Distinguishing Copies and Repetitions

Chris Collins (New York University)
Erich Groat (University of Nottingham)

Abstract: The purpose of this paper is to review proposals made concerning the difference between copies and repetitions in minimalist literature. Our conclusion is that if one adopts the well-motivated and austere assumptions of minimalist syntax, then every proposal we examine faces serious shortcomings. Although the difference between copies and repetitions can be maintained in some of these analyses, none of them explain the distinction, but merely provide stipulations capturing it.

Keywords: copies, repetitions, chains, occurrences, multi-dominance, phases, phase-level memory

1. Introduction: Copies and Repetitions

In minimalist syntax, the structure building operation is Merge, defined as in (1) (see Collins 2002, Seely 2006 and Chomsky 2013):

(1) Merge(X,Y) = {X, Y}

If X and Y are independent (neither one contains the other), then (1) is an instance of External Merge. If X contains Y (or vice versa) then (1) is an instance of Internal Merge. Internal and External Merge are two separate cases of the Merge operation, not distinct operations (see Chomsky 2004).

As an illustration of Internal Merge, Consider the following passive sentence:

(2) John was seen <John>.

In (2), *John* has been externally merged as the complement of the passive verb *seen*. Subsequently, *John* is internally merged as the subject of the clause. In minimalist literature, these two occurrences are called copies since they are related by Internal Merge. The angled brackets around the rightmost occurrence of *John* in (2b) indicate non-pronunciation (at Transfer/Spell-Out/Externalization). For ease of reference, we will henceforth refer to the leftmost occurrence of *John* as the first occurrence, and the rightmost occurrence as the second occurrence.

The structure of (2) needs to be distinguished from the structure of (3):

(3) John saw John.

In (3), *John* has been externally merged as the complement of the verb. Subsequently *John* is externally merged as the subject of the clause. In minimalist literature, these two occurrences are called repetitions (see for example Chomsky 2013: 40), since they are not related by Internal Merge. Note that in this case, both occurrences are pronounced.

The question is how to distinguish examples like those in (2) and (3). That is, how does one distinguish the notion of copy from the notion of repetition? Concretely, how does Transfer/Spell-Out/Externalization determine that the two occurrences in (2) are related by Internal Merge, but the two occurrences in (3) are not?

In this paper, we will consider a number of proposals that have been given in the literature. We will show that in each case, the proposal faces difficulties. In particular, any solution to the issue of distinguishing copies and repetitions consistent with minimalist aims must meet the following criteria: (a) no operations other than Merge should be used to build structure, (b) nothing beyond lexical items and the structures built from them by Merge should be interpreted by the interfaces, and (c) the definition of Merge should not be made more complex than the definition in (1). No current proposal satisfies all of these criteria. We conclude that no adequate proposal exists in minimalist syntax for distinguishing copies and repetitions.

We divide the presentation into three sections based on three basic types of analysis: Added Structure Approaches (section 2), Configurational Approaches (section 3), and Phase-Level Memory Approaches (section 4). We analyze representative analyses for each approach. For brevity's sake, we will not analyze computational approaches (such as Stabler 1997) that adopt assumptions that lie outside the austere regime we are assuming.

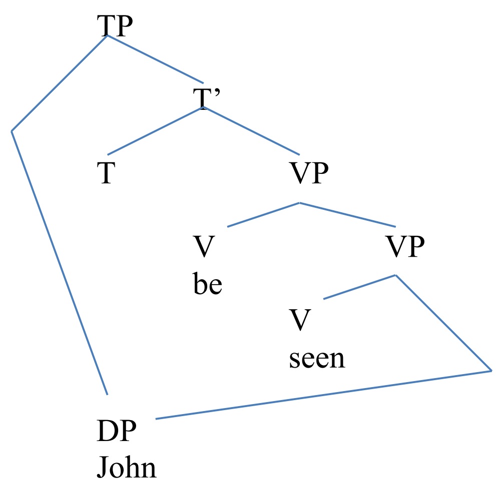
In order to limit the length of the article, we do not give definitions of widely used notions such as *contain*, *occurrence*, *syntactic object*, *workspace*, etc. See Collins and Stabler 2016 for explicit definitions of these notions.

2. Added Structure Approaches

2.1 Multidominance

In a multidominance theory (see Citko 2011a, 2011b for overview and extensive references), the sentence in (2) has the following structure. (We assume labels, and put aside irrelevant details such as the nature of the implicit argument and the v/vP.)

(4)



In (4), *John* is immediately dominated twice, once by the TP and once by the VP.

One issue that comes up right away is that (4) is a graph theoretic object. In minimalism, Merge forms sets {X,Y}, so the correct representation of (4) would be:

(5) {John, {T, {be, {seen, John}}}}

Compare (5) to the structure of (3), which is given in (6). (We ignore v/vP and the vP internal subject for convenience.)

(6) {John, {T, {see, John}}}

In both (5) and (6), *John* has two occurrences. We require that there is only one syntactic object in (5) with two occurrences, but in (6) there are two syntactic objects, each with its own occurrence. However, in the multi-dominance approach, nothing in the structures in (5) and (6) allows one to conclude that there is one syntactic object in (5), but two syntactic objects in (6). This remains the case regardless of the specific definition of occurrence used (for example, the "sisterhood" definition of Chomsky 2001, or the "path" definition of Collin and Stabler 2016).

In order to implement this intuition, something additional needs to be added to the representations. One possibility is a diacritic, like an index (see Chomsky 1995, Collins and Stabler 2016). Given this modification, the structures are as follows:

(7) {John_1, {T, {be, {seen, John_1}}}}
(8) {John_1, {T, {see, John_2}}}

Now $John_1$ has two occurrences in (7), whereas in (8) $John_1$ and $John_2$ have one occurrence each. (7) can be taken as the minimalist Merge based implementation of the multi-dominance analysis.

The problem with the structures in (7) and (8) is that they require a diacritic to work. And diacritics of this nature are clear violations of the Inclusiveness condition, stated in (9):

(9) **Inclusiveness** (Chomsky 2001: 2-3)
[Inclusiveness] bars introduction of new elements (features) in the course of computation: indices, traces, syntactic categories or bar levels, and so on.

Chomsky (1995) already recognized this problem. The use of indices or diacritics, he writes, "… is a departure from the Inclusiveness condition, but one that seems indispensable: it is rooted in the nature of language, and perhaps reducible to bare output conditions." (Chomsky 1995: 227).

A number of authors, some of whom we will discuss here such as Kitahara 2000: 153-154, Martin and Uriagereka 2011, and Muñoz Pérez 2018, have criticized the use of indicies in the sense of Chomsky's quote precisely because it violates Inclusiveness.

Note that minimalist interpretations of multi-dominance work within a system of Merge-based structure building, in which all syntactic objects are either LIs or sets: for X to be immediately dominated by both SO_1 and SO_2 means that the set X is a member of the sets SO_1 and SO_2. It might be thought that a richer system, such as the formalisms of graphs or trees, might allow us to capture the copy/repetition distinction more easily. But such a move is quite undesirable from a minimalist perspective. First, it would complicate the definition of syntactic object: graphs and trees are relations (sets of ordered pairs) between some set of nodes; such relations and nodes are not natural to minimalism. For example, we would need to express immediate dominance relations as pairs, such as {<X, Y>, <X, Z>}, instead of simply having

Merge create the set X={Y, Z}.

But second, and more importantly, such a move does not provide any new insights into the distinction we desire to make, but only forces us back to diacritics. Imagine the case of two occurrences of Z in a multidominance structure, such that Z is dominated by both A and B (i.e. the graph includes the ordered pairs <A, Z> and <B, Z>). What information could be added here that would tell us whether these two relations pointed to copies or repetitions of Z? Again, it would seem that the only way would be via some diacritic on Z, say Z_1 vs. Z_2. Of course this is no improvement.

One might think that a more restricted sort of graph, such as the traditional tree, which explicitly bans multi-dominance, could come to the rescue. But in fact the situation gets slightly worse. Barring diacritics, nothing distinguishes the Z elements in <A,Z> and <B,Z>. But any Z dominated by two nodes through movement would create a banned structure: it would not be a tree. Movement would not be possible.

Thus, multi-dominance approaches do not shed light on the copy/repetition distinction, always requiring some sort of additional indexation/diacritic. We do not deny that the copy/repetition distinction can be implemented in some way in a multi-dominance approach, but the approach itself does not naturally lead to the distinction.

2.2 Chains

In a chain based theory (see Collins and Stabler 2016 for formalization and references), (2) is distinguished from (3) in terms of the chains formed. A chain may be defined as a sequence of occurrences. The first occurrence in the sequence is the head of the chain. The last in the sequence is the tail. An occurrence is defined as a location in a tree (as before, the specific definition of occurrence makes no difference to the following points). The chains in (2) and (3) are illustrated in (10):

(10) a. One chain: <first occurrence of *John*, second occurrence of *John*>
b. Two chains: <first occurrence of *John*>, <second occurrence of *John*>

In (10a), *John* has two occurrences and these occurrences are linked by a chain. In (10b), *John* once again has two occurrences, but they are not linked by a chain, rather there are two trivial (one member) chains. On this theory, there is no need for diacritics (unlike in the multi-dominance approach). The work of the diacritics is being done by the chains.

The main problem with chain based accounts is the massive machinery that they require to get off the ground. First, a chain is a sequence of occurrences, so minimally one needs definitions of chain and occurrence. Second, one needs to assure that when Internal Merge takes place, a chain is formed. Third, formation of a chain goes way beyond the simple definition of Merge in (1), which does not form chains. In effect, one is introducing a new operation (Form Chain, see Nunes 2004) into the theory. Lastly, one needs to define Transfer/Spell-Out/Externalization to make reference to chains rather than simply to syntactic objects created by Merge. All of these steps are non-trivial (on dispensing with chains, see Epstein and Seely 2006, chapter 2).

3. Configurational Approaches

3.1 Theta-Role Distinctions

One property that distinguishes copies and repetitions, at least in the case of A-movement, is theta-role assignment. Consider the following configuration of occurrences of a given DP:

(11) $DP_1 \ldots DP_2$
where DP_1 c-commands DP_2, and there is no DP_3 c-commanded by DP_1 and c-commanding DP_2.

Here is a simple proposal for distinguishing repetitions from copies.

(12) If DP_1 is in a theta-position in (11), then DP_1 and DP_2 must be repetitions. Otherwise they are copies.[1]

Consider the example (13) below, in which we have shown the subject theta-position in Spec vP (we continue to ignore object shift for ease of exposition). Here, the first *John₁* in Spec TP is not in a theta position; it is therefore a copy of *John₂*, which it c-commands. (Note that no relation is established with *John₃*, since John₂ intervenes.)

(13) {John_1, {T, {John_2 {v {see, John_3}}}}}

However, *John₂* in Spec vP is in a theta position; therefore, it is a repetition of John_3, which it c-commands.

But theta-roles are *semantic* notions. Minimalist assumptions would require that the phonological component is effectively oblivious to any semantic information. Thus the phonological component has no access to the distinction as defined in (12) above. But we know that the distinction is crucial to the determination of which occurrences are pronounced. Thus, recourse to theta-roles does not appear to be a plausible explanation of the tight *connection* between phonological and semantic consequences of the distinction.

Even if theta-roles are features visible to syntactic computation, as in Hornstein (1999), the same problems arise. The question is whether these features, and the fact that they have been checked, are visible at Spell-Out, since they are not phonological features.

Furthermore, it is not only arguments receiving a theta-role that are capable of movement, meaning there are further cases to contend with. PPs, for example, commonly undergo movement, so (12) above would have to be revised; as it stands, all occurrences of a given PP would be copies, yet we know that identical PPs co-occur, such as the PP *on Mondays* below:

(14) On Mondays I think that on Mondays he drinks.

Factors beyond theta-roles are clearly involved.

Interestingly, it is notable that were such an approach plausible, it would make the distinction between External and Internal Merge irrelevant for the determination of the copy/repetition distinction; all structure-building could be External Merge, with the distinction

---

[1] Similar proposals have been discussed in personal communication with Noam Chomsky, Dennis Ott, Andreas Blümel, and Daniel Seely.

falling out of the semantic interpretation of theta-roles. The consequences of this (for example, in the domain of Probe-Goal relations) would need to be explored.

3.2 Phasal Distinctions

Martin and Uriagereka (2014) suggest an approach in which lexical items, and the syntactic objects built from them, are treated within Narrow Syntax purely as non-individuated syntactic objects (without any indices or diacritics to distinguish copies from repetitions). They put it this way: "Our proposal is that the distinction between repetition and copy depends solely on syntactic context: if two occurrences of lexically identical elements occur close enough to one another (in a sense to be determined), the system regards them as the same." (pg. 172) The process exploits the Phase Impenetrability Condition of Chomsky (2000):

(15) **Phase-Impenetrability Condition (PIC)**
In a phase α with head H, the domain of H is not accessible to operations outside of α, only H and its edge are accessible to such operations.

Generally, when two occurrences are accessible to each other by the PIC, they are interpreted as two copies of (a chain involving) a single item; if they are not, they are treated as separate repetitions. They state: "…all identical syntactic objects contained in the domain of a transferred phase head being interpreted as copies." (pg. 174)

They provide the following example for A-movement. Martin and Uriagereka (2014: 173) specify that "…here and below indices are given for clarity and are not intended to be part of the representations…".

(16) a. Students believe that students were criticized.
b. [$_{CP}$ [$_{TP}$ students$_4$ [$_T$ [$_{vP}$ students$_3$ [$_v$ [$_{VP}$ believe [$_{CP}$ that [$_{TP}$ students$_2$ [$_T$ [$_{PartP}$ were [$_{VP}$ criticized students$_1$]]]]]]]]]]]

The claim is that the two lower occurrences, *students$_1$* and *students$_2$*, are interpreted as an A-chain made of two copies of *students,* while two the higher occurrences, *students$_3$* and *students$_4$*, form a separate A-chain, also of two copies. These pairings reflect precisely which occurrences are accessible to the others by the PIC: note that *students$_1$* and *students$_2$* are not accessible to *students$_3$* and *students$_4$*, due to the intervening phase CP, which has no occurrence of *students* on its edge, effectively breaking the set of occurrences into two chains.

However, major issues arise immediately for basic cases of A′-movement:

(17) a. Guess [$_{CP}$ who students criticized]
b. [$_{CP}$ who$_3$ [$_{TP}$ students [$_T$ [$_{vP}$ who$_2$ [students [$_v$ [$_{VP}$ criticized who$_1$]]]]]]]

Martin and Uriagereka point out that in (17b), *who$_1$* would not be accessible to *who$_2$*, since the lower occurrence is rendered inaccessible by Transfer of the vP phase.

It seems that any solution will necessarily involve a formal distinction between Internal and External Merge in their approach. That is, occurrences formed by IM always constitute a single chain. As they state: "…to account for properties of A-bar chains, we concluded that internal merge in some sense creates chains/copies immediately…" (pg. 178). This undermines the strictly configurational solution.

3.3 Featural Distinctions

Muñoz Pérez (2018) advances an approach that allows two or more syntactic objects that are not featurally identical to be related and form a single chain at the interfaces. Restricting identification to syntactic features only, he examines the way unvalued features are valued in the course of a derivation; these features alone will determine what copies of an SO may form a chain not through identity but through inclusion: if A includes all of the features values of B, then A and B will be interpreted as a single chain. (For brevity, we put aside certain details of his proposal that are orthogonal to our discussion.)

Muñoz Pérez follows Adger and Svenonius (2011) in taking the syntactically active components of syntactic objects to be exclusively their formal features, which are organized as pairs. In the case of a valued feature, the value is drawn from set of values, such as ACC/NOM, +/–wh, and so on:

(18) **Valued feature**

a. A valued feature is an ordered pair <Att,Val>, where
b. Att is drawn from the set of attributes, {A, B, C, D, E, …}, and
c. Val is drawn from the set of values, {a, b, …}

In the case of unvalued features, the value is represented as Ø, the empty set, following Adger 2010:

(19) **Unvalued feature**

a. An unvalued feature is an ordered pair <Att, Ø> where
b. Att is drawn from the set of attributes, {A, B, C, D, E, …}, and
c. Ø needs to be replaced with an element from the set of values, {a, b, …}

For Muñoz Pérez, movement from a position in which SO has an unvalued feature to a (c-commanding) position where that feature is valued creates a distinct occurrence of the SO; the feature receives its value in the raised position alone. Consider the following example (see Munoz Perez 2018: 6):

(20) [Cosmo [T was [arrested Cosmo]]]

The derivation begins with *Cosmo* in its base position with an unvalued Case feature K, i.e. it bears the unvalued syntactic feature <K, Ø>. Although Muñoz Pérez does not specify exactly what the valuation mechanism is, he presumes that movement of *Cosmo* to a Case position will create a copy of that syntactic object, but with its K feature valued NOM:

(21) [Cosmo$_{\text{<K, NOM>}}$ [T was [arrested Cosmo$_{\text{<K, Ø>}}$]]]

These two "occurrences" of *Cosmo* are not in fact the same syntactic object, since they differ in feature-values. How then are they identified as copies, rather than independent repetitions? The answer lies in examining the feature values they have: The lower *Cosmo* does not bear any valued feature that the higher *Cosmo* lacks; the higher copy "includes" the feature values of the lower. By hypothesis, then, they may be considered non-distinct. Muñoz Pérez formalizes this hypothesis with the principle Inclusion-S (his example 14):

(22) **Inclusion-S**

A constituent β is non-distinct from a constituent α if for every value of β there is an identical value in α.

Muñoz Pérez introduces this notion of non-distinctness into a principle of chain formation, replacing the condition of coindexation with an Inclusion condition:

(23) **Chain Condition**

Two constituents α and β are part of the same chain iff

a. α c-commands β,
b. β is non-distinct from α (by Inclusion-S), and
c. there is no δ between α and β such [that] (i) β is non-distinct from δ, or (ii) δ is non-distinct from α.

In (21) above, we see that all three conditions are met for the two *Cosmos*: (a) the first (leftmost) c-commands the second, (b) the second is non-distinct form the second (since it has no valued features distinct from the valued Case feature of the higher; in fact it has no valued features at all), and (c) no other non-distinct element intervenes.

Implicit here is the condition that, in a given chain, every feature be valued on (at least) one member of the chain. This is satisfied in (21), since the higher *Cosmo* of the chain has a valued Case feature, though the lower one does not.

Contrastingly, in the following example, the first *Cosmo* is a part of the same chain as the second as before, and the third is part of the same chain as the fourth. However, the third and fourth are in a separate chain from the first two (we assume valuation of Accusative in Spec VP of the ECM verb):

(24) a. Cosmo is believed Cosmo to be ill.
b. [Cosmo$_{\langle K,\ NOM\rangle}$ T is [$_{vP}$ Cosmo$_{\langle K,\ Ø\rangle}$ v [$_{VP}$ Cosmo$_{\langle K,\ ACC\rangle}$ [believed [to be [Cosmo$_{\langle K,\ Ø\rangle}$ ill]]]]]]

The second *Cosmo* c-commands third, but the third bears a valued Case feature <k, ACC> that the second lacks, in violation of condition (b) in (23) above. They are therefore in distinct chains, as desired.

This analysis replaces indexation with an intuitive notion of non-distinctness that emerges in the course of feature-valuation. Muñoz Pérez naturally extends this analysis to wh-movement, which for him involves the valuation of wh-features (his ω-features).

Though this approach has intriguing properties, we believe it faces several difficulties. First, it still assumes the creation of some sort of a data structure called a "chain," an object not created by Merge, but nonetheless legible to both PF and LF interfaces; this is undesirable for the reasons mentioned in section 2.2. Second, how exactly are features valued? After all, a *new* syntactic object is created in Muñoz Pérez's system by *copying* the old one and replacing an unvalued feature with a valued one; only later is it united with it to form a chain. In the case of a complex syntactic object, would the entire structure need to be rebuilt? If so, do we really have any notion of "movement" here, or is structure-building entirely the result of External Merge (see also discussion of theta-approaches above)? Third, why is non-distinctness asymmetrical with respect to c-command; that is, why must the lower syntactic object not bear any valued feature borne by the higher, but not vice-versa? Why does (23b) not *also* include the condition "α is non-distinct from β?" This would entail a requirement for complete non-distinctness, which is after all the very definition of identity of sets.

Finally, and most importantly, we believe there is a significant counterexample to Muñoz Pérez's analysis, which we will analyze in some detail. Consider

(25) Cosmo, we believe Cosmo to adore.

We standardly assume that the topicalized *Cosmo* has undergone A-bar movement. In Muñoz Pérez's analysis, A-bar movement results in a chain whose elements bear a ω-feature, unvalued in all positions but Spec CP in which it is valued (via some left-peripheral feature; we will assume focus, or "FOC", as its value here). Here the valued position is the matrix Spec CP. We will see that a problem emerges due to the identically valued Case features of what we *want* to be two distinct repetitions of *Cosmos* originating in two theta-positions of the embedded clause; this identity creates ill-formed, uninterpretable chains.

Let's go through the derivation. We will assume that a Case feature is valued ACC in the specifier of the VP within in a transitive vP (see Chomsky 2008, 2013; we will ignore the head-movement associated with such analyses). Notationally, we will distinguish what SOs that we *hope* to occur in distinct repetitions with a diacritic ´, i.e. *Cosmo* vs. topicalized *Cosmo´*.

First, we build the embedded vP. *Cosmo´* begins with unvalued Case (K-) and ω-features; *Cosmo* with just an unvalued K-feature. At this point, assuming cyclic structure-building, *Cosmo´* has raised to its Case-valuing position:

(26) [$_{vP}$ Cosmo$_{<K,\ Ø>}$ [v [Cosmo´$_{<K,\ ACC>,<ω,\ Ø>}$ [adore Cosmo´$_{<K,\ Ø>,<ω,\ Ø>}$]]]]

Now, there are two cases to consider: one-fell-swoop movement of *Cosmo'* to Spec CP, and successive-cyclic movement of some kind.

Let us first consider one-full-sweep movement of *Cosmo´* to Spec CP, with no intermediate copies made. First, we build the matrix VP, with the external argument *Cosmo* valuing Case in Spec VP of *believe*:

(27) [$_{VP}$ Cosmo$_{<K,\ ACC>}$ [believe [T to [$_{vP}$ Cosmo$_{<K,\ Ø>}$ [v [Cosmo´$_{<K,\ ACC>,<ω,\ Ø>}$ [adore Cosmo´$_{<K,\ Ø>,<ω,\ Ø>}$]]]]]]]

Next, we add the matrix vP and TP projections (including the internal argument *we*), and move *Cosmo´* into Spec CP:

(28) [**Cosmo´**$_{\mathbf{<K,\ ACC>,<ω,\ FOC>}}$ [C [we$_{<K,\ NOM>}$ [T [we$_{<K,\ Ø>}$ [v [$_{VP}$ **Cosmo**$_{\mathbf{<K,\ ACC>}}$ [believe [T to [$_{vP}$ Cosmo$_{<K,\ Ø>}$ [v [Cosmo´$_{<K,\ ACC>,<ω,\ Ø>}$ [adore Cosmo´$_{<K,\ Ø>,<ω,\ Ø>}$]]]]]]]]]]]]

The intervening occurrences of *we* (which form a chain of their own) can be factored out of the evaluation of distinctness by part (c) of the Chain Condition.[2] But note that the boldfaced *Cosmos* are non-distinct: the lower *Cosmo* has no valued feature that the *Cosmo´* c-commanding it lacks. Furthermore, although the DP *we* intervenes, there are no non-distinct interveners – so all three conditions of (23) above are satisfied. *Cosmo´* in Spec CP and *Cosmo* in Spec VP are incorrectly predicted to be elements of a single chain.

Also included in this chain is *Cosmo*$_{<K,\ \emptyset>}$ in the embedded Spec vP, since it is non-distinct from *Cosmo*$_{<K,\ ACC>}$. This chain would thus appear to Topicalize the subject, not the object, of the embedded clause. Finally, the final two occurrences of *Cosmo´* (in embedded Spec VP and complement of *adore*) form a second chain. This chain is ill-formed, as none of its elements have a valued ω-feature. Such a derivation fails to correctly generate (25) above.

However, under phase theory, elements undergoing A-bar movement out of a phase are required to move to the Edge of any phase it moves out of, resulting in successive-cyclic movement leaving intermediate copies. Let us consider a derivation of that kind. Below, the *Cosmo´* seeking to value its ω-feature has moved to the edge of the lower vP phase to a position above the subject:

(29) [$_{vP}$ Cosmo´$_{<K,\ ACC>,<\omega,\ \emptyset>}$ [$_{v'}$ Cosmo$_{<K,\ \emptyset>}$ [v [Cosmo´$_{<K,\ ACC>,<\omega,\ \emptyset>}$ [adore Cosmo´$_{<K,\ \emptyset>,<\omega,\ \emptyset>}$]]]]]

At this point in the derivation Transfer applies (perhaps simultaneously with movement; see Chomsky 2000). The interfaces must herewith be provided with the information of what the chains are in (29). The problem arises immediately: *Cosmo* in the lower Spec v´ is non-distinct from the *Cosmo´* in the Spec vP that c-commands it; they therefore form a chain, as if it were the subject, not the object, that was undergoing A-bar movement. Meanwhile, a distinct A-chain is formed between by the two *Cosmo´s* in VP. This chain has no elements on the phase edge, and thus no hope of ever valuing its unvalued ω-feature. Such a derivation cannot converge.

Thus, under standard assumptions of structure-building and movement, (25) above does not seem to be derived.

A single counterexample does not bring down any theory, of course, but the nature of the analysis of this case strongly suggests that Inclusive-S is too weak a condition for distinguishing repetitions (elements in separate chains) from copies (elements in the same chain).

3.4 Select and Case

Kitahara 2000 proposes distinguishing copies and repetitions of pronouns by relating the distinct Select operations that introduce the pronouns into the derivation to the checking of the Case

---

2 There is, however, a problem with the Chain Condition in (23) as it stands: Successive-cyclic movement creates structures with interveners that prevent the formation of a single chain containing all of the elements. Consider movement of John below:

(i) Cosmo´´$_{<K,\ NOM>}$ seems <Cosmo´>$_{<K,\ \emptyset>}$ to have been misunderstood <Cosmo>$_{<K,\ \emptyset>}$.

According to (23), *Cosmo´´* and *Cosmo* cannot be in the same chain: *Cosmo´* intervenes, *Cosmo´* is non-distinct from *Cosmo´´*, and *Cosmo* being non-distinct from it, violating (23.c). This might be addressed by having the condition characterize not co-membership in a chain, but something like “successive links in a chain.”

feature of those pronouns. He proposes the following procedure:

(30) An NP gets interpreted upon the checking of its Case feature in the course of a derivation.

Given this principle, consider the following two sentences with their structures:

(31) a. He seems to be believed to be smart.
b. [$_{TP}$ He seems [$_{TP}$ he to be believed [$_{TP}$ he to be [AP he smart]]]]

(32) a. He thinks that he is smart.
b. [$_{TP}$ He [$_{vP}$ he thinks that [$_{TP}$ he is [$_{AP}$ he smart]]]]

Kitahara notes that in the derivation for (31b), there is only one Case-checking position for *he* (the matrix Spec TP position). Therefore, by (30), there is a single interpretation of *he*. In (32b), there are two Case-checking positions for *he* (the matrix Spec TP and the embedded Spec TP). Therefore, there are two interpretations of *he*. Crucially, in obtaining this result, no use of indices or diacritics is made.

Kitahara's approach raises numerous questions. First, what exactly is meant by "interpretations" of a pronoun? The notion of separate interpretations of a SO is meant to yield results obviously related to the existence of separate repetitions, but it is not clear how interpretation proceeds. Is this a matter solely of semantic interpretation, or phonological interpretation, or both? Does "to be interpreted" mean to be assigned a theta role? What about the phrases containing the yet-to-be-interpreted elements: do they too fail to be interpreted until every feature is checked? These matters are unclear.

Second, Kitahara seeks to connect the notion of independent operations of Select of an LI with independent interpretations of that LI. "So, in each derivation, the number of applications of Select, accessing *he*, equals the number of times *he* is interpreted…" (Kitahara: 156) Following Chomsky (1995), he presumes that a Numeration keeps track of the number of times an LI is selected by means of an index. For example, (31) above will require only one selection of *he*, while (32) will require two; the numerations will then be (33) and (34) respectively:

(33) N1=(… , (he, 1), …)
(34) N2=(… , (he, 2), …)

Each time a lexical item is used, the index is decreased by one; for any Numeration, the index of each LI must be reduced to zero by the end of the derivation.

However, it is not the case that selection is directly associated with interpretation in this way. Consider Kitahara's structure (11d) in example (35) below.

(35) {$_{TP}$ he seems {$_{TP}$ he to be believed {$_{TP}$ he to be {$_{AP}$ he smart}}}}

Kitahara claims that *he* raises successive-cyclically from its base position in AP to Spec of the matrix TP. In the specifier of the matrix TP, the case of *he* is deleted, and *he* is interpreted. "The derivation of [35] thus performs the single interpretive procedure of *he* applying at this point, thereby yielding a single interpretation of *he*."

However, Kitahara's principles seem also to be consistent with the following derivation.

He is selected and externally merges with *smart.* Then, another occurrence of *he* is selected and externally merges with *to be he smart.* Similarly, the third and fourth occurrences of *he* are externally merged in their respective positions. At this point, as in Kitahara's analysis, the *he* in the matrix Spec TP is deleted, and *he* gets interpreted. Therefore, even though there have been multiple instances of Select, there has been only a single interpretation of *he.* The connection between Select and interpretation falls apart.

Kitahara also suggests, "… if two occurrences of *he,* for example, are formed by distinct applications of Select accessing *he,* each occurrence of *he* necessarily bears a Case feature. This entails the presence of two distinct Case positions for convergence … [I]n each of these positions, *he* deletes its Case feature, and is interpreted." But the derivation given above clearly shows that this reasoning doesn't hold. Indeed, not only can multiple applications of Select result in multiple occurrences of *he* with unchecked Case feature through External Merge: multiple occurrences with unchecked Case features emerge equally well through *Internal* Merge. There is no inherent connection at all between the number of Select operations and the number of Case features that require deletion. Kitahara's proposal, based on Select, Case-checking and interpretation, thus fails to capture the difference between occurrences and repetitions.

4. Phase-Level Memory

Chomsky has suggested in several places that the issue of distinguishing copies and repetitions can be addressed by introducing the notion of phase level memory. Some quotes are given in (36-37):

(36) Chomsky 2008

"There must be some way to identify internally merged α with its copy, but not with other items that have the same feature composition: to distinguish, say, 'John killed John' or 'John sold John to John' (with syntactically unrelated occurrences of John), from 'John was killed John' (with two copies of the same LI John). That is straightforward, satisfying the inclusiveness condition, if within a phase each selection of an LI from the lexicon is a distinct item, so that all relevant identical items are copies. Nothing more than phase-level memory is required to identify these properties at the semantic interface C-I, where the information is required."

(37) Chomsky, Gallego and Ott 2017

"At TRANSFER, phase-level memory suffices to determine whether a given pair of identical terms Y, Y′ was formed by IM. If it was, then Y and Y′ are copies; if it was not (i.e., it was formed by EM), Y and Y′ are independent repetitions."

The two quotes both invoke phase level memory, but in different ways. The quote in (36) talks about selection of lexical items, and the quote in (37) talks about whether Internal Merge (IM) or External Merge (EM) takes place.

Let us consider the following structures for our original examples in (2) and (3). Under each are listed the kinds of phase level memory described by the above quotes:

(38) {John, {T, {be, {seen, John}}}}

Selection: *John* was selected from the lexicon only once, so the first and the second occurrences are copies.

EM vs. IM: the first occurrence of *John* was formed by IM, so the first and the second occurrences are copies.

(39) {John {v, {saw, John}}}
Selection: *John* was selected from the lexicon twice, so the first and second occurrences are distinct repetitions.
EM vs. IM: The first occurrence of *John* was formed by EM, so the two occurrences of *John* are distinct repetitions.

The first approach, (36), brings up many questions. In what way are occurrences of *John* marked as distinct in (39)? The quote suggests that this can be done by looking at the selections of lexical items from the lexicon. How are these lexical selections recorded and stored in phase level memory? Of course, one can infer the lexical selections that have taken place by looking at the sequence of workspaces. If the lexical item *John* is not in W_1 but it is in the following W_2, then *John* was selected from the lexicon. But, once one establishes the lexical selections, how is this information relayed to the interfaces? Furthermore, in the case of complex syntactic objects such as *the governor of California*, exactly how would the selections of the lexical items out of which it is built be associated with it?

The second approach, (37), brings up a similar set of issues. How is the fact that *John* was formed by IM stored in memory? Note that in minimalist syntax, there is no separate operation IM, there is just Merge that has two separate cases. So one needs a way to identify whether or not IM has applied to form (38) or (39). Again, the only way to do this is to look at a previous workspace, in this case the immediately preceding workspace (W_{n-1}), and compare it to the current workspace (W_n); there is no way that one can tell whether the two occurrences are related by IM simply by looking at the contents of W_n.

To order to investigate this issue, let's make the assumption (see Chomsky, Gallego, Ott 2017; Collins and Stabler 2016), that if X and Y are members of a workspace W, then Merge replaces X and Y with {X, Y} (removing X and Y from the workspace). Clearly in a derivation of (38) the structure {T, {be, {seen, John}}} will be a member of W_{n-1}. Now the question is whether *John* was also a member of W_{n-1}. If not, then we can conclude that only IM was involved, and hence the occurrences of *John* in (38) must be copies. But suppose that *John* was a member of W_{n-1}. Then the only way to know whether IM or EM was involved in forming (38) is by looking at W_n and seeing if it contains *John* as a member. If so, (38) must have been formed from IM (since the extra occurrence of *John* in the workspace is not used up).

These issues can be discussed in the context of a particular derivation of (38):

(40) a. W_0: {T, be, seen, John}
b. W_1: {T, be, {seen, John}}
c. W_2: {T, {be, {seen, John}}}
d. W_3: {{T, {be, {seen, John}}}}
e. W_4: {{John, {T, {be, {seen, John}}}}}

Transfer examines the syntactic object {John, {T, {be, {seen, John}}}} contained in W_4 and decides whether or not to spell-out the lower occurrence of *John*. Given phase level memory, Transfer makes reference to W_3, and sees that *John* is not a member of W_3. Therefore, the higher occurrence of *John* in {John, {T, {be, {seen, John}}}} in W4 must have been formed by IM.

Therefore, at Transfer, the lower occurrence of *John* is not spelled-out.

The above analyses show that it is *possible* to distinguish copies from repetitions by examining a series of workspaces; this series can embody "phase-level memory." But this raises a host of issues.

First, it means that the data structure sent to the semantic and phonological components is no longer a syntactic object proper, but a new object: a sequence of workspaces, an object not formed by Merge. Second, if post-syntactic computation has access to a sequence of workspaces, what are the interpretive constraints on such access? An unforeseeable range of interpretive possibilities seems to be unleashed depending on how access to previous workspaces is constrained. Finally, and most importantly, though it is perfectly possible to make the copy/repetition distinction by examination of the sequence of workspaces, the only motivation provided for such phase-level memory is precisely our need to make the distinction: yet this distinction does not *fall out* of phase-level memory in any natural way, but must be stipulated as a (complex) interpretive algorithm. Why are copies and repetitions defined as they are in the first place, given a sequence of workspaces? What motivates this particular distinction or any particular algorithm at all?

5. Conclusion

In this paper we have reviewed various proposals to distinguish copies and repetitions in the minimalist syntax literature. We have shown than none of these are satisfactory. Multi-dominance violates Inclusiveness. Chains require a new operation Form Chain. Phase level memory draws a distinction between copies and repetitions, at the cost of introducing phase level memory into the syntactic computation (with all the questions that such phase level memory raises).

The problem can be understood as an apparent limit on the power of Merge to create objects that naturally encode the distinction. The structure-building involved in Internal Merge is identical to that in External Merge – the creation of a new set {X, Y} out of pre-existing sets X and Y (Epstein et al 1998, Chomsky 2000, 2008) – leaving no trace of the origins of X or Y. The attractiveness of a theory that takes such a simple operation as Merge as its sole structure-building operation compels us to resist arbitrary complications, yet no approach so far has been able to make sense of the distinction between copies and repetitions without doing so.

Chomsky (2004: 110) sums up an important advance of recent years:

(41) "NS is based on the free operation Merge. SMT entails that Merge of α, β is unconstrained, therefore either external or internal. Under external Merge, α and β are separate objects; under internal Merge, one is part of the other, and Merge yields the property of "displacement," which is ubiquitous in language and must be captured in some manner in any theory. It is hard to think of a simpler approach than allowing internal Merge (a grammatical transformation), an operation that is freely available. Accordingly, displacement is not an "imperfection" of language; its absence would be an imperfection."

There is no question that the move to a Merge-based view of structure-building has provided tremendous insight into the otherwise mysterious phenomenon of displacement. But until the matter of distinguishing copies from repetitions has been settled, our understanding is far from complete. It is possible that the approaches we have examined might be improved in a way that solves the problem, or that some new solution will appear within the current theoretical assumptions; it is also possible that the problem will require some radical re-thinking.

References

Adger, David. 2010. A minimalist theory of feature structures. In Anna Kibort and Greville Corbett (eds.), *Features*. Oxford University Press.

Adger, David and Peter, Svenonius. 2011. Features in minimalist syntax. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*. Oxford University Press.

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

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. MIT Press.

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

Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In Adriana Belletti (ed.), *Structures and Beyond,* 104–131 (originally published as: Chomsky, Noam. 2001. Beyond Explanatory Adequacy: MIT Occasional Papers in Linguistics 20. MIT Working Papers in Linguistics.)

Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, 133-166. Cambridge, MA: MIT press.

Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130, 33-49

Chomsky, Noam, Angel Gallego, and Dennis Ott. 2017. Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. (https://ling.auf.net/lingbuzz/003507)

Citko, Barbara. 2011a. Multidominance. In Cedric Boeckx (ed.), *Handbook of Linguistic Minimalism,* 119-142. Oxford University Press, Oxford.

Citko, Barbara. 2011b. *Symmetry in Syntax: Merge, Move and Labels*. Cambridge University Press.

Collins, Chris. 2002. Eliminating Labels. In Samuel David Epstein and T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 42-64. Blackwell, Oxford.

Collins, Chris and Joachim Sabel. 2016. A C/I-Interface condition on In Remnant Movement. In Günther Grewendorf (ed.), *Remnant Movement (SGG 123)*, 93-131. De Gruyter.

Collins, Chris and Edward Stabler. 2016. A Formalization of Minmalist Syntax. *Syntax* 19, 43-78.

Epstein, Samuel D., Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations.* Oxford: Oxford University Press.

Epstein, Samuel D. and T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.

Gärtner, Hans Martin. 2002. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Akademie Verlag: Berlin.

Hornstein, Norbert. 1999. *Movement and Control.* Linguistic Inquiry, Volume 30, Number 1, Winter 1999. MIT Press. 69–96.

Kitahara, Hisatsugu. 2000. Two (or More) Syntactic Categories vs. Multiple Occurrences of One. *Syntax* 3, 151-158.

Martin, Roger and Juan Uriagereka. 2014. Chains in Minimalism. In Peter Kosta et. al. (eds.), *Minimalism and Beyond: Radicalizing the interface,* 169-194. John Benjamins Publishing Company.

Muñoz Pérez, Carlos. 2018. Recognizing Copies: On the Definition of Non-Distinctiveness. *Glossa* 3(1): 40, 1-29.
Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. MIT Press.
Seely, T. Daniel. 2006. Merge, Derivational C-Command, and Subcategorization in a Label-Free Syntax. In Cedric Boeckx (ed.), *Minimalist Essays*. Amsterdam: John Benjamins.
Stabler, Edward. 1997. Derivational Minimalism. In Christian Retoré (ed.), *Logical Aspects of Computational Linguistics*, 68-95. Springer.