John Benjamins Publishing Company

This is a contribution from *Minimalism and Beyond. Radicalizing the interfaces*.
Edited by Peter Kosta, Steven L. Franks, Teodora Radeva-Bork and Lilia Schürcks.
© 2014. John Benjamins Publishing Company

This electronic file may not be altered in any way.
The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.
Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).
Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Chains in Minimalism*

Roger Martin and Juan Uriagereka
Yokohama National University / University of Maryland

This paper considers how the system identifies multiple occurrences of a syntactic object α as a chain, a set of copies. For Chomsky (1995, 2000, 2001), copies can arise only by movement (internal merge); lexical items introduced by external merge are stipulated to be distinct tokens, coded by indexation, or by introducing some other special concept of numeration, which violates the *inclusiveness condition*. We argue that this type/token dichotomy is unnecessary and that copies can be distinguished from repetitions in terms of syntactic context alone. This yields interesting situations where two arguments introduced by external merge may also be recognized as a chain, and we propose that obligatory control and parasitic gaps should be analyzed in exactly these terms.

1. Introduction

Our concern in this paper is the notion of chain. In Section 2 we review how chains have been conceived within the minimalist program, most notably by Chomsky (1995, 2000, 2001, 2008). We argue that these conceptions rely on making an ontological distinction between types and tokens, invoking indices or some similar coding device, which violates basic principles of conservation. In Section 3 we present an alternative view, in which the difference between "copy" and "repetition" is determined contextually. We then extend the program to account for obligatory control and parasitic gaps in Section 4. Section 5 concludes the paper.

2. Some minimalist concepts of chain

Chomsky (1995) defines a (non-trivial) chain as a pair of positions (or *occurrences*; Chomsky 2000), (K, L) of a constituent α in a phrase marker Σ, where a position is taken to be a co-constituent ("sister") of α in Σ. Thus a chain is a set of phrase markers denoting the positions in which some element α occurs, in some sense, simultaneously. For example, in (1a), assuming the copy theory of movement, the element *students*

© 2014. John Benjamins Publishing Company
All rights reserved

is moved from its position of initial merger to create the specifier of TP, as shown in (1b). We thus have a chain consisting of two occurrences of *students* as shown in (1c).[1]

(1) a. Students were criticized.
b. [$_{CP}$ [$_{TP}$ students [T [$_{PartP}$ were [$_{VP}$ criticized students]]]]]
c. CH = <{were, {criticized, students}}, criticized>

As Chomsky points out, however, to ensure proper interpretation at the interface, we need to be able to distinguish a situation like (1) from that in (2). The sentence in (2a) has the structure in (2b), with four occurrences of *students* constituting the two chains in (2c).

(2) a. Students believe students were criticized.
b. [$_{CP}$ [$_{TP}$ students [T [$_{vP}$ students [*v* [$_{VP}$ believe [$_{CP}$ that [$_{TP}$ students [T [$_{PartP}$ were [$_{VP}$ criticized students]]]]]]]]]]
c. CH_1 = <{were, {criticized, students}}, criticized>
CH_2 = <{T, {students, {*v*, {believe, {that {students, {T, {were, {criticized, students}}}}}}}}}, {*v*, {believe, {that {students, {T, {were, {criticized, students}}}}}}}>

Chomsky argues that chains, defined as multiple occurrences of a single α, can arise only by way of movement given that "we distinguish among distinct selections of a single lexical item…" (1995: 251). In particular, assuming that at some stage S in a derivation D, there exists a set of syntactic objects {SO_1, …., SO_n}, the operation Select is "a procedure that selects a lexical item LI from the numeration, reducing its index by 1, and introduces it into the derivation as SO_{n+1}" (1995: 226), namely, introduces it as an *indexed* syntactic object.[2]

It is obvious that two different notions of index are being referred to here, as can be seen more clearly in the following passage:

> Suppose the lexical item LI in the numeration N has index *i*. If a derivation is to be generated, Select must access LI *i* times, introducing it into the derivation. But the syntactic objects formed by distinct applications of Select to LI must be distinguished; two occurrences of the pronoun *he*, for example, may have entirely different properties at LF [that is, two repetitions as opposed to copies of *he*, as in *he thinks he is clever*; M&U]. *l* and *l′* are thus marked as distinct for C_{HL} if they are formed by distinct applications of Select accessing the same lexical item of N. Note that this is a departure from the inclusiveness condition, but one that seems indispensible… (Chomsky 1995: 227)

One type of index exists solely in the numeration. In fact, it can be said that a numeration differs from a lexical array precisely in that it specifies not only which lexical items are accessed in the derivation, but also exactly how many times each lexical item is selected (hence the index in question gets "reduced by 1" with each application of Select). On the other hand, a lexical array can be thought of as a mere subset of the

© 2014. John Benjamins Publishing Company
All rights reserved

lexicon, only specifying the first of these two conditions. The second notion of index seen here is that which gets assigned under Select, marking any two selections (even two selections of the same lexical item) as distinct syntactic objects in the derivation.[3]

Chomsky (2000: 114), reflecting on his earlier analysis, takes a revisionist stance on the issue: "In [Chomsky 1995] indices are introduced, but not within C_{HL} itself — rather in extending lexical arrays to numerations." Yet, as we have just seen, indices *are* introduced within C_{HL} in Chomsky 1995, as part of the operation Select. Here, however, Chomsky claims that indices are only introduced in the numeration and furthermore seems to suggest that this suffices not only to specify the number of times a given lexical item is selected in a derivation, but also to recognize each of those selections as distinct syntactic objects.

It is not clear how Chomsky intends to achieve this. One implementation would be to take the numeration to be a set of indexed syntactic objects from the start. Thus in the earlier sense a numeration might look like {(students, 2), …}, according to which Select must apply to the lexical item *students* twice, adding two distinctly indexed members to the set of syntactic objects in the derivation, {$student_1$, $student_2$,…}. Whereas an alternative conception might assume the numeration to instead look something like {$student_1$, $student_2$, …}. Since this set corresponds directly to the previously assumed set of syntactic objects in the derivational workspace, the operation Select (and its associated indexing mechanism) is not a necessary part of the computational mapping from numeration to the interfaces.[4]

Chomsky's (2001) discussion, however, suggests that he has a rather different idea in mind:

> The new concept [numeration] is added, while another concept is eliminated: chains are determined by identity, *with no need for indices or some similar device to distinguish chains from repetitions*… (Chomsky 2001: 11; emphasis added)

Here Chomsky claims that indices play no role in distinguishing different selections of the same lexical item (repetitions) from copies formed in the derivation. However, for *the numeration* to establish this without indices, it seems to us that it would have to be significantly more articulated than just a set of elements from the lexicon, taking on properties reminiscent of those typically ascribed to D-structure. Furthermore, any distinctions made in the numeration would have to be retrievable at the interfaces, presumably involving some sort of derivational memory.[5]

We will not dwell on the differences between these takes on the numeration. However the relevant properties are established — by introducing indices or by enriching the numeration in some other way — inasmuch as these properties do not exist in the lexicon, such a move still seems to constitute a violation of inclusiveness.

The essence of the problem is the type/token distinction. If the lexicon specifies only types, there is no natural way to capture tokens in the system. Suppose, on the other hand, that a derivation only deals with lexical types. The input to such a computational procedure can easily be thought of as a mere subset of the lexicon (more

© 2014. John Benjamins Publishing Company
All rights reserved

precisely, a *lexical array*, as opposed to numeration, in Chomsky's terminology).[6] In fact, Chomsky (2000: 114), aware of the issues at hand, invites us to consider this sort of possibility, noting that the only significant consequence of such a move is that "it would require a new notion of chain." He continues: "if [simple unarticulated lexical arrays, taken to be mere subsets of the lexicon] are extended to numerations Num, a chain can be defined as a sequence of identical α's — more accurately, a sequence of *occurrences* of a single α" (Chomsky 2000: 114). In the next section, we will pursue Chomsky's suggestion as we outline an approach "to distinguish chains from repetitions" that does not rely on indices or force us to complicate the input to derivations beyond an unarticulated subset of the lexicon.[7]

3. Towards a new understanding of chains in minimalism

Suppose that the derivation starts not with a numeration but with a lexical array, a set of lexical items that have been activated for use in a given derivation. A lexical array, in this sense, is still very much part of the fabric of the lexicon; it has not been turned it into some entirely new entity. Yet if all one has is types, what does it mean to access them *more than once* in a derivation? In particular, why are multiple syntactic objects that are identical from a lexical perspective sometimes regarded as one object, whereas sometimes they are regarded as multiple objects, at least in their interpretation? In other words, what does it mean to have copies as opposed to repetitions?

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.[8] There is no real difference from this perspective between repetitions and copies. A repetition is simply an occurrence of some syntactic object that does not manage to collapse together with some other occurrence of the same syntactic object when interpreted at the interfaces. Thus, ontologically, there are only two sorts of things: items in the lexicon and interactions among lexical items in a given derivation, as predicted under inclusiveness.

The following analogy from visual perception might help understand the intuition. Suppose that, at intervals of half a second, you are shown a sequence of images: one of a man, then one of a bird, then the same image of the man, then one of a butterfly, and finally the man again. Under those circumstances you will see five distinct images in rapid succession, three of which appear to be repetitions of the same image. However, if the bird is put at the beginning, and the butterfly at the end, so that the three identical images of the man become adjacent to one another, what one perceives is a total of three distinct images — the three images of the man collapse into what is interpreted as a single image, even if the presentation of this image may appear to be longer in duration than that of the other two images. Also, even with this ordering (the three identical images adjacent), if the five images are presented with "gaps" of a

© 2014. John Benjamins Publishing Company
All rights reserved

certain degree of duration (say, one second) between each successive image, one will again perceive each of the three repetitions of the image of the man as a distinct image.[9]

3.1 An analysis of A-chains

With the above analogy in mind, consider once again the example in (2), repeated in (3) (here and below indices are given for clarity and are not intended to be part of the representations).

(3) 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$_1$]]]]]]]]]]

We want to ensure that *students*$_1$ and *students*$_2$ are treated as copies, as are *students*$_3$ and *students*$_4$. Furthermore, no member of the latter pair of occurrences is a copy of either member of the former, but rather must be seen as repetitions. One way to understand this is in terms of the locality of phases. Assume the phase impenetrability condition (PIC) proposed by Chomsky (2000: 108) in (4).

(4) *Phase-Impenetrability Condition*
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.

Notably, for each of the pairs *students*$_1$/*students*$_2$ and *students*$_3$/*students*$_4$, the lower occurrence is accessible to the higher one, according to (4). On the other hand, if we consider the pair *students*$_2$ and *students*$_3$, the lower occurrence is not accessible to the higher one. Thus we might assume that if X and Y are identical syntactic objects, and X is accessible to Y, then X and Y are interpreted as a chain (i.e., as copies not repetitions). On this view, when two elements can be interpreted as a chain, they must be. Notice that this is in sync with the analogy from visual perception above; when the two adjacent images are close enough together (in terms of the time lag in between presentations), there is no choice as to whether to interpret them as one image or two — what you see is what you get.

Consider next the structure in (5a), which can only be interpreted as (5b), involving two chains headed by two repetitions of *students*, and not as (5c) with one chain composed of three copies of *students*.

(5) a. [$_{CP}$ [$_{TP}$ students$_3$ [T [$_{vP}$ students$_2$ [v [$_{VP}$ criticized students$_1$]]]]]]
b. Students criticized students.
c. *Students criticized.

In (5a), *students*$_2$ and *students*$_3$ will be interpreted as a chain since the former is accessible to the latter. However, it would seem that *students*$_1$ and *students*$_2$ should also

© 2014. John Benjamins Publishing Company
All rights reserved

be treated as copies given what we have said so far, as both of these occurrences are contained in the same phase.[10] Thus it looks like we need a somewhat stricter notion of locality than (4) provides, at least for the purpose of copy/chain determination. Needless to say, the desire is for the relevant locality, whatever it may turn out to be, to follow from the natural dynamics of the system, and not to have to be stipulated as a principle of chain formation.

Suppose the domain of a phase head H is transferred to the interfaces at the completion of the phase. This would yield the effects of (4), if we make the further assumption that material contained in a transferred domain is not accessible to further syntactic operations. If so, it would be natural to assume that chains are "cashed out" in a similar manner, with all identical syntactic objects contained in the domain of a transferred phase head being interpreted as copies.[11] For example, at the completion of the *v*P phase in (5a), the single occurrence $students_1$ in the domain of *v* would be transferred to the interface as a chain CH_1.[12] Then at the completion of the CP phase, $students_3$ and $students_2$ in the domain of C are transferred as a distinct chain, CH_2 ($CH_1 = students_1$ is no longer accessible at this point, having already been transferred).

3.2 Some issues concerning A-bar chains

The sort of analysis sketched above works well for distinguishing copies from repetitions in A-chains. However, we must also take into account A-bar chains, which raise further issues. Consider the indirect question in (6a), along with the relevant part of its structure in (6b).

(6) a. Guess [$_{CP}$ who students criticized]
b. [$_{CP}$ who$_3$ [$_{TP}$ students [T [$_{vP}$ who$_2$ [students [*v* [$_{VP}$ criticized who$_1$]]]]]]]

The structure in (6b) adheres to the standard assumption that *wh*-movement takes place successive cyclically through the edge of each phase, hence who_2 in the outer-spec of *v*P. Yet if we follow the above logic, who_1 and who_2 should not be analyzed as forming a chain: who_1 is transferred together with the VP at the completion of the *v*P phase, and therefore it will be recognized as distinct from any occurrences occurring outside of this domain due to inaccessibility. Thus there must be some way for an occurrence to remain accessible in the derivation even after transfer applies, allowing the chain to extend its reach into higher phases. At least we want this to be possible in examples like (6), but not in (5).

It is conceivable that who_1 in (6b) is indeed interpreted as a unique (singleton) A-chain at the *v*P phase, hence who_1 and who_2 do not constitute (part of) a chain, just as expected on the current view. That is, we might assume, adopting the analysis of Chomsky (2008), that who_2 is the lowest occurrence in a uniform A-bar chain, and that this and the distinct A-chain who_1 are construed as an "operator-variable construction" at the interface.[13]

© 2014. John Benjamins Publishing Company
All rights reserved

However, even if this two-chain analysis of operator-variable constructions is correct, we still need to say something more in order to account for A-bar chains. The reason is that who_2 and who_3 in (6b) will also not be recognized as copies, since who_2 but not who_3 will be transferred as part of the domain of C at the completion of the CP phase.

We do not see many ways to provide a satisfactory treatment of A-bar chains. The one that seems most promising is based on the assumption that internal merge automatically yields a chain. In other words, if two occurrences arise by way of internal merge, they must be interpreted as copies at the interface. The question is how to instantiate such an idea.

Recall, that, on our view, whether an occurrence is a repetition or a copy of some other occurrence is a matter of interpretation at the interface and depends solely on the syntactic context. Suppose that whereas external merge creates a new term, or syntactic object, in a phrase marker, internal merge works differently in that it, in some sense, "stretches" the very same syntactic object across two different syntactic contexts.

Another analogy might help to understand the intuition we are entertaining. Imagine two sheets of origami paper, one black and the other grey, combined by stapling the bottom edge of the grey sheet to the top edge of the black sheet. We could continue this "derivation" by taking a new sheet of grey paper and stapling it to black sheet, so that now we have a total three sheets combined into a single object with the bottom-to-top order grey-black-grey (see Figure 1). However, suppose that rather than introducing a new grey sheet, we instead fold the sheets from the initial step in such a way that we staple the bottom edge of the black sheet to the top edge of the grey sheet (see Figure 2).

In this case, the resulting object consists of just two sheets of paper, yet it is topologically more complex than the previous object with three sheets. Needless to say, there is something very similar about the nature of these "origami operations," but at the same time we have created very different sorts of objects in the two cases.

We speculate that external merge makes a phrase marker larger, whereas internal merge turns it into a topologically more complex sort of object. This raises a vast array of questions — including the very nature of phrase structure and how it is to be formalized — that fall beyond the scope of the current paper. The point that we would like to make here is that, if this is the right way of thinking, the question of repetition vs. copy does not even arise in the case of internal merge, as there is only one occurrence to speak of, albeit one that exists simultaneously in multiple syntactic contexts. Also note that, by distinguishing two types of merge in this way, we have not reintroduced inclusiveness-violating "tokens" into the system. The relevant notions are definable in terms of lexical items and their combinatorial interactions alone.

Readers may worry about the special nature of *internal* merge, as opposed to *external* merge, given Chomsky's (2001, 2008) claim that there is no substantive difference between the two varieties. However, while it may or may not be true that there is a shared component of these two operations, it is very well established that internal

© 2014. John Benjamins Publishing Company
All rights reserved

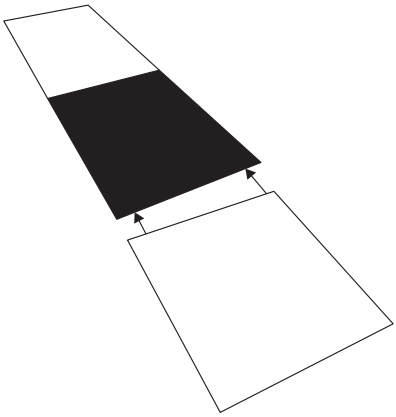

Figure 1.

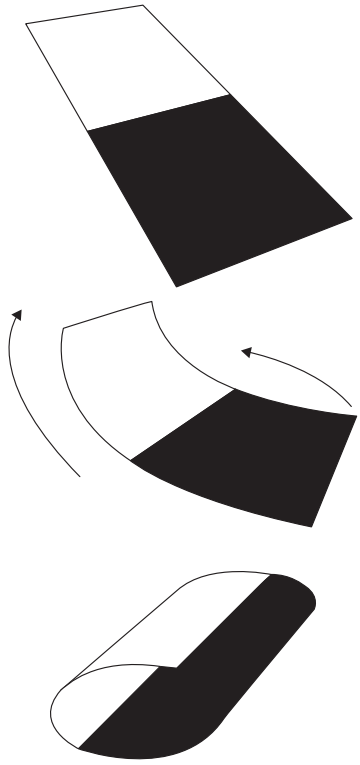

Figure 2.

merge (i.e., displacement) at least differs from external merge in that only the former involves a context-sensitive component — however it is to be instantiated (e.g., by way of *Agree*, as in the probe-goal system in Chomsky 2000, 2001, 2008).[14]

Another assumption that we will make in order to deal with A-bar chains concerns situations in which, descriptively speaking, only part (i.e., the tail and/or some intermediate link) of the chain (call this Ch_{sub}) is contained in the domain of phase P. In such cases, we assume that Ch_{sub} is not literally "handed over" to the interfaces when the domain is transferred at the completion of P. This does not entail, however, that Ch_{sub} is still accessible to further computation after transfer applies. Rather, we continue to assume that *only the copy in the edge of the phase can be accessed*.[15]

In this light, we should note that who_2 in a structure like (6b) may in some instances still be *active* at the point when the domain TP undergoes transfer.[16] In particular, it is not always the case that the unvalued feature that makes movement to the edge of an intermediate phase legitimate in the first place becomes valued as a result of that movement; otherwise further movement would not be possible. If who_2 is literally sent to the interfaces when TP is transferred, it should result in the derivation crashing. Thus an active chain link contained in a transferred domain, although presumably no longer accessible to syntactic operations, must somehow be able to remain in the

© 2014. John Benjamins Publishing Company
All rights reserved

derivation until feature valuation takes place (rendering all links inactive) at which point the chain *as a whole* is ready to be sent to the interfaces.[17]

We should also consider examples like (7a) and (7b), which are shown together with two possible structures (of the embedded clause) in (7c) and (7d).

(7) a. *They know [$_{CP}$ who who criticized]
b. They know [$_{CP}$ who criticized who]
c. [$_{CP}$ who$_5$ [$_{TP}$ who$_4$ [T [$_{vP}$ who$_3$ [who$_2$ [v [$_{VP}$ criticized who$_1$]]]]]]]
d. [$_{CP}$ who$_4$ [$_{TP}$ who$_3$ [T [$_{vP}$ who$_2$ [v [$_{VP}$ criticized who$_1$]]]]]]

Following Chomsky (2008), we assume that the structure of the embedded clause in (7b) must be as in (7d), not as in (7c). The reason for this is that the external argument *who*$_2$ is not in the search domain of probe *v* and therefore cannot be the target of internal merge forming a copy in the outer-spec of *v*P. In (7a), on the other hand, the direct object *who*$_1$ can presumably serve as a goal for probe *v*, assuming it has an appropriate unvalued feature making it active. If so, internal merge should be able to apply to yield *who*$_3$ in (7c).[18]

As Chomsky (2008) notes, this sort of analysis leaves the superiority effect witnessed in (7a) unexplained. In fact, the same sort of derivation is presumably necessary in (8), involving topicalization, which does not exhibit a superiority effect.

(8) a. Students, students criticized.
b. [$_{CP}$ students$_5$ [$_{TP}$ students$_4$ [T [$_{vP}$ students$_3$ [students$_2$ [v [$_{VP}$ criticized students$_1$]]]]]]]

We follow Chomsky (2008) in setting aside the question of what is behind the superiority effect witnessed in (7a), and focus here on the question of why *students*$_2$ and *students*$_3$ in (8b) (and presumably *who*$_2$ and *who*$_3$ in (7c) as well) are necessarily treated by the system as repetitions and not copies.

As it turns out, the correct distinctions can easily be made if we adopt Chomsky's (2008) proposal that chains are always uniform. More precisely, the assumption is that since *students*$_3$/*who*$_3$ is an A-bar position and *students*$_2$/*who*$_2$ is an A-position, there cannot be a chain involving these occurrences. Furthermore, if *students*$_3$/*who*$_3$ has an unvalued feature — namely, whatever feature makes displacement to the edge of the phase possible — whereas *students*$_2$/*who*$_2$ does not, the relevant occurrences would also be distinct in terms of their feature makeup, which might also prevent them from being interpreted as copies.[19]

3.3 Summary

Let us briefly summarize our main proposals up to now. We have suggested that the input to the derivation is nothing more than a set of lexical types and, furthermore, that there are no inclusiveness violating operations that transform a lexical type into a

© 2014. John Benjamins Publishing Company
All rights reserved

syntactic token. We proposed that two identical syntactic objects are interpreted, necessarily, as copies if they are very close to one another, where we stated the relevant locality in terms of phase theory. The basic idea behind this locality is that two occurrences contained in the same transferred domain are treated as copies. However, to account for properties of A-bar chains, we concluded that internal merge in some sense creates chains/copies immediately, perhaps because it does not introduce a new syntactic object into the derivation, but rather stretches an existing one across two contexts. If we think of internal merge in this way, it is certainly tempting to simply stipulate that external merge creates repetitions so that the copy/repetition distinction follows directly from the internal/external merge distinction, by definition. However, in the next section, we show how that, by not making such a stipulation, and allowing for situations where two occurrences introduced by external merge may be recognized as copies, we are able to provide an analysis of chain-like phenomena that have resisted a satisfactory treatment in terms of movement (internal merge): obligatory control and parasitic gaps.

4. Some extensions of the program

In this section, we will consider some consequences of the above approach to chains. First, in Section 4.1, we briefly review the analysis of obligatory control by Martin and Uriagereka (2013), which follows naturally from the proposals made above. While this analysis maintains many of the virtues of the so-called movement theory of control (Hornstein 1999), it differs from such theories in that it is compatible with a strictly configurational theory of θ-roles. In Section 4.2, we extend our analysis of OC to account for parasitic gaps, in a way that does not require us to complicate the theory of merge/search by allowing sideward movement.

4.1 Obligatory control

We agree with Hornstein (1999) that obligatory control (OC) should not be analyzed as involving an empty formative PRO or a special (null) Case stipulated to explain its distribution (Chomsky and Lasnik 1993; Martin 1996, 2001).[20] At the same time, we do not subscribe to the idea that the controller-controlee relation involves movement (internal merge) to a θ-position, primarily because, in our view, θ-roles are best treated as the interpretations assigned by the interface to configurations — namely, the structures resulting from the application of context-free external merge — not as features entering into the probe-goal system. We take the latter to be the only source for internal merge, assuming some version of last resort (for extensive arguments in favor of a configurational approach to θ-roles see, for example, Chomsky 1995, Hale and Keyser 1993, Uriagereka 2008). Furthermore, even if one were to assume some non-θ-type feature to serve as the driving force for internal merge to a θ-position (such as c-selection features in Bowers 2008), this would still be incompatible with an approach

© 2014. John Benjamins Publishing Company
All rights reserved

that takes θ-roles to be configurations: a chain is simply not a configuration, as the notion is typically understood.[21]

Martin and Uriagereka (2013) analyze OC in terms of chains and, at the same time, are able to maintain that θ-roles are configurations since the relevant chains do not involve movement to a θ-position. The approach turns out to be reminiscent of earlier *Equi* deletion approaches to OC (e.g., Rosenbaum 1967), which assume the controller-controlee relation to involve two distinct occurrences of identical NPs, one of which gets deleted in the course of the derivation.

Let's begin by considering OC with object control, illustrated by (9a), which we assume has a structure like (9b).[22]

(9) a. John persuaded Mary to kiss him.
b. [$_{CP}$ [$_{TP}$ John$_2$ [T [$_{vP}$ John$_1$ [v [$_{VP}$ Mary$_2$ [persuaded [$_{TP}$ to [$_{vP}$ Mary$_1$ v [$_{VP}$ kiss him]]]]]]]]]]

Crucially, we assume that both *Mary*$_1$ and *Mary*$_2$ are necessarily introduced into the phrase marker by external merge in order for θ-role assignment to take place, as discussed above. Another crucial assumption about the structure of (9a) in (9b) is that the infinitival complement is not CP, but rather TP.[23] If so, when the matrix VP undergoes transfer at the completion of the matrix *v*P phase, *Mary*$_1$ is still accessible, thus *Mary*$_1$ and *Mary*$_2$ will be recognized by the system as copies, not repetitions.

One of the notable properties of OC is that only the structurally closest NP can serve as the controller — the so-called Minimal Distance Principle (Rosenbaum 1967), which we take to be descriptively correct. In our terms, in order for (9a) to be interpreted as subject control, it would need to have a structure like (10), where crucially *John*$_1$ and *John*$_2$, both necessarily introduced by external merge, are recognized as copies.

(10) [$_{CP}$ [$_{TP}$ John$_3$ [T [$_{vP}$ John$_2$ [v [$_{VP}$ Mary$_1$ [persuaded [$_{TP}$ to [$_{vP}$ John$_1$ v [$_{VP}$ kiss him]]]]]]]]]]

However, this is not possible since, as we have just seen, the VP complement of the matrix *v* will be transferred at the completion of this phase. Since this domain contains *John*$_1$ but not *John*$_2$, these two occurrences will not be recognized as a chain. Since *John*$_1$ still has an unvalued Case feature, and is not part of a chain containing some link with a valued Case feature, the derivation will crash.[24] Thus the Minimal Distance Principle follows directly.

Before turning to subject control, let us pause for a moment to address some questions that arise concerning the above analysis, having to do with how the chain consisting of two occurrences of *Mary* in (9b) is interpreted at the interfaces. Interpretation at the sensory-motor (SM) interface is straightforward. The chain is interpreted in exactly one, usually the highest or most inclusive, position, as is almost invariably the case at this particular interface.[25] Interpretation at the CI interface raises more interesting questions. If the whole chain is available for interpretation at the CI interface, it

© 2014. John Benjamins Publishing Company
All rights reserved

might seem that there should be no problem computing the appropriate meaning of the sentence. However, such a ploy will not work.

As has already been discussed above, we are assuming a strictly configurational theory of θ-role assignment, and for some argument α to be in *a configuration* means for it not to be a (non-trivial) chain. In other words, if α is in a superposition of mutually exclusive states — namely, a chain state (the state of being in multiple configurations simultaneously) as opposed to being in a unique determinate configuration — we assume it cannot be assigned a θ-role. Interpreting α, we argue, entails establishing its exact position, which requires the chain to collapse.[26] However, if what we have just articulated is in on the right track, it seems we have reached a paradox. The relevant α in (9a) must be interpreted in two distinct θ-positions, and furthermore there is presumably a (unique) position from which α takes scope, which in principle may be different from both of the θ-positions. But if the chain must collapse to interpret α in any one of these positions, how can it be interpreted in all of them?[27]

Uriagereka (2008) argues, in part for these reasons but also on other grounds, for a distributed approach to meaning, involving independent 'conceptual' (C) and 'intentional' (I) interfaces. Following this line, we might assume the C system, which deals with θ-roles and other thematic-conceptual aspects of meaning, considers only configurations of the type {α, β} resulting from external Merge. On this view, collapse is not necessary in the C system, since it does not even recognize chains. On the other hand, suppose the I system, which deals with scope, operator-variable binding, reference relations, and other aspects of intentional meaning takes a more global view, in that it sees a transferred domain in its entirety, thereby recognizing a set of occurrences as either copies or repetitions in the ways described above. In order to yield the command relations that determine intentionality, chains must collapse into unique positions.[28]

Although we set aside many details concerning the exact architecture and derivational dynamics (see Uriagereka 2008), let's briefly consider the situation that arises in (9b) in these terms. We assume that when the C system examines the local configurations resulting from external merge {α, β}, which it interprets in a strictly compositional manner, it encounters two distinct θ-configurations, {Mary, K} and {Mary, L}. Each of these configurations involves an occurrence of *Mary*, both of which are treated as independent arguments and assigned independent (and possibly differing) θ-roles. When the I system, however, examines the transferred domain in this instance, its sees two occurrences of *Mary*, which it recognizes as copies, as discussed above. This chain then collapses to yield a unique scope-taking position for intentional interpretation. Dividing the labor of semantic interpretation between C and I in this way allows us to interpret two occurrences as distinct objects for the purposes of θ-role assignment, but as the same object with it comes to reference/scope. As noted above, this is conceptually very similar to a deletion-based analysis that postulates two NPs at D-structure, one of which gets deleted prior to LF.

Let's now turn to OC involving control by a subject, as in (11a), which we assume has the structure in (11b).

© 2014. John Benjamins Publishing Company
All rights reserved

(11) a. John tried to kiss Mary.
b. [$_{CP}$ C [$_{TP}$ John$_3$ [T [$_{vP}$ John$_2$ [v [$_{VP}$ tried [$_{TP}$ to [$_{vP}$ John$_1$ [v [$_{VP}$ kiss Mary]]]]]]]]]]

In order for (11b) to give rise to (11a), all three occurrences of *John* must be recognized as copies. However, if the matrix *v*P is a phase, the VP complement of the phase head *v* will undergo transfer at the completion of the phase, which results in *John*$_1$ being interpreted as distinct from *John*$_2$ and *John*$_3$. Since the uninterpretable Case feature of *John*$_1$ is not valued, this results in a derivational crash. The situation appears to be exactly the same as what we described for (10) above, and also what we presumably would want to say about examples like (12), with *believe* in place of *try*, which can come out only as (13a), not as (13b).[29]

(12) [$_{CP}$ C [$_{TP}$ John$_3$ [T [$_{vP}$ John$_2$ [v [$_{VP}$ believes [$_{TP}$ to [$_{vP}$ John$_1$ [v [$_{VP}$ kiss Mary]]]]]]]]]]

(13) a. John believes John to like Mary.
b. *John believes to like Mary.

For our analysis to work, it must be the case that the matrix *v*P in (9)/(10) and (12) is a phase, whereas in (11) it is not. Even assuming that phase-hood is somehow related to the presence of φ/Case on the phase head (Chomsky 2008), which seems to be at least descriptively correct, we cannot simply say that *try* differs from *persuade*/*believe* in that it lacks φ/Case, since it can take a nominal object, as in (14).

(14) John tried something.

Rather, it looks like we need to say that in the construction […v [$_{VP}$ try α]], v has φ/Case *when α is a nominal phrase but not when α is an infinitival complement*. In other words, whether or not *v* has φ/Case features depends in part on the complement structure of the VP.[30]

As surprising as this conclusion may sound, there is morphological evidence in Basque showing that something along these lines is necessary. Consider the following data from San Martin and Uriagereka (2002).

(15) a. Nik hori saihiatu dut
I$_{ERG}$ that$_{ABS}$ try$_{PRT}$ have$_{1PER/3PER}$
'I have tried that.'
b. Ni [pisua galtzen] sahiatu naiz
I$_{ABS}$ weight lose$_{INF}$ try$_{PRT}$ be$_{1PER}$
'I have tried to lose weight.'

In (15a), with a nominal complement, we see the typical pattern exhibited by transitive verbs in Basque, where the subject is marked with ergative Case and the object with absolutive Case. Furthermore, the form of the auxiliary is *have*, which shows

© 2014. John Benjamins Publishing Company
All rights reserved

both subject and object agreement. On the other hand, the pattern witnessed in (15b), where the complement is an infinitival clausal, is that typically witnessed with unaccusative verbs in Basque. The argument in subject position is marked with absolutive Case, not ergative, and the form of the finite auxiliary is *be*, which agrees only with the surface subject. Although we will not be able to go into an explanation of these facts here (see San Martin and Uriagereka 2002), we wish merely to note that they are predicted by the conclusion reached above.

Finally, consider the impossibility of lexical subjects in OC complements, as illustrated in (16).

(16) a. *John tried Bill/him/himself to kiss Mary.
b. $[_{CP}$ C $[_{TP}$ John$_2$ [T $[_{vP}$ John$_1$ [v $[_{VP}$ tried $[_{TP}$ to $[_{vP}$ Bill$_1$/him$_1$/himself$_1$ [v $[_{VP}$ kiss Mary]]]]]]]]]

The ungrammaticality of (16a) is straightforward on our analysis. The embedded subject cannot form a chain with an element in the higher clause (unlike (11), where it must do so) since the relevant occurrences are not identical. Therefore, the uninterpretable Case feature of the embedded lexical subject will remain unvalued and the derivation will crash.

In this subsection we have shown that the analysis of chains presented in Section 3 allows for a treatment of OC in terms of chains that does not assume PRO. Crucial to this analysis is the fact that two occurrences that have been introduced into a phrase marker by external merge can in principle be interpreted as copies in certain situations. Our analysis thus treats OC as involving a chain, but at the same time remains compatible with a strictly configurational treatment of θ-relations. In the following subsection we will further extend this approach to account for parasitic gaps.

4.2 Parasitic gaps

Within the minimalist program, parasitic gap constructions have been analyzed in terms of chains formed by movement, most notably by Nunes (2004). However, this type of approach relies crucially on the availability of sideward movement, which involves a serious complication of the notion minimal search so as to allow a probe to find a goal in a disconnected phrase marker.[31] In this section, we outline how the analysis of chains presented above can account for parasitic gaps in a way that does not depend on sideward movement.

Parasitic gap constructions in (17) exhibit a number of intriguing properties.

(17) (Guess) $[_{CP}$ who [you met *t* [before students criticized *e*]]]

For example, a parasitic gap cannot be c-commanded by the real gap (Engdahl 1984, Chomsky 1982), as shown by ungrammaticality of (18).

(18) *(Guess) $[_{CP}$ who [*t* met you [before students criticized *e*]]]

© 2014. John Benjamins Publishing Company
All rights reserved

We will focus on explaining this so-called anti-c-command requirement, namely, the contrast between (17) and (18).

First, consider the derivation of (17). Suppose the derivation constructs the adjunct phrase as in (19a), in essentially the same way as described for (6) above, alongside the independently constructed (19b). Assume that (19a) and (19b) can now be merged, to yield the structure in (19c).

(19) a. $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students [v $[_{VP}$ criticized who$_1$]]]]]]]]
b. $[_{vP}$ you [v $[_{VP}$ met who$_4$]]]
c. $[_{vP}$ [you [v $[_{VP}$ met who$_4$]]] $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students [v $[_{VP}$ criticized who$_1$]]]]]]]]]

The derivation continues as in (20), with *who*$_5$ being created in the spec of *v*P.[32]

(20) $[_{vP}$ who$_5$ [[you [v $[_{VP}$ met who$_4$]]] $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students [v $[_{VP}$ criticized who$_1$]]]]]]]]]]

Transfer then applies to the matrix VP, resulting in *who*$_4$ being interpreted as an A-chain.

Continuing further with the derivation, we reach the matrix CP phase, at which point we assume C enters a probe-goal relation with *who*$_5$, creating *who*$_6$ in spec-CP, as shown in (21).

(21) $[_{CP}$ who$_6$ [C $[_{TP}$ you [T $[_{vP}$ who$_5$ [[you [v $[_{VP}$ met who$_4$]]] $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students [v $[_{VP}$ criticized who$_1$]]]]]]]]]]]]]]

At the completion of this phase, the domain TP is transferred to the relevant interfaces. Crucially, we assume that *who*$_{2/3}$ is treated as a copy of *who*$_{5/6}$ (here and below we employ the informal notation $\alpha_{i/j}$ to indicate a situation where a single syntactic object α is in multiple configurations due to internal merge) since they are both contained in the domain and accessible to one another. As a consequence, *who*$_{2/3}$ does not have to be literally handed over to the interface at this step, *being recognized as part of a chain whose reach extends outside of the domain*. In other words, at the point of transfer *who*$_{5/6}$ is simultaneously inside and outside of the domain, hence cannot be sent to the relevant interfaces. At the same time, *who*$_{2/3}$ is contextually determined to be a copy of *who*$_{5/6}$, therefore it is able to remain in the derivation as well. Finally, the root CP is transferred, at which point the A-bar chain (*who*$_{5/6}$, *who*$_{2/3}$) is handed over to the interfaces, where it collapses in the highest position, in which an operator can be interpreted. Assuming this A-bar chain is "entangled" with the two A-chains *who*$_4$ and *who*$_1$, we yield the one-to-multiple operator-variable interpretation of a parasitic gap construction.[33]

Let's now look at the derivation of the ungrammatical (18). The structure and derivation of the adjunct clause is the same as in (17), and has been illustrated already

© 2014. John Benjamins Publishing Company
All rights reserved

in (19a). The structures in (22a–b) correspond directly to those in (19b–c) above, the only difference being that the "real gap" is a subject not an object.

(22) a. $[_{vP}$ who$_4$ $[v$ $[_{VP}$ met you]]]
b. $[_{vP}$ [who$_4$ $[v$ $[_{VP}$ met you]]] $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students $[v$ $[_{VP}$ criticized who$_1$]]]]]]]]]

We follow Chomsky (2008) in assuming that there can be no step equivalent to (20) in the derivation of (18) (recall the discussion of (7) above). Thus the next relevant stage of the derivation to consider is the structure at the completion of the root CP phase, shown in (23)

(23) $[_{CP}$ who$_6$ [C $[_{TP}$ who$_5$ [T $[_{vP}$ [who$_4$ $[v$ $[_{VP}$ met you]]] $[_{CP}$ who$_3$ [before $[_{TP}$ students [T $[_{vP}$ who$_2$ [students $[v$ $[_{VP}$ criticized who$_1$]]]]]]]]]]]]

When the domain TP is transferred at the completion of the root phase, unlike what we saw in the case of the derivation of (17), the *wh*-phrase in the edge of the adjunct phrase, *who*$_{2/3}$, will not be recognized as a copy of any higher occurrence. Even if it turns out that *who*$_{4/5}$ commands *who*$_{2/3}$, the nature of the positions is such that the former is defined as an A-occurrence and the latter as an A-bar occurrence. We assume, following Chomsky (2008), that a chain must be uniform, and that A-positions and A-bar positions do not interact with respect to copy recognition, perhaps because they exist in different dimensions.[34]

We should also address the question of why *who*$_{2/3}$ cannot move further, to the outer-spec of *v*P, creating the required "link" between the main clause and adjunct clause A-bar chains, equivalent to the intermediate copy formed by movement of *who*$_4$ in (20). The answer is straightforward if we assume that the adjunct must adjoined to *v*P (or perhaps TP; see below), in which case *who*$_3$ is not in the search domain of *v*. However, we must also question why *who*$_{2/3}$ cannot serve as a goal of the root probe C, thus being immediately recognized as copy of *who*$_6$. In this particular case, even if the movement in question were possible, such a derivation would presumably leave the A-chain consisting of *who*$_{4/5}$ "unconnected" to the A-bar chain, hence unable to be bound as a variable.[35]

In any event, there is good reason to believe that the movement in question is impossible in principle. Perhaps the most well known property of parasitic gaps is that they must indeed be parasitic on a "real" gap. Thus (24) is ungrammatical, illustrating that adjuncts are in general islands for extraction.

(24) *(Guess) $[_{CP}$ who [you met John [before students criticized *t*]]]

However, the island status of adjuncts does not follow immediately from anything we have said so far. Furthermore, we cannot say that adjuncts are opaque domains for all relations, since we have crucially appealed to the system's recognition of *who*$_{5/6}$ and *who*$_{2/3}$ in (21) as copies, despite the fact that they are separated by an adjunct boundary.

© 2014. John Benjamins Publishing Company
All rights reserved

Thus the generalization appears to be that a probe cannot look into an adjunct to find a matching goal, whereas the process that examines a transferred domain and determines whether or not two occurrences in that domain are copies, which clearly does not involve search by a probe, is not subject to such a condition. This distinction is natural, we might add, given that establishing a probe-goal relation by way of minimal search is a computational process that takes place in the confines of narrow syntax, whereas copy/chain identification is essentially a matter of interpretation, which takes place at the interface. Also, it remains to be answered exactly why search by a probe into an adjunct, even the edge of an adjunct, is prohibited, but the problem is more general. Elements inside of complex specifiers and adjuncts do not give rise to intervention effects. This raises the question of exactly how minimal search works when a probe encounters certain ambiguities, such as when probing into the structure K = {α, β}, where both α and β are complex. We will not be able to resolve these complex issues here, deferring them for future research (see Uriagereka 2012 for an analysis of subject condition effects in terms of multiple spell-out).

Finally, note that, given our assumptions so far, we predict that an adjunct phrase containing a parasitic gap cannot be attached any higher than to a projection of *v*, the reasoning being that the crucial occurrence serving as the "intermediate link" in the specifier of *v*P would otherwise not command into the adjunct. This might seem problematic given that adjuncts can apparently be stranded under so-called VP ellipsis (which we will take to be *v*P ellipsis). As (25a) shows, an adjunct can be part of the *v*P, undergoing ellipsis together with it, but as the relative grammaticality of (25b) demonstrates, it is also possible for the adjunct to appear as a remnant, in which case we assume it must be located outside *v*P.

(25) a. This paper, John criticized *t* before reading it and Bill did [$_{vP}$ e] too.
b. This paper, John criticized *t* before reading it and Bill did [$_{vP}$ e] after reading it.

However, if the adjunct contains a parasitic gap, as in (26), stranding of the adjunct under ellipsis is impossible, as the ungrammaticality of (26b) indicates.[36]

(26) a. This paper, John criticized *t* before reading *e* and Bill did [$_{vP}$ e] too.
b. *This paper, John criticized *t* before reading *e* and Bill did [$_{vP}$ e] after reading.

This contrast, between (25b) and (26b), is exactly what is predicted under our analysis.

In this subsection, we have sketched an approach to parasitic gaps in terms of the ideas about chains presented in Section 3. Our analysis captures the chain-like nature of such constructions in a natural way, and notably it does not require us to complicate the notion of search, as in the sideward movement (or parallel merge) account. We have demonstrated that the anti-c-command requirement follows in a straightforward way in our approach.

© 2014. John Benjamins Publishing Company
All rights reserved

5. Conclusion

In this paper, we have entertained the idea that accessing a lexical item α multiple times in a derivation gives rise to a sequence of occurrences of α, assuming no ontological distinction between types in the lexicon (or lexical array) and tokens in a phrase marker. The system treats members of this sequence that fall next to one another, within some well-defined *space*, as a chain (a sequence of "copies"), which collapses as a single occurrence for the purposes of interpretation at the intentional and sensory-motor interfaces. This approach does not rely on any special coding devices to distinguish copies from repetitions, thus there is no departure from inclusiveness. We have furthermore argued that our proposals can explain the core chain-like properties of OC and parasitic gaps, in a way that is both very straightforward and natural given the independent workings of the system.

The ideas sketched here are programmatic in nature, so many important but difficult issues have been set aside. As part of future research, we hope to extend the empirical reach of the program to see whether it can be fruitfully applied to account for a range of other phenomena that exhibit chain-like behavior, but which are not easily amenable to standard treatments in terms of movement.

Notes

* This paper was originally presented at *The Minimalist Program: Quo Vadis?* at the University of Potsdam in 2011. We would like to thank the organizers, especially Peter Kosta and Lilia Schürcks, for a stimulating workshop and for their efforts in bringing the current volume to life. We offer our gratitude to an anonymous reviewer, whose thoughtful and extensive review proved to be very helpful. We also would like to acknowledge Bob Berwick, Michael Jarret Baume, Jeff Bubb, Tomohiro Fujii, Ángel Gallego, Eric Groat, Wolfram Hinzen, Norbert Hornstein, Hisatsugu Kitahara, Diego Krivochen, Howard Lasnik, Terje Lohndal, Jairo Nunes, Massimo Piattelli-Palmarini, Doug Saddy, students in a course (taught by Uriagereka) at the Universitat Autònoma de Barcelona in 2013, and members of the audience of the aforementioned workshop for providing useful comments and suggestions on this work at various stages. The research reported here has been supported by a Grant-in-Aid for Scientific Research (#23520458; PI: Roger Martin) from the Japan Society for the Promotion of Science.

1. For ease of illustration, we ignore labels and some details of the structure that are not relevant to the discussion.

2. Concerning the indexation algorithm, prior to any applications of Select, the set of syntactic objects will necessarily be empty. Although Chomsky does not make it explicit, presumably the very first application of Select creates SO_1; in other words, if there is no SO already in the derivation, n = 0.

© 2014. John Benjamins Publishing Company
All rights reserved

3. See also Collins and Stabler (2011) for an explicit formalization of the system in Chomsky (1995), which highlights the problems discussed here. Thanks to an anonymous reviewer for pointing out the relevance of this work.

4. But presumably a similar or equivalent operation would be implicated in the formation of a numeration from the lexicon.

5. In more recent works, Chomsky (2008, 2013) seems to move away from the claim that the concept of numeration is crucial in determining identity. For example, Chomsky (2008) states: "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…" He further notes that this avoids the violation of inclusiveness that arises in earlier work, and in this sense the approach is more in the spirit of the proposals made in this paper. However, Chomsky stipulates that each independent extraction of an item from the lexicon yields a *distinct* syntactic token. Furthermore, these nuances must again somehow be communicated to the interfaces, which would require a significant amount of derivational "look back." Chomsky (2008) suggests, as seen in the quote above, that the required memory can be limited to phases. Chomsky (2013) similarly appeals to the idea that all operations and transfer apply simultaneously at the phase level. It is not clear how this proposal accounts for A-bar chains, which can span across an indefinite number of phases. See Section 3.2 below for further discussion.

6. We set aside whether the derivation must access a predetermined (proper) subset of the lexicon, i.e., a lexical array, or whether it can access the lexicon directly, as argued for by Collins (1997), assuming the former for the purpose of the discussion.

7. Kitahara 2000 is another explicit attempt to deal with the copy/repetition problem. Kitahara proposes that interpretation takes place whenever an NP's Case features are checked and argues that this eliminates the need for the inclusiveness violating aspects of Chomsky's (1995) system. This idea — that cyclic interpretation is relevant to distinguishing copies from repetitions — is similar in many respects to the one we pursue in this paper, though he does not, for example, deal with the problems raised by A-bar chains. See also Kobele 2006 for much related discussion.

8. A question that often arises is what is meant by *lexically identical elements*. The cleanest case is identical lexical items that have been accessed more than once in a derivation. However, complex phrases pose exactly the same questions. A radical approach would be to assume that the copy/repetition distinction is only relevant to lexical items, and that what we call two occurrences of a complex phrase like *the students* should actually be understood as two occurrences of *the* and two occurrences of *students*, so that in *the students were arrested*, there are two chains, one headed by *the* and one headed by *students*. Another approach is to assume that identity of complex phrases is recognized compositionally at the interface. One occurrence of, say, the set {the, students} is recognized as identical to some other occurrence of the very same set. A related possibility is that complex phrases are "flattened out" by the process multiple spell-out, essentially becoming the equivalent of lexical compounds, in which case the identity is more easily computed; e.g., *the-student-of-chemistry* as opposed to {the {student, {of, chemistry}}}. Uriagereka (2012) makes such a proposal for complex "left-branches," explaining, for example, why they do not allow subextraction, so the question would be what to say about complex "right-branches."

9. The situation just described is consistent with treating cognitive categories as topological closed spaces, akin to bubbles. Clearly two adjacent bubbles may fuse into a (larger) bubble,

© 2014. John Benjamins Publishing Company
All rights reserved

which is topologically the same entity; this fusion, however, doesn't happen at a distance: the relevant discontinuity results in each bubble keeping its own spatial identity. Similarly, Mercier, Kawai and Errington (2013) and Zhou, Martinez and Fredberg (2013) provide compelling arguments for treating cells as bubble-like elements. The main issue is whether these metaphorical approaches to cells, which have a clear import and consequences in biology, can be seriously extended to cognition.

10. Here *students*$_1$ is not accessible to *students*$_3$ under the PIC. But if the *copy of* relation is transitive, we would still predict the ungrammatical (5c) to be possible.

11. This would have to be augmented with the assumption that a chain, once transferred, cannot be extended or combined with another chain in subsequent phases.

12. In this paper we are noncommittal about whether or not objects in general move to a specifier position (generalized EPP) in English or other languages. However, we assume that such movement, if it does take place, is to the specifier of VP not *v*P, as in the feature inheritance analysis of Chomsky (2008).

13. In other words, there are no non-uniform chains *per se*. However, numerous difficult questions arise concerning this analysis of operator-variable constructions, including how the independent A and A-bar chains relate (or become necessarily entangled), and how this relation, whatever it is, gets translated semantically as variable binding. There are other technical questions as well, such as what happens to the various uninterpretable features assumed to be contained by LI *who* (e.g., Case and *wh*-features). We assume, essentially following Chomsky, that if a feature contained by element is valued/erased, it is somehow valued/erased in every copy of that element as well. So if the A-chain is a completely distinct from the A-bar chain, Case, for example, had better only be contained by the former, uninterpretable *wh*-features only by the latter, and so on.

14. An anonymous reviewer suggests that this departure from Chomsky's assumptions regarding the two varieties of merge is undesirable since it undermines the attempt to trivialize the transformational component of the grammar. However, we believe we need to start by assuming certain basic facts about language, such as the existence of context-sensitive operations, and then try to explain them in an optimal way.

15. Otherwise we lose the rationale for successive cyclic movement (movement to the edge of an intermediate phase) that phases may provide. If some α with an uninterpretable feature is not moved to the edge of the phase, this feature will never be valued since α is no longer accessible; α containing the still offending feature will be transferred to the interfaces as such, resulting in a crash. On the other hand, as we discuss below, it is not clear why movement to an intermediate edge is able to prevent the derivation from crashing in the first place, since the relevant uninterpretable feature presumably remains unvalued at the point of transfer.

16. Although it depends on the exact timing of transfer and other operations taking place at the phase level (such as valuation), it is likely that *who*$_2$ in a sentence like (6a) is in fact valued at the point of transfer. However, our point holds if we further embed the structure in (6b) — e.g., in a sentence like *Guess who professors think students criticized* — so that *who*$_3$ is an "intermediate trace." There are potential consequences related to this. If the relevant feature of *who*$_2$ is valued at the point of transfer of the domain TP in a sentence like (6a), the system cannot determine whether *who*$_2$ is an "intermediate trace" (part of an incomplete chain), or whether it is the head

© 2014. John Benjamins Publishing Company
All rights reserved

of a chain that is ready to be sent to the interfaces, simply by examining whether it has an unvalued feature or not. Thus the algorithm for determining this must be based on the different nature of the objects resulting from external merge and internal merge discussed above. On this view, who_2 and who_3 are in fact *one and the same* occurrence, which is both inside and outside of the domain, simultaneously, at the relevant point of transfer. Therefore the system does not attempt to interpret the chain at this stage.

17. Weakening transfer in this way may have the further consequence of making it impossible to derive the effects of the PIC in (4) in the way alluded to above.

18. Recall that we are assuming that who_1 and who_3 are in fact the same syntactic object if internal merge has applied, as discussed above. A related question concerns how the first link in a uniform A-bar chain is created, if we assume the two-chain theory of operator-variable constructions. It is not clear, if we take such an approach seriously, whether who_3 should be taken to come about by way of internal merge, or whether we should assume that it is introduced by external merge. Either way, it seems that for the A-bar chain to become "entangled" with the A-chain, the tail of the former has to be to very local to the latter.

19. This depends on whether the relevant features remain and are visible at the point of interpretation. In this context, it is worth noting that the phenomenon of Case may also be relevant when considering how the system sorts out copies from repetitions. As Uriagereka (1997) notes (see also Uriagereka 2008), we must ask not only why uninterpretable Case features exist, but also why there are different Case *values*, such as nominative, accusative, and dative. Uriagereka goes on to suggest that the purpose of Case, and in fact the array of Case values attested, may be for the system to have a way to "sort out" arguments appearing within a clause, which otherwise might be indistinguishable in terms of their lexical makeup (Kitahara 2000 also views Case as being closely related to the mechanism that distinguishes copies from repetitions, although he states things in rather different terms). If this is true, there may be some redundancy between this sort of analysis and the one we have been discussing above. However, note that although it might appear that different Case values would suffice to distinguish subjects from objects in simple transitive sentences (e.g., *He likes him*), this can be avoided by embedding the relevant examples under an ECM verb (e.g., *I believe him to like him*), where the fact that the two accusative pronouns are interpreted as repetitions, apparently demonstrates that an analysis based on Case alone is not sufficient. On the other hand, Case may be the only thing distinguishing the two otherwise identical internal arguments in a ditransitive sentence, if these arguments are contained in the domain of the same phase.

20. On our view OC is equated with infinitival clauses that can only appear with an empty subject. We assume that non-obligatory control (NOC), on the other hand, corresponds to so-called *for-infinitivals*, infinitival clauses that can appear with either an empty or lexical subject (the latter typically immediately preceded by *for*). One possible set of exceptions to this generalization involves control in questions (e.g., *Mary knows how to fix the car*). This is essentially the view expressed by Chomsky and Lasnik (1977), Lasnik and Fiengo (1974), and Williams (1980), and differs significantly from the approach taken by Hornstein (1999) and many others, where the OC/NOC distinction is seen as a distinction between local vs. long-distance control (or *Equi* vs. *Super-Equi* in the sense of Grinder 1970).

21. Chomsky has stressed this point in numerous places (for example Chomsky (1995: 313) and Chomsky (2000: 103, 143)). Suppose a θ-configuration is something like {arrested, students}. Yet

© 2014. John Benjamins Publishing Company
All rights reserved

in a sentence like *students were arrested*, what is transferred to the interface on Chomsky's view is the chain CH = <{were {arrested, students}}, arrested>, which is not the θ-configuration in question. If the chain is viewed instead as the pair <{students {were {arrested, students}}}, {arrested, students}>, then we could at least appeal to the fact that it contains the relevant θ-configuration. However, containing a θ-configuration and being one are again very different things. It has often been argued that interpretation can only apply to a single link of a chain, and not to all (or even multiple) links simultaneously. This is obvious on the "sensory-motor side" (explaining it is a different matter; see Nunes 2004 for one attempt) and has been argued to be true for "meaning" as well by Hornstein (1998) and Martin and Uriagereka (2008), where it is shown that binding and scope must be determined in a single position in the chain. One way to think of this is that interpretation requires *collapsing* the chain, viewing the interpreted element in a *unique* configuration, not in multiple configurations simultaneously. With this in mind, consider the fact that scope and binding possibilities are not restricted to the θ-configuration in a chain. In other words, collapse need not be at the θ-position. One might suggest that the chain can be looked at differently for the purposes of thematic interpretation and for intentional interpretation (e.g., scope and binding). However this is the same as saying that the "meaning side" consists of two components, a conceptual one and an intentional one, that are distinct enough that a chain can collapse differently in each of these components. This is essentially what we argue below.

22. Here we set aside difficult issues surrounding the nature of the EPP and tentatively assume that the subject of the infinitival does not raise to the TP-spec position. Nothing hinges on this decision.

23. Alternatively, we might assume that it is a *v*P, which has a number of interesting consequences. For example, if *to* is an instantiation of the head *v*, as would be virtually necessary under the *v*P complement hypothesis, it provides an explanation for the fact that *to* most naturally follows negation, as opposed to what we see with elements assumed to be located in T (e.g., *John persuaded Mary not to kiss him*). Of course, even if the infinitival complement was taken to be a TP or a CP, one could still stipulate that *to* is located in *v*. More directly supporting of the *v*P complement hypothesis, however, is the observation by Lasnik and Fiengo (1974) that OC complements must have agentive/causal subjects. This would follow if what is selected is a projection of *v*, which is responsible for the Agent/Cause θ-role, as opposed to any other type of "little" verbal projections, such as those involved in intransitive or experiencer constructions. If OC complements were categories larger than *v*P, however, such selection would be difficult to state.

24. Apparently, there is no "escape hatch" for A-movement. In other words, movement to the edge of a phase is restricted to operators, for some reason. We should also note that the interpretation of $John_1$ and $John_2$ as copies in (10) is quite plausibly blocked due to the fact that $Mary_1$ intervenes as well (recall the effect of the images of the bird and the butterfly when interspersed among the images of the man in the analogy form visual perception discussed above), raising the usual questions of redundancy.

25. There are some apparent exceptions to the generalization that a chain is phonetically realized only in one of its positions; see Nunes 2004 for discussion. Also, some researchers have analyzed certain word order effects, including so-called covert movement, as involving pronunciation of a lower copy, as opposed to the highest one; for a detailed overview, see Bošković and Nunes 2007.

© 2014. John Benjamins Publishing Company
All rights reserved

26. The idea of collapse needs clarification. Aside from the analogy we gave about visual perception of discrete images in motion above, why should collapses occur? Obviously, two clones or identical twins do not "collapse" into a single individual, no matter how close to one another they stand. In a sense, collapse is something that is more familiar in the subatomic world than the classical one. In quantum terms, more specifically, the wave/particle duality is essentially the notion that an entity simultaneously possesses localized properties (qua particle) and distributed properties (qua wave). If we think of the former as something like position of an entity α (the local configuration created by external Merge of α to β) and the latter as the set of configurations across which α is distributed, we can perhaps view collapse of the chain along the lines of collapsing a wave function (for a more detailed presentation of this kind of interpretation of collapse, still mostly at an intuitive level, see Martin and Uriagereka 2008).

27. It should be noted that the same sort of paradox arises in Chomsky's (1995, 2000, 2001, 2008) analyses, even for a sentence like *someone seems to everyone to have been elected*. Chomsky assumes that what gets sent to the CI interface is a chain consisting of multiple occurrences of *someone*. The chain cannot be assigned a θ-role since it is not a configuration, as discussed above. But even if we assume that the chain can be undone, reconstructing *someone* to its "base" configuration, although we account for θ-role assignment, we lose an explanation of the fact that *someone* can take wide scope over *everyone*.

28. We assume that the sensory-motor interface is essentially the same in that it interprets chains, which collapse into a unique determinate configuration, yielding the unambiguous command relations that arguably play a crucial role in determining linear order.

29. It should be noted that exactly the same issue arises in the movement approach to OC. For example, to rule out (i) Hornstein (1999) appeals to the inverse Case filter, namely the condition that a Case assigner must discharge its Case; *believe* fails to check its Case feature and the derivation crashes.

(i) [John$_i$ T [t_i *v* believes [t_i to have [t_i *v* [met Mary]]]]]

To allow (ii), then, it must be assumed that *try* need not have a Case feature.

(ii) [John$_i$ T [t_i *v* tried [t_i to [t_i *v* [meet Mary]]]]]

30. As (13) demonstrates, not all verbs lose their ability to assign Case when taking an infinitival complement and no nominal complement. It is worth recalling that Exceptional Case Marking in (13a) is not common across languages. For example, in Romance the pattern of grammaticality in (13) is reversed: only control is possible.

31. A related approach is Kasai's (2010) multi-dominance analysis, which is based on parallel merge (Citko 2005). Insofar as this sort of approach invokes internal merge, it suffers the same problem as sideward movement, namely, the complication of the mechanism of merge/search.

32. Whether this involves a probe-goal relation triggering internal merge, or whether *who*$_5$ is inserted by external merge, relates to problems that arise concerning the two-chain analysis of operator-variable constructions. Although these questions are far from trivial, for the purposes of the argument here it suffices to assume that *who*$_5$ arises in whatever way the equivalent element in (6) arises.

© 2014. John Benjamins Publishing Company
All rights reserved

33. Again we set aside the important question of exactly how who_4 and who_1 are related to $who_{5/6}$ and $who_{2/3}$ respectively (namely, the general problem of how A-bar chains and A-chains relate, especially within the two-chain analysis of operator-variable constructions).

34. Although this may seem like a mysterious notion, we should note that the locality involved in the recognition of two occurrences as copies may then *reduce to structural adjacency*. This is conceptually desirable (recall the analogy from visual recognition above) but possible only if the A-dependencies, A-bar dependencies, and presumably also head-dependencies, can somehow be thought of as existing in separate dimensions within the phrase marker.

35. This depends on how an A-bar chain (operator) becomes entangled with an A-chain (variable), if these really are separate chains (namely, if the two-chain analysis is correct). If internal merge (hence minimal search) is not required, but rather the entanglement can arise simply by virtue of the chains being very close to one another, then whether this suffices as a solution would depend on the exact nature of that locality.

36. An anonymous reviewer notes that an informant did not agree with the contrast in question. To our ears, the contrast is significant enough. The same reviewer also points out that an example involving gapping such as (i) appears to be fine.

(i) This paper, John criticized before reading, and that paper after reading.

Although we are not certain about the judgment of (i) (it sounds to us to be somewhat better than (26b), though perhaps still not perfect), the question in our terms is whether gapping involves ellipsis — a topic without consensus in the literature.

References

Bošković, Željko and Nunes, Jairo. 2007. "The copy theory of movement: A view from PF." In *The Copy Theory of Movement*, Norbert Corver and Jairo Nunes (eds), 13–74. Amsterdam: John Benjamins. DOI: 10.1075/la.107.03bos

Bowers, John. 2008. "On reducing control to movement." *Syntax* 11: 125–143. DOI: 10.1111/j.1467-9612.2008.00103.x

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

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

Chomsky, Noam. 2000. "Minimalist inquiries: The framework." In *Step by Step: Essays on Minimalism in Honor of Howard Lasnik*, Roger Martin, David Michaels and Juan Uriagereka (eds), 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. "Derivation by phase." In *Ken Hale: A Life in Linguistics*, Michael Kenstowicz (ed.), 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. "On phases." In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Robert Freidin, Carlos P. Otero, Maria Luisa Zubizarreta (eds), 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. "Problems of projection." *Lingua* 130: 33–49. DOI: 10.1016/j.lingua.2012.12.003

Chomsky, Noam and Lasnik, Howard. 1977. "Filters and control." *Linguistic Inquiry* 8: 425–504.

© 2014. John Benjamins Publishing Company
All rights reserved

Chomsky, Noam and Lasnik, Howard. 1993. "The theory of principles and parameters." In *An International Handbook of Contemporary Research*, Joachim Jacobs, Arnim Von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds), 506–569. Berlin: Mouton de Gruyter.

Citko, Barbara. 2005. "On the nature of merge: External merge, internal merge, and parallel merge." *Linguistic Inquiry* 36: 475–497. DOI: 10.1162/002438905774464331

Collins, Chris. 1997. *Local Economy*. Cambridge, MA: MIT Press.

Collins, Chris and Stabler, Edward. 2011. "A formalization of minimalist syntax." Unpublished manuscript.

Engdahl, Elizabeth. 1984. "Parasitic gaps." *Linguistics and Philosophy* 6: 5–34. DOI: 10.1007/BF00868088

Grinder, John. 1970. "Super equi-NP deletion." In *Papers from the Sixth Regional Meeting - Chicago Linguistic Society, Chicago Linguistic Society* (ed.), 297–317. University of Chicago.

Hale, Kenneth and Keyser, Samuel Jay. 1993. "On argument structure and the lexical expression of grammatical relations." In *The View from Building 20: Essays in Honor of Sylvain Bromberger*, Kenneth Hale and Samuel Jay Keyser (eds), 53–110. Cambridge, MA: MIT Press.

Hornstein, Norbert. 1998. "Movement and chains." *Syntax* 1 (2): 99–127. DOI: 10.1111/1467-9612.00005

Hornstein, Norbert. 1999. "Movement and control." *Linguistic Inquiry* 30: 69–96. DOI: 10.1162/002438999553968

Kasai, Hironobu. 2010. "Parasitic gaps under multiple dominance." *English Linguistics* 27 (2): 235–269.

Kitahara, Hisatsugu. 2000. "Two (or more) syntactic categories vs. multiple occurrences of one." *Syntax* 3 (3): 151–158. DOI: 10.1111/1467-9612.00030

Kobele, Gregory. 2006. "Generating Copies: An Investigation into Structural Identity in Language and Grammar." Ph.D. dissertation, UCLA.

Lasnik, Howard and Fiengo, Robert. 1974. "Complement object deletion." *Linguistic Inquiry* 5: 535–571.

Martin, Roger. 1996. "A Minimalist Theory of PRO and Control." Ph.D. dissertation, University of Connecticut.

Martin, Roger. 2001. "Null case and the distribution of PRO." *Linguistic Inquiry* 32: 141–166. DOI: 10.1162/002438901554612

Martin, Roger and Uriagereka, Juan. 2008. "Competence for preferences." In *Gramatika Jaietan: Patxi Goenagaren Omenez (Grammar Festival: A Tribute to Patxi Goenaga)*, Xabier Artiagoitiak and Joseba Lakarrak (eds), 561–572. Bilbao: University of the Basque Country.

Martin, Roger and Uriagereka, Juan. 2013. "Equi redux." In *Deep Insights, Broad Perspectives: Essays in Honor of Mamoru Saito*, Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi and Koji Sugisaki, Asako Uchibori (eds), 292–308. Tokyo: Kaitakusha.

Mercier, Romain, Kawai, Yoshikazu and Errington, Jeff. 2013. "Excess membrane synthesis drives a primitive mode of cell proliferation." *Cell* 152 (5): 997–1007. DOI: 10.1016/j.cell.2013.01.043

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

© 2014. John Benjamins Publishing Company
All rights reserved

San Martin, Itziar and Uriagereka, Juan. 2002. "Infinitival complementation in Basque." In *Erramu Boneta: A Festschrift for Rudolf P. G. de Rijk*, Xabier Artiagoitia, Patxi Goenaga and Joseba Lakarra (eds), 597–610. Bilbao: University of the Basque Country.

Uriagereka, Juan. 1997. "Formal and substantive elegance in the minimalist program." In *The Role of Economy Principles in Linguistic Theory*, Chris Wilder, Hans-Martin Gärtner and Manfred Bierwisch (eds), 170–204. Berlin: Akademie Verlag.

Uriagereka, Juan. 2008. *Syntactic Anchors*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511481482

Uriagereka, Juan. 2012. *Spell-Out and the Minimalist Program*. New York: Oxford University Press.

Williams, Edwin. 1980. "Predication." *Linguistic Inquiry* 11: 208–238.

Zhou, Enhua H., Martinez, Fernando D. and Fredberg, Jeffrey J. 2013. "Cell rheology: Mush rather than machine." *Nature Materials* 12: 184–185. DOI: 10.1038/nmat3574

© 2014. John Benjamins Publishing Company
All rights reserved