphology appearing in two places.

In the discussion of remnant movement in Section 2.2, we saw that it is necessary to make certain assumptions about how deletion proceeds in order to ensure that trace copies of phrasal material in the topicalized vP are correctly deleted – Copy Reduction must be blind. This issue did not arise here, since the apparent movement of the verb did not generate any additional copies that needed to be reduced. In fact, strictly speaking, verb-doubling verb phrase topicalization is not even remnant movement if we adopt Conflation. As we will see in the following section, this allows us to sidestep many issues that come along with trying to linearize multiple chains in Nunes's (2004) approach to Chain Reduction. Since there is no head *movement*, we do not have to deal with linearizing copies of the same element in multiple chains.

4 Head movement as syntactic movement

In this section I turn to how verb-doubling is derived if head movement is modeled under the copy theory of movement, which I refer to as the CTM HYPOTHESIS, or just CTM for

---

9 We might need to be more careful with what *uppermost* means, but if we assume that it is calculated with reference to structural adjacency that appears to be part of the derivation of Conflation, T° is not in a local relationship with the v° in the topicalized vP.

short. I focus on Nunes's (2004) original proposal that certain morphologically complex heads can be reanalyzed, rendering them opaque to the linearization algorithm. This, I argue, is too stipulative. I also look briefly at two related proposals: Bastos (2001) proposes that infinitival morphology on the topicalized verb invokes morphological reanalysis, and Landau (2006) proposes the head of the topicalized vP must be prosodically focused. Neither of these explanations is empirically adequate.

4.1 Head movement as movement and morphological reanalysis

The CTM hypothesis posits that heads are copied from their base positions and adjoined to the next highest head (Roberts 2011). This process is iterative in the sense that the resulting complex head can subsequently be moved, resulting in more copies of the first head. This is demonstrated in (19), where iterated head movement results in three copies of Z°. (For simplicity, I will represent head adjunction with the + symbol.)

(19)

```
                    XP
                  /    \
        Z°+Y°+X°        YP
              ↖       /    \
               Z°+Y°        ZP
                    ↖      /    ⋱
                     Z°
```
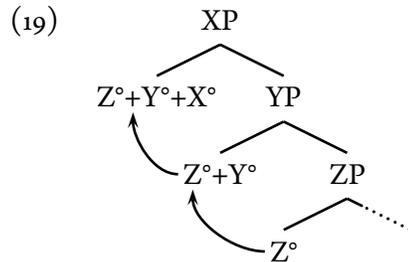
The main issue with modeling head movement under the CTM hypothesis is that it will require us to stipulate that certain links in the head movement chains are invisible to linearization. This is essentially the tack that Nunes (2004) takes, and I review this here. This does, of course, get us the right result, but there is no principle underlying when this is allowed to happen. Consequently, I will argue that the conflation approach is superior, since we need not specify that specific copies are invisible to the linearization algorithm.

If CTM hypothesis is right, VVPT will involve multiple chains: The vP movement chain and the head movement chains:

(20) $[\,[\,[\,\underbrace{V^5+v\,[\,V^4\,...\,]}_{}\,]\,]\,C\,[\,V^3+v+T\,[\,\underbrace{V^2+v\,[\,V^1\,...\,]}_{}\,]\,]\,]$
with CH 4, CH 2, CH 1, and CH 3 labeled on the chains.
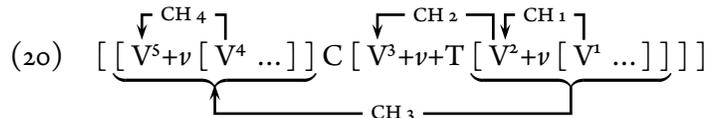
First, it is worth remarking on the straightforward cases of Chain Reduction. The lower copy of vP should be reduced, which will delete copies $V^2(+v)$ and V^1 since they are inside of this vP. V^2+v and V^1 would be deleted anyway. The copy of V^2 forms a chain with V^1, and this would license the deletion of V^1. V^3+v forms a chain with V^2+v, licensing the deletion of V^2+v. Finally, V^4 will be deleted since it forms a chain with V^5.

This leaves V^3 and V^5, the copies that are apparently pronounced under VVPT. This at first seems like a welcome result, but the fact is that under Nunes's original implementation, we expect verb doubling to be ungrammatical. V^3 and V^5 are copies of the same element and should therefore be treated as non-distinct. If irreflexivity holds, then one of these copies should need to be reduced since they should not be able to be linearized with respect to each other. V^5 is not part of the same chain as V^3, and so neither copy directly licenses the

deletion of the other. Despite the fact that they should violate irreflexivity, the structure is still pronounceable.

To further complicate matters, the posited blindness of Chain Reduction actually predicts that V^5 should be reduced. Recall from the discussion of phrasal remnant movement in Section 2.2 that Chain Reduction can over-apply in the sense that an instruction to delete a specific copy might delete multiple copies. This is exactly what we expect to happen here. Under Nunes's proposal, V^5+v should be reduced because V^3+v will license the the reduction of V^2+v, assuming that Chain Reduction targets the $v°$ that is a sister to VP. Indeed, reduction of V^5 would solve the linearization problem just described, but it predicts, incorrectly, that only one verb should be pronounced.

Essentially, we have the inverse of the problem described in Section 2.2. Recall that there we saw that phrasal material moved out of a topicalized vP could not be pronounced twice. The blindness of Chain Reduction was posited to account for this fact, but here it makes the incorrect prediction that verbal material should not be pronounced twice, either. So there are two problems: First, the two copies of the verb should not be linearizable. Second, it is predicted that one copy should be reduced when neither is.

Nunes reasons that since both copies are pronounced, it must be the case that one of them is somehow being ignored for the purposes of linearization, allowing them to bypass irreflexivity. He proposes that certain morphologically complex heads can be MORPHOLOGICALLY REANALYZED. This causes material internal to a (morphological) word to become invisible to the linearization process, and therefore, they will be ignored by Chain Reduction.

Nunes thus proposes that the V^3+v+T complex is morphologically reanalyzed, meaning that its subparts are no longer visible to the linearization algorithm. This has two effects. First, it means that V^3 will no longer be visible to the linearization algorithm. This means that it can now be linearized with regard to copies V^5, solving the linearization problem. Second, V^3+v is no longer visible to the linearization algorithm. Consequently, it no longer licenses the reduction of V^2+v. Since V^2+v is not targeted for reduction, V^5+v will not be targeted for reduction, either, explaining why V^5+v is pronounced.

4.2 The problem with morphological reanalysis

The main issue with this approach is that there is no evidence that morphological reanalysis actually happens in VVPT or any other construction. The only thing that it appears to do is allow for for multiple copies of the same element to be pronounced. There is no other proposed morpho-phonological reflex.

The problem is that this means there is no independent way to predict what will be morphologically reanalyzed and what will not be (Bastos 2001:117). It cannot be the case that V°-to-T° movement always invokes reanalysis, since V°-to-T° movement must be able, under the CTM hypothesis, to trigger the deletion of lower copies of the verb when there is no vP topicalization (Nunes 2004:169, n.40). If it did not, we would expect lower copies of the verb to be pronounced whenever verb movement has occurred. This is ungrammatical in all of the languages that I know of, including, *e.g.*, Portuguese:

(21) *O   João   lav-ou   lav-ar   o   carro.
     the   João   wash-PST   wash-INF   the   car.

One might be tempted to propose that V°+v°+T° must be reanalyzed when and only when there is verb phrase topicalization, but it is not clear exactly how one would invoke this in the grammar in a fully predictive model rather than in a way that simply restates the facts.

The fact is that Morphological Reanalysis is invoked only when it is necessary to explain why two things are pronounced. Nunes proposes that if morphological reanalysis occurs, then multiple copies of a head will be pronounced. But his analysis actually works in reverse: If two copies are pronounced, then the conclusion is that one of them must have been Morphologically Reanalyzed. Without any independent evidence that reanalysis occurs, this is practically equivalent to simply stating which copies will be pronounced – the very problem that we are trying to explain.

And this is why VVPT shows that the CTM hypothesis is problematic: Head movement generates a large number of copies that need to be correctly deleted, but there is no principled way of accounting for which copies must be deleted and which get pronounced. Under the conflation approach detailed in Section 3, none of these issues arise. In fact, there is no need to propose that something like morphological reanalysis exists.

4.3 Alternatives to Morphological Reanalysis

As noted above, the Morphological Reanalysis operation is what permits multiple copies of the same element to be pronounced in Nunes's (2004) system. If it could be shown that there is some principle behind when it applies, then many of the concerns about it would be alleviated. Alternatively, if some other principle forced copies to be pronounced independently, then it might be possible to replace morphological reanalysis with some other principle. However, I know of no convincing case that accounts for all the data under discussion.

Bastos (2001:126) proposes that the V° +v° is reanalyzed, claiming that the post-syntactic process that introduces infinitival morphology in the fronted vP induces Morphological Reanalysis of the verb. The typical view of this is that some post-syntactic operation inserts infinitival morphology either in v° or as a last-resort mechanism to render the structure pronounceable. Bastos argues that this infinitival morphology causes the resulting complex head to be distinct from the copy in T°, and therefore, the two copies can be linearized with respect to each other.

The problem with this is that not all fronted vPs have infinitival morphology. Passive morphology is doubled when passive vPs are fronted, and it is not possible to have infinitives in this case (Vicente 2009):

(22)   [Entreg-<u>ada</u>   / *entreg-<u>ar</u>   al   ganador]$_k$, la   medalla$_i$ ha   sido
       awarded-PASS.FEM   /   award-INF   to.the   winner,   the   medal   has   been
          entreg-<u>ada</u> t_k.
          awarded-PASS.FEM

'Awarded to the winner, the medal has been.'

Here, in addition to the copies of the verb root, the copies of the passive morphology should cause linearization problems since there is no reason to assume that the copies of the passive verb are non-distinct. Consequently, Bastos's proposal is not sufficient to account for all of the data.

Landau (2006) suggests that pronunciation of the verb in the fronted *v*P in Hebrew is the result of a phonological requirement, imposed by Top°, that the verb receive a characteristic intonation associated with topicalization. This could be taken as external motivation for pronouncing this copy of the verb that does not rely on Morphological Reanalysis.

But it is not clear that this explanation generalizes beyond Hebrew. As Landau (2006:39–40) discusses, Hebrew and Russian may impose different prosodic requirements on fronted verb phrases (see Abels 2001 on Russian). Consequently, it is not clear that the need to stress the verb in the topicalized *v*P is a viable cross-linguistic explanation for why that verb must be pronounced.

4.4 Summary: Why non-movement is better than movement

The reason the Conflation approach fares better than the copy-theoretic approach is because the latter generates so many additional copies of the verb. Under the copy-theoretic approach, we have to remove certain copies of the verb from the linearization computation to avoid linearization problems since the copies of the verb are not distinct.

The Conflation approach sidesteps this by generating fewer copies of the verb. Phonological features percolate up the tree. But the heads to which they percolate remain distinct since they are not copies. Conseqeuntly, we do not run into the linearization problems that are found if head movement is copy theoretic. The linearization problems simply do not arise.

This, on its own, is not a sufficient reason to reject the copy-theoretic approach. The real issue is that we are forced to stipulate which copies are removed from the computation or to stipulate reasons for their removal. This is, in essence, what Morphological Reanalysis is for. But there is no principled way of determining which copies will be reanalyzed and which will not. The CTM approach basically requires us to state that some copy will be reanalyzed but gives us no way to understand why certain copies must be reanalyzed.

This is what separates the CTM approach from the Conflation approach. VVPT falls out from the interaction of Conflation with standard *v*P movement. There is no need to appeal to Morphological Reanalysis or other outside considerations. Thus, although Chain Reduction is simpler under Conflation, this is not just a parsimony argument. Conflation makes better predictions about VVPT than does copy-theoretic head movement, and for this reason it is to be favored.

5   Conclusion and outlook

In this paper, I have argued that non-movement approaches to head movement phenomena make better empirical predictions for verb-doubling verb phrase topicalization than copy-theoretic approaches. Under Nunes's (2004) theory of Chain Reduction, there is no adequate way of deciding which copies will be pronounced at PF if head movement is derived in copy-theoretically. Non-movement approaches sidestep the issue, resulting in bet-

ter empirical coverage with fewer stipulations.

5.1 Approaches other than Conflation

As I mentioned in the Section 3, I adopt Conflation since it is the most straightforward non-movement implementation of head movement that does not need modification. Non-movement approaches other than Conflation should work similarly, although they may need further modification.

If head movement is truly a PF phenomenon, as Chomsky (2001) suggests, then the data under discussion in this paper should be accounted for just as well assuming that PF movement is not copying. Other approaches may fare differently. Platzack's (2013) Agree- and EPP-based formulation of head movement effects works similarly to Conflation. However, it relies on an EPP feature on one of the heads to trigger pronunciation of the head chain on that head. Some modification would need to be made to allow for double pronunciation of the verb. It is not clear to me that this is formulable without loosing some of the key insights of Platzack's approach. I am less sure about what predictions Brody's (2000) Mirror Theory might make. In particular, it remains unclear to me how phrasal movement of part of the head-complement structure would interact with the principle Mirror. Nonetheless, Mirror Theory should handle the data discussed in this paper straightforwardly.

5.2 Related phenomena

I have not addressed in this paper issues introduced by many Germanic languages. VPT in these languages often involves some version *do*-support and, in the right circumstances, can even involve apparently headless verb phrase fronting:

(23) a. % German (Bayer 2008)

[$_{vP}$ Einen guten Charakter besitzen] <u>tut</u>   der Klaus auf alle
a   good character own.INF does.PRES the Klaus in any
Fälle.
cases

'Klaus has a good character in any case.'

b. ? German (Müller 1998)

(Ich glaube) [$_{vP}$ Kindern   Bonbons t_i] gibt$_i$ man besser nicht
I   believe   children.DAT sweets.ACC   gives one better not.
t_{VP}.

'(I believe that) give candies to children, one had better not.'

The first case is problematic regardless of the theory of head movement one adopts, primarily because Germanic languages (other than English) typically require V°-to-C° movement in matrix clauses. This does not happen here; instead, the default verb *tun*, 'do', is inserted in C°. Nonetheless, verb movement to C° is predicted to happen here rather than

do-support regardless of whether one adopts Conflation of the CTM hypothesis. I suspect that this has more to do with how *do*-support is implemented in German than it does with head movement itself, and so I set this aside here.

The second case is more problematic for the hypothesis presented here. A copy of the verb should be pronounced in the position of the trace in the fronted verb phrase, but this does not occur. As discussed in Section 4.1, this is actually the way we expect remnant *v*P topicalization to behave if head movement acts like phrasal movement, following Nunes (2004). Landau (2006:61) points out that it is unclear how productive these constructions really are (hence the judgment of ? above), and so I will follow him and set them aside for future research. This could be a case of exceptional multiple topicalization, which is typically banned in German, and that might explain why speakers find it marginal.

5.3 Can we get rid of Morphological Reanalysis? Not yet

I argued above that the Conflation approach to head movement is better than the CTM approach because it does not need to appeal to Morphological Renalysis to explain the verb-doubling phenomenon. It is worth nothing, however, that Nunes uses Morphological Reanalysis to explain several other phenomena, including multiple *wh*-copies in Germanic, doubled clitics in Argentinian Spanish, duplication of postpositions in Panara, and focus duplication in Brazilian Portuguese.

A discussion of each of these phenomena is beyond the scope of this paper, suffice it to say that none of them clearly depend on head movement. As such, whatever benefits a non-movement analysis of head movement might confer to verb doubling cannot readily be transfered to these phenomena. However, the criticisms of Morphological Reanalysis raised in Section 4 remain. In these cases, Morphological Reanalysis is proposed because multiple copies of the same element are pronounced, but an explanation as to why it occurs is still lacking.

So long as clear, independent evidence that this operation has applied is lacking, it is profitable to seek alternative accounts of the phenomena it is meant to explain. That is not to say that no such process exists – it very well could – but the effects it is meant to explain may be the result of a heterogeneous set of morpho-syntactic processes. Nunes, for instance, suggests that Morphological Reanalysis might be a part of cliticization, and here I have suggested it falls out from the right theory of head movement. The hope is that if we had a handle on what this set of morpho-syntactic processes is and how they work, we should be able to predict exactly when pronunciation of multiple copies of the same element is possible.

the UMass Amherst Department of Linguistics.

References

Abels, Klaus. 2001. The Predicate Cleft Construction in Russian. In Steven Franks & Michael Yadroff (eds.), *Formal Approaches to Slavic Linguistics 9*, 1–19. Bloomington, IN: Michigan Slavic Publications.

Bastos, Ana Cláudia Pinto. 2001. *Fazer, eu faço! Topicalização de constituintes verbais em português brasileiro*. Universidade Estadual de Campinas MA thesis.

Bayer, Josef. 2008. What is Verb Second? Ms. Universität Konstanz. http://ling.uni-konstanz.de/pages/home/bayer/pdf/Verb-Second.pdf.

Brody, Michael. 2000. Mirror Theory: Syntactic Representation in Perfect Syntax. *Linguistic Inquiry* 31(1). 29–56. https://dx.doi.org/10.1162/002438900554280.

Cable, Seth. 2004. Predicate Clefts and Base Generation: Evidence From Yiddish and Brazilian Portuguese. Ms. Massachusetts Institute of Technology. http://people.umass.edu/scable/papers/Phon-N-V-Dissim.pdf.

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

Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, Mass: MIT Press.

Embick, David & Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32(4). 555–595. https://dx.doi.org/10.1162/002438901753373005.

Gärtner, Hans-Martin. 1998. Review of "The Copy Theory of Movement and Linearization of Chains in the Minimalist Program". *GLOT International* 8(3). 16–20.

Hale, Kenneth & Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, Mass.: MIT Press.

Harley, Heidi. 2004. Merge, Conflation and Head Movement: The First Sister Principle revisited. In Keir Moulton & Matthew Wolf (eds.), *Proceedings of the North East Linguistic Society 34*, Amherst, Mass.: GLSA Publications.

Harley, Heidi. 2013. Getting Morphemes in Order: Merger, Affixation, and Head-movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing Syntax*, Oxford: Oxford University Press.

Källgren, Gunnel & Ellen F. Prince. 1989. Swedish VP-Topicalization and Yiddish Verb-Topicalization. *Nordic Journal of Linguistics* 12. 47–58. http://dx.doi.org/10.1017/S033258650000192X.

Kayne, Richard. 1994. *The Antisymetry of Syntax*. MIT Press.

Landau, Idan. 2006. Chain Resolution In Hebrew V(P)-fronting. *Syntax* 9(1). 32–66. https://dx.doi.org/10.1111/j.1467-9612.2006.00084.x.

Müller, Gereon. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.

Müller, Gereon. 2004. Verb-Second as vP-First. *Journal of Comparative Germanic Linguistics* 7(3). 179–234. https://dx.doi.org/10.1023/B:JCOM.0000016453.71478.3a.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement* (Linguistic Inquiry Monographs 43). Cambridge, Mass: MIT Press.

Platzack, Christer. 2013. Head Movement as a Phonological Operation. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing Syntax*, 21–43. Oxford: Oxford University

Press.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20(3). 365–424. http://www.jstor.org/stable/4178634.

Roberts, Ian. 2011. Head Movement and the Minimalist Program. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, Oxford: Oxford University Press.

Travis, Lisa DeMena. 1984. *Parameters and Effects of Word Order Variation*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Vicente, Luis. 2007. *The syntax of heads and phrases: A study of verb (phrase) fronting*: Universiteit Leiden dissertation. http://www.lotpublications.nl/Documents/154_fulltext.pdf.

Vicente, Luis. 2009. An Alternative to Remnant Movement for Partial Predicate Fronting. *Syntax* 12(2). 158–191. https://dx.doi.org/10.1111/j.1467-9612.2009.00126.x.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages* Oxford Studies in Comparative Syntax. New York: Oxford University Press.

Zwart, C. Jan-Wouter. 2001. Syntactic and Phonological Verb Movement. *Syntax* 4(1). 34–62. https://dx.doi.org/10.1111/1467-9612.00036.