Labeling by Minimal Search: Implications for Successive-Cyclic A-Movement and the Conception of the Postulate ''Phase''

Samuel David Epstein
Hisatsugu Kitahara
T. Daniel Seely

We argue that Chomsky's (2013) ''label identification by minimal search'' explains ''obligatory exit'' from intermediate positions, not only in the successive-cyclic Ā-movement phenomena that Chomsky analyzes, but also in (phase-internal) successive-cyclic A-movement. Moreover, it does so by employing simplest Merge and third-factor minimal search for label identification. Our extension of Chomsky's analysis to A-movement operates without any appeal to Merge-over-Move or to lexical arrays or subarrays. This in turn renders the concept ''phase'' itself no longer necessary in analyzing the core cases of illicit A-movement, shown to reduce to labeling failure. Implications of this result and the nature of the long-standing evidence for strict cyclicity are discussed.

Keywords: A-movement, phases, labeling, minimal search, simplest Merge

1 Simplest Merge and Label Identification by Minimal Search

One recent development in Minimalist theory (Chomsky 2004, 2007, 2008, 2013) is that Merge, formulated in the simplest form, applies *freely*[1] (i.e., without ''teleological purpose'')[2] as long as it conforms to third-factor principles such as the no-tampering condition and the condition of inclusiveness. Merge, by hypothesis, is no longer operating *in order to* create a configuration that allows interface-illegitimate features to be checked; it is not ''purposeful'' in the sense of early

We are deeply indebted to Noam Chomsky for his interest in and countless contributions to our work, including to this article. We are grateful to Noam for coming to the LSA Summer Institute at the University of Michigan in 2013 and talking to the students in a course the three of us co-taught, entitled ''Derivational Minimalism.'' In addition, we thank all the students in our Institute course for their interest, participation, and terrific questions and contributions to the class. Marlyse Baptista, Vicki Carstens, Norbert Hornstein, Acrisio Pires, and Rick Lewis were also kind enough to sit in on some of our classes, and their contributions were truly invaluable. We owe Erich Groat our sincere appreciation for a detailed, thoughtful, written ''critique'' of an earlier version of this article. We are also indebted to Ruriko Kawashima for extensive and incisive feedback on previous versions of this article. In addition, we thank Masashi Nomura for his valuable comments on the manuscript. We are also grateful to the following groups for listening to aspects of this research and providing important commentary: the Keio Study Group of Generative Grammar, the University of Michigan Syntax Research (aka ''Syntax Support'') Group, and Generative Linguistics in Musashino. Aspects of this research were presented at the 31st conference of the English Linguistic Society of Japan in a workshop entitled ''Simplest Merge and Label Identification by Minimal Search.''

Finally, we thank the Program for the Advancement of Next Generation Research Projects, Keio University, for its support of this research.

[1] See Boeckx 2010 and Ott 2010 for discussion.

[2] See Epstein 2007 for extensive discussion of what he calls ''internalist functional explanation'' as it concerns the theory of ''purposeful'' application of Merge.

Linguistic Inquiry, Volume 45, Number 3, Summer 2014
463–481

doi: 10.1162/ling_a_00163

Minimalism in that it is no longer driven by convergence conditions (e.g., the valuation of ϕ-features or Case features). Concepts such as ''derivational economy'' no longer determine—internal to the narrow syntax (NS)—when Merge must apply and when it cannot.[3] Rather, Merge is free; it optionally applies (and so crashing happens; that is, the system is not crash-proof).[4]

What prompted the development from purposeful Merge to freely applying Merge? In large part, it was due to the unification of internal and external Merge; that is, IM and EM have been reanalyzed as instantiations of the *single* operation Merge (with *internal* and *external* simply expositional mnemonics for *one* operation accessing different resources).[5] It follows that if IM and EM are but one operation, then what is true of ''one'' is true of the ''other.'' Thus, since EM applications are by hypothesis not driven by convergence conditions (e.g., the valuation of ϕ-features or Case features), IM cannot be driven by the quest for convergence. Similarly, EM was hypothesized to be ''driven'' (at least in part) by θ-theory. But since IM is arguably not driven by θ-theory (but see Bošković 1994, O'Neil 1997, and Hornstein 1999 for important contrary views), and since IM and EM are two possible applications of the single operation Merge, then EM cannot be driven by θ-theory either. If so, there is nothing left for Merge to be driven by (not for convergence and not for coherence). Consequently, it is not ''driven'' at all, which means that it is simply *available* to apply, and does so optionally and freely (similar, in its freedom, to the Move α operation within Government-Binding (GB) Theory). It applies not because it must, but because it can.

Under the simplest conception of unified Merge, Chomsky (2013) argues that Merge(α, β) yields {α, β} with no label projection or linear order, and independently of the character of α and β.[6] We call this *simplest Merge*.

(1) *Simplest Merge*
Merge(α, β) → {α, β}

Given (1), a syntactic object SO constructed from α and β by Merge is just {α, β}. Merge puts the two objects α and β into a relation,[7] the output being represented as a two-membered set. Nothing more. Thus, unlike the output of Merge in Chomsky 1995a and much subsequent work, the output of Merge in this conception does not overtly encode a label; contra Chomsky's (1995a) notation, there is no constructed set {δ, {α, β}}, where δ represents the label (as either H(α) or H(β)) identifying the relevant properties of {δ, {α, β}} (see Chomsky 1995a,b). Rather, under simplest Merge there is just {α, β}.

For an SO to be interpreted, however, it is necessary to know what kind of object it is. Chomsky (2013) takes *labeling* to be the process of finding the relevant (object identification)

[3] Note too that Chomsky (2007:22) eliminates the Activity Condition; with respect to raising to Spec,T ''. . . no recourse to the activity condition is needed.''

[4] See Frampton and Gutmann 2002 for detailed discussion of what they call a crash-proof model of syntax. For recent discussion of crash-proof grammars, see also Putnam 2010.

[5] See Kitahara 1997 for an earlier attempt to reformulate Merge and Move as consisting of more elementary operations. See also the important discussions in Nunes 2001, 2004.

[6] To the best of our knowledge, Collins (2002) was the first within the generative tradition to argue for the elimination of labels. For extensive further discussion of the elimination of labels and of the ''projection'' component of Merge, see Seely 2006; for further important discussion, see Boeckx 2006, Jayaseelan 2008, and Hornstein 2009, among others.

[7] See Epstein et al. 1998 for the proposal that Merge creates the only syntactically significant relations.

information of {α, β} generated by simplest Merge. He proposes that such labeling is ''just minimal search, presumably appropriating a third factor principle, as in Agree and other operations.'' There are two cases in question: {H, XP} and {XP, YP}.[8] Chomsky outlines how minimal search operates to find the label of each case, as follows:

(2) Suppose SO = {H, XP}, H a head and XP not a head. Then minimal search will select H as the label, and the usual procedures of interpretation at the interfaces can proceed. (Chomsky 2013:43)

(3) Suppose SO = {XP, YP}, neither a head. Here minimal search is ambiguous, locating the heads X, Y of XP, YP, respectively. There are, then, two ways in which SO can be labeled: (A) modify SO so that there is only one visible head, or (B) X and Y are identical in a relevant respect, providing the same label, which can be taken as the label of the SO. (Chomsky 2013:43)

In {H, XP}, minimal search immediately finds a lexical item H (a bundle of features, provided by the lexicon) and a two-membered set XP (a set-theoretic object constructed by Merge). The former, being a lexical item, makes available what matters to the interface systems, while the latter, being a set, does not; it requires further search. Thus, H is identified as the label of {H, XP}.[9]

In {XP, YP}, minimal search is ambiguous, locating (with equally minimal ''depth of search'') each of the two heads X and Y of XP, YP, respectively. It is assumed that such failure to identify a unique head in {XP, YP} prevents labeling, and since labels are required for interpretation at the conceptual-intentional interface (CI), if the object lacking the label appears at CI, it violates Full Interpretation. For Chomsky (2013), however, there are two ways in which unlabeled {XP, YP} can be salvaged: (a) this SO must be modified so that there is only one visible head (i.e., X or Y (not X and Y)) or (b) X and Y share some prominent features and the shared features are taken as the relevant identifying information (i.e., as the label of {XP, YP}).[10]

To see how minimal search finds the label in each case, let's examine concrete instances. Take SO_i = {v, VP}. Here minimal search finds v as the label of SO_i since v is unambiguously identifiable by applying (2); in any {H, XP}, the head H is always found with ''less search'' than any feature-bearing (hence relevant information-bearing) element within XP.

Next, take SO_j = {NP, {v, VP}}. Here minimal search is ambiguous, locating two relevant heads, N and v. If the syntactic object is left as is, labeling fails and Full Interpretation is violated at CI in that CI cannot find the information it needs to identify the categorial status of this object; such identification of status is hypothesized by Chomsky (2013) to be a necessary prerequisite

[8] We leave aside the only other possibility, {H, H}. See Chomsky 2013 for discussion. We also leave aside (with Chomsky 2013) issues with adjunction. For recent discussion of ''head-head'' merger in a language without ϕ-feature agreement, see Saito 2012, 2013a,b.

[9] A question remains as to how the conceptual-intentional systems use the information that, say, {H, XP} is an H-type thing; but the idea is that they can't interpret {H, XP} without this information (see Chomsky 2013 for discussion).

[10] Note that the featural intersection of X, Y is taken as the label of {XP, YP}. Interestingly, as a reviewer points out, Chomsky (1995b) proposed this idea, but at the time rejected the existence of projection of an intersect of features borne by X and by Y since X, Y might not have intersecting features.

to CI (properly) interpreting the object. As pointed out above, one way to label SO_j is to raise NP to a higher position (= strategy (3A)), which yields (after merger of T to SO_j and subsequent subject raising) SO_k = {NP, {T, {NP, {v, VP}}}}. Chomsky (2013) takes the single NP (call it α) in SO_k to be in the domain D (in this case, in the set SO_j now embedded within SO_k) if and only if every occurrence of α is a term of D.[11] Given this, NP is taken to be not in SO_j (a term of SO_k) because SO_j does not contain every occurrence of NP as its term; rather, NP is taken to be in SO_k because SO_k contains every occurrence of NP as its term. Informally, the lower copy of NP is ''invisible'' when minimal search conducts a search for the label of {NP, {v, VP}}; it therefore ''sees'' only {v, VP} when it ''looks at'' {NP, {v, VP}}.[12] Thus, the movement of NP makes the v unambiguously identifiable; that is, minimal search finds the only ''visible'' head v as the label of SO_j. Note that this is Chomsky's account of ''EPP-driven'' subject raising, but without any appeal to the EPP (Extended Projection Principle), in any of its mysterious incarnations (see, among many others, Epstein, Pires, and Seely 2005, Epstein and Seely 2006, Bošković 2007, and the references cited regarding the ''oddity'' of the EPP). Rather, if there is no movement (and no expletive), then there is labeling failure, which in turn creates a barrier to CI interpretation.[13]

Turning to SO_k, namely, {NP, {T, {NP, {v, VP}}}}, it is also of the form {XP, YP}. Can minimal search find the label of SO_k? Chomsky answers ''yes,'' even though SO_k is a set contain-

[11] The notion ''term-of'' was originally defined as follows (Chomsky 1995a:399):

(i) a. K is a term of K.
b. If L is a term of K, then the members of the members of L are terms of K.

Given simplest Merge, however, the definition (i) requires the following modification, as proposed by Seely (2006:201):

(ii) a. K is a term of K.
b. If L is a term of K, then the members of L are terms of K.

In this article, we adopt this simplified definition of the notion ''term-of.'' See Seely 2006, Chomsky 2008:n.16, and Epstein, Kitahara, and Seely 2012b for relevant discussion. For recent discussion of the possible importance of reducing (ib) to (iib), see Epstein 2013.

[12] The invisibility of lower copies is supported by cases such as English *wh*-movement (i) (see Chomsky 2007, 2008) and Icelandic dative subjects (ii) (see Jónsson 1996, also cited in Bobaljik 2008). In each case, a lower copy of the moved element, occupying Spec,v, does not interfere with the minimal search of T.

(i) What do they like?
[wh [C [T [wh [Subj [v VP]]]]]]

(ii) Jóni líkuðu þessir sokkar.
Jon.DAT like.PL these socks.NOM
'Jon likes these socks.'
[C [NP.DAT [T [NP.DAT [v [V NP.NOM]]]]]]

Given data such as (i) and (ii), Epstein, Kitahara, and Seely (2012b) propose that α (goal) is invisible to the minimal search of β (probe) if some occurrence of α is outside β's search domain.

[13] It should be noted that with expletive insertion (as in *There is a man in the room*), the lower portion (in this case, [a man in the room]) will not have a label. It could be that there is ''short movement'' of *a man* to the ''specifier'' of some higher head, rendering the lower copy of *a man* invisible and hence allowing the lower portion in question to be labeled, as P. Now the question arises of how (i), with short movement of *a man*, is to be excluded within Chomsky's (2013) account of the EPP.

(i) * ____ will be a man in the room.

We leave the important questions of how such existentials with *there* are to be allowed, and how those without *there* are to be excluded, to future research regarding Chomsky's (2013) compelling reduction of EPP to labeling by minimal search.

ing two nonheads. Chomsky suggests that in such cases, the ϕ-features shared by the two relevant heads N and T (in finite clauses) can be the label of SO_k (= strategy (3B)); that is, when there are some prominent features shared by X and Y, minimal search *can* identify the label for {XP, YP}—namely, the features (in this case, the ϕ-features) appearing on both heads X and Y.

Chomsky (2013) takes successive-cyclic Ā-movement to support his analysis. Consider *wh*-movement in (4).[14] (We use *t(race)* only for expository purposes.)

(4) [$_\beta$ In which Texas city did they think [$_\alpha$ *t* [C [$_{TP}$ the man was assassinated *t*]]]]?

Immediately after the *wh*-phrase *in which Texas city* (hereafter, the *wh*-PP) is internally merged to the specifier (Spec) of the embedded C, the embedded clause α is of the form of {XP, YP}, where XP is the *wh*-PP and YP is {C, TP}, as shown in (5).

(5) . . . (think) [$_\alpha$[$_{XP}$ in which Texas city] [$_{YP}$ C [$_{TP}$ the man was assassinated *t*]]]

Crucially, the embedded C (selected by *think*) bears no Q-feature and as a result there is assumed to be no prominent feature shared by XP and (the Q-feature-lacking) YP. If XP remains in this intermediate position, minimal search cannot find a label for α, since there is no prominent feature (e.g., ϕ-features or Q-feature) shared by X (the head of the *wh*-PP) and Y (the head of {C, TP}). However, if the *wh*-PP raises to a higher position (rendering the lower copy of the *wh*-PP ''invisible'' inside α), then minimal search can find a unique ''visible'' head—namely, C—as the label of α. Notice that the matrix clause β of (4) is also of the form {XP, YP}, but there is a feature shared by X and Y (namely, the Q-feature of the *wh*-PP and the Q-feature presumably borne by the interrogative mood–marked C of the matrix); hence, Q can be the label of β.[15]

Chomsky's (2013) analysis thus has the following property: the ''obligatory exit'' from an intermediate position of *wh*-movement (e.g., **They think in which Texas city the man was assassinated?*) is accounted for by a labeling failure (ultimately, a CI violation of Full Interpretation). The analysis makes no appeal to a mismatch of features between *think*, which selects a [−Q] C, and the [+Q] *wh*-PP occupying the Spec of this [−Q] C. It appeals neither to an explicit Spec-head relation (defined via m-command/government) in CP; nor to any of the technical devices that have been nonexplanatorily invoked in past research like coindexing of Spec and head in CP; nor to an S-Structure level of representation (as in important prior analyses; recall Rizzi's (1997) *Wh*-Criterion and Lasnik and Saito's (1984, 1992) S-Structure condition that blocks a [+Q] *wh*-phrase from occupying the Spec of a [−Q] C in English). Nor does Chomsky's analysis appeal to the EPP (obligatory Spec,T or Spec,C), as noted above for A-movement. No such descriptive technicalia are invoked.

[14] We ignore a lower copy of *the man* here, since it is irrelevant to our discussion. (In response to a reviewer's comment, this is not to say that a lower copy is irrelevant to all phenomena, for example possible reconstruction.)

[15] As a reviewer notes, Bošković (2007) similarly proposes an ''attraction-free'' account of movement (i.e., one not driven by properties of the target), and Bošković's arguments for a non-feature-checking system can extend to Chomsky 2013. See also Bošković 2008, which provides an analysis of ''freezing'' effects of movement (see Epstein 1992, Rizzi 1997). Under the proposals in Bošković 2008, feature checking has a freezing effect on movement, and it is not clear how these effects can be accounted for in a ''free Merge'' system, where feature checking does not constrain the application of Merge. We leave this question for future research (see Epstein, Kitahara, and Seely 2013 for relevant discussion).

Chomsky (2013) further examines data such as (6a–b).

(6) a. They thought the man was assassinated in which Texas city?
b. *They thought [$_{\alpha}$ in which Texas city [C [$_{TP}$ the man was assassinated *t*]]]?

English allows *wh*-in-situ constructions in relevant contexts, such as quiz shows.[16] But if a *wh*-phrase raises to form the embedded clause α and remains there, as in (6b), minimal search cannot find a label for α, since α is of the form {XP, YP}, and there is no prominent feature (e.g., ϕ-features or Q-feature) shared by X and Y, as just discussed. The fact that α has no label bars (6b) at CI, exactly the right result.

The same pattern can be observed in a multiple *wh*-question context. Consider (7a–b).

(7) a. Who thought the man was assassinated in which Texas city?
b. *Who thought [$_{\alpha}$ in which Texas city [C [$_{TP}$ the man was assassinated *t*]]]?

Here again, if a *wh*-phrase raises to form the embedded clause α and remains there at CI, minimal search cannot find the label of α because α is of the form {XP, YP}, and there is no prominent feature (e.g., ϕ-features or Q-feature) shared by X and Y. The fact that α has no label correctly bars (7b) at CI.

It is important to note that the ''no-label'' situation, observed in (6b) and (7b), arises as a consequence of simplest Merge and third-factor minimal search. Moreover, it accounts for the ''obligatory exit'' from intermediate landing sites in *wh*-movement with no representational projection of labels as in Chomsky 1995a, and with no GB-style representationally defined technicalia like Spec-head/government/m-command relations, coindexing, or an S-Structure level of representation to which specifically syntactic filters (e.g., the Case Filter or the EPP) apply.

In what follows, we argue that such a ''no-label'' situation *also* arises in cases of (phase-internal) successive-cyclic A-movement, just as it does in cases of successive-cyclic Ā-movement, with welcome empirical results that in turn raise the possibility of eliminating the concept of ''phase'' itself.[17]

2 Successive-Cyclic A-Movement: A Selective Review

Chomsky (2000) defined Move as the composite operation that combines Merge and Agree, and he demonstrated that (an economy or third-factor) preference for simpler operations over more complex ones explained relevant phenomena, without appeal to an S-Structure level at which the descriptive Case Filter applies. For example, consider (8a–b).

(8) a. *There is likely [$_{TP}$ a man to be *t* in the room].
b. There is likely [$_{TP}$ *t* to be a man in the room].

[16] See Pires and Taylor 2007 for extensive discussion of the licensing conditions on such *wh*-in-situ cases. As a reviewer points out, related issues are discussed in Vlachos 2012.

[17] Like Chomsky (2013), we leave the status of head movement aside; see also footnote 8.

Each derivation involves the derivational stage [T [be a man in the room]], where T requires that something occupy Spec,T (the so-called EPP).[18] At this point, Merge of *there* in Spec,T and Move of *a man* to Spec,T are (in principle) both available to NS, and (in principle) either could be applied in order to satisfy the EPP. Of course, however, (8a) is ungrammatical. What excludes it? Preference for simpler operations over more complex ones selects Merge of *there* in (what will be embedded) Spec,T over Move of *a man* to (what will be embedded) Spec,T precisely because Merge (of *there*) is simpler than Move consisting of both Merge and Agree (Chomsky 2000); hence, the derivation of (8b) is ''selected'' over that of (8a). Thus emerges the ''last resort'' character of movement, accounting for such so-called ''Merge over Move'' phenomena. Note of course that Merge of *there* is possible only if *there* is an item present in the lexical array (LA) available to Merge. So, the initial choice of LA is crucial, and this device (i.e., the LA) must be part of NS.

Postulation of LAs, however, turned out to be necessary but not sufficient, empirically. Consider (9a–b).[19]

(9) a. There is a possibility [$_{CP}$ that a man will be *t* in the room].
b. A possibility is [$_{CP}$ that there will be a man in the room].

(9a) is now seemingly paradoxical. The derivation of (9a) contains *there* (and recall, inserting *there* is less complex than moving *a man*; hence, it is required), yet Move of *a man* to Spec,T is also possible. But it should not be possible, given that insertion of *there* is a simpler operation filling Spec,T. So somehow, when the embedded CP is being built in (9a), in particular Spec,T, Move can be preferred over Merge—the exact opposite of what was motivated to exclude (8a), namely, ''Merge over Move.'' Thus, the theory of that era confronted a seeming contradiction.

Chomsky (2000) presented a solution to this problem. The basic idea is that structures are built bottom-up, cyclically and—importantly—in chunks called ''phases,'' each associated with its own separate lexical subarray (SA), extracted from the LA. So, in effect, first we form the SA containing all and only the lexical material needed to generate the embedded CP phase, which in the case of (9a) includes *that*, *will*, *be*, *a man*, *in*, *the*, *room.* When *a man* moves, insertion of *there* is not an option since (crucially) *there* is—during the independent generation of this phasal CP—absent from the SA. Thus, movement of *a man* in fact does not violate ''Merge over Move,'' since nothing was available to merge in Spec,T when movement to Spec,T took place. ''Merge over Move'' is thus maintained as a nonstipulated operational preference. It provides an account of (9a), overcoming the apparent contradiction, and with no construction- or language-specific mechanisms. Moreover, it follows in part from nonlinguistic principles of complexity (applying X and Y (here, Merge and Agree) is more complex than applying X (here, Merge) alone); that

[18] We ignore *to* here, since it is not relevant for our main concerns; but see footnote 24.

[19] Examples such as (9a–b) and their importance were, to the best of our (and an *LI* reviewer's) knowledge, independently noted (in personal communications to others) by both Alec Marantz and Juan Romero. For discussion of the possible elimination of ''Merge over Move,'' see Castillo, Drury, and Grohmann 1999, 2009 and Epstein and Seely 1999, 2006.

is, appeal to Universal Grammar–specific axioms is minimized by (what are now called) third-factor principles that are by hypothesis not specific to UG.

In short, given postulation of SAs (extracted from LAs), (9a–b) do not compete. In the derivation of (9a), the (embedded) CP is constructed from SA_i, which contains no expletive; hence, Move of *a man* to Spec,T is the only option available at the derivational stage generating the (embedded) CP. In the derivation of (9b), the (embedded) CP is constructed from SA_j, which contains an expletive; hence, Merge of *there* in Spec,T is selected over Move of *a man* to Spec,T at the corresponding stage, because Merge (of *there* in Spec,T) is simpler than Move (of *a man* to Spec,T).

3 Successive-Cyclic A-Movement: A New Analysis

As we have just shown, Chomsky (2000) provided a principled account of the ''last resort'' character of movement by implementing (10a–d).

(10) a. Postulating Move as the composite operation that combines Merge and Agree
b. Postulating lexical array LA
c. Postulating lexical subarray SA (extracted from LA)
d. Postulating Merge over Move (deducibly)

In this section, we argue that (10a) loses its conceptual and empirical support under simplest Merge. In addition, (10b–d) lose their original empirical motivation if labeling is just minimal search. We argue that under the simplifying assumptions of minimal search, data such as (8a–b) and (9a–b) receive a principled explanation, one that is unified in a key respect with the corresponding cases of Ā-movement like (6b) and (7b), which Chomsky (2013) analyzes as labeling (hence CI) failures. Specifically, by showing that remaining in an intermediate landing site of A-movement also results in a failure of labeling, we provide a unified account at least for the core data presented here.[20]

First, recall simplest Merge, defined in (1) and repeated here.

(1) *Simplest Merge*
$\text{Merge}(\alpha, \beta) \rightarrow \{\alpha, \beta\}$

Under the simplest conception of Merge, Chomsky (2004, 2005) argued that Merge and Move are unified; they are just two possible instances—EM and IM—of $\text{Merge}(\alpha, \beta) \rightarrow \{\alpha, \beta\}$. EM takes α and β, where α originates external to β. IM takes α and β, where α originates internal to β; and given the no-tampering condition, IM necessarily yields two copies of α, one external to β (the landing site) and the other internal to β (the departure site). Chomsky notes that both EM and IM come for free: it would require stipulation to bar either of them. Thus, Move can no longer be expressed as the composite operation that combines (external) Merge and Agree (10a),

[20] Goto (2013) independently points out that the core data presented here result in labeling failure at CI.

since (external) Merge and Move are the very same operation—namely, Merge.[21] Consequently, ''Merge over Move'' (10d) is not maintainable. In effect, (10a) and (10d) are eliminated.

Turning now to (10b–c), let us reconsider the data motivating the postulation of LAs and SAs. First, consider (8a–b), repeated here, where TP is replaced by α (whose categorial status is to be determined by labeling, under minimal search). We take *there* to be a complex SO, not a simple head.[22]

(8) a. *There is likely [$_{\alpha}$ a man to be *t* in the room].
b. There is likely [$_{\alpha}$ *t* to be a man in the room].

If, as in (8a), *a man* raises to form the embedded clause α and remains there at CI, minimal search cannot find the label of α, because α is of the form {XP, YP}, and there is no prominent feature (e.g., ϕ-features or Q-feature) shared by X and Y.[23] The fact that α has no label bars (8a) by Full Interpretation at CI. Now if *there* is inserted to form the embedded clause α, as in (8b), and *there* is in fact an XP (perhaps of the form {D, pro}; see Uriagereka 1988), then α is (once again) a label-less set {XP, YP}, where crucially *there* and infinitival raising T do not share ϕ-features.

Thus, *there* cannot remain in this position if the derived representation is to survive Full Interpretation at CI. If *there* bears at least one ϕ-feature and undergoes further movement, then minimal search finds the only visible head T as the label of α, and a completely labeled, hence Full-Interpretation-compliant, CI representation is generated; that is, the ''no-label'' situation does not arise in the final complete representation (8b).[24] Notice that this analysis also predicts that the ''no-label'' situation arises if *there* is merged to form the embedded clause α and it stays there at CI, as in **There is likely there to be a man in the room* or **It seems there to be a man in the room* (see Lasnik 1992).

As shown above, the contrast exhibited by (8a) and (8b) now arguably follows with no reference to the notions LA (10b), SA (10c), or ''Merge over Move'' (10d). Given this ''labeling'' analysis, let us now consider (9a–b), repeated here.

(9) a. There is a possibility [$_{CP}$ that a man will be *t* in the room].
b. A possibility is [$_{CP}$ that there will be a man in the room].

[21] In addition, composite operations are in general unwelcome as we search for the primitive, minimal, undecomposable operations of NS.

[22] See Chomsky 2013:46 for relevant discussion of the complex status of what appears to be a head.

[23] Note that there are languages (such as Modern Greek and Portuguese) where T lacking the tense feature bears ϕ-features; the consequences of this for labeling await further research.

[24] Notice that free Merge in fact allows one-fell-swoop A-movement, a type of derivation that Epstein and Seely (2006) propose—and that they in fact assume is forced given computationally efficient satisfaction of the interfaces (Strong Minimalist Thesis). However, unlike Epstein and Seely (2006), Chomsky (2008:153) assumes there is phase-internal successive-cyclic A-movement. Although the facts are not entirely clear to us (see Epstein and Seely 2006 for discussion), allowing phase-internal successive-cyclic A-movement may be preferable to the one-fell-swoop analysis in Epstein and Seely 2006, with respect to capturing reconstruction (e.g., anaphoric binding) effects in so-called A-chains (e.g., *John seems to Mary* ___ *to seem to himself* ___ *to be* ___ *happy*). Interestingly, the free-Merge analysis assumed here would allow both one-fell-swoop A-movement and successive-cyclic A-movement. For further discussion, specifically about reconstruction effects in A-chains, see for example Lasnik 1999, Bošković 2002, and references therein.

Suppose we dispense with the notions LA, SA, and ''Merge over Move.'' Then there is no competition between insertion of *there* and movement of *a man*. Each counts as an option available to NS. Thus, (9a–b) are both generable. There is no longer any need to postulate LAs (10b) or SAs (10c) to account for data such as (8a–b) and (9a–b).[25] Thus, it seems we can eliminate (10a–d), repeated here.

(10) a. Postulating Move as the composite operation that combines Merge and Agree
b. Postulating lexical array LA
c. Postulating lexical subarray SA (extracted from LA)
d. Postulating Merge over Move (deducibly)

Summarizing, we suggest that Chomsky's (2013) analysis explains ''obligatory exit'' from intermediate positions, not only in successive-cyclic Ā-movement, but also in (phase-internal) successive-cyclic A-movement. Moreover, we suggest that it does so by employing simplest Merge (with no explicit additional label notation, as in Chomsky 1995a) and third-factor minimal search for labels (labels being arguably a natural requirement necessary for CI interpretation), without any appeal to an S-Structure level, a Spec-head/government/m-command relation, coindexing, or S-Structure filters such as the Case Filter or the EPP.

If this ''labeling'' analysis is on track, then LAs and SAs are eliminated from NS, thereby simplifying NS. But notice now that the elimination of LAs and SAs entails the elimination of the notion ''phase'' itself, where a ''phase'' of the derivation is defined as a syntactic object derived from an SA, extracted from an LA. In the next section, we briefly discuss this issue and conclude our discussion.

4 A Remaining Question: Inducing Strict-Cyclic Derivation without the Concept ''Phase''?

If ''obligatory exit'' from intermediate landing sites of successive-cyclic movement can in fact be accounted for by minimal search label identification, for both A- and Ā-movement, and without appeal to the notion ''phase'' defined in terms of SAs or LAs, what becomes of the motivated cyclic computation induced by phase-based analyses? Is there some other way to induce such cyclic derivation but without defining phases in terms of SAs or LAs, which we conjecture to be eliminable, under the ''labeling'' analysis? In Epstein, Kitahara, and Seely (EKS) 2012b, we suggest one possibility: namely, we seek to deduce the Phase Impenetrability Condition (PIC; i.e.,

[25] Similarly, there is no need to stipulate LAs, SAs, or ''Merge over Move'' to account for the following contrast noted (with syntactically parallel examples) in Chomsky 2000:104:

(i) a. I expected [$_{TP}$ someone to be $t_{someone}$ in the room].
b. *I expected [$_{TP}$ t_I to be someone in the room].

Here too, each derivation involves the stage [T [be someone in the room]], and two options—EM of *I* in Spec,T and IM of *someone* to Spec,T—are (by hypothesis) available to NS. However, if EM of *I* in Spec,T, merging an argument into a non-θ-position, creates an illegitimate CI object, violating Full Interpretation, then the observed contrast naturally follows, and Chomsky's (2000) θ-theoretic condition (i.e., pure Merge in a θ-position is required of (and restricted to) arguments) can be eliminated. See also Epstein, Kitahara, and Seely 2012a for relevant discussion.

the timing of Transfer (and also which SO undergoes Transfer)) from the necessity of eliminating anomalous (derivationally defined) syntactic relations, generated by Chomsky's (2007, 2008) proposed ''countercyclic'' application of IM.

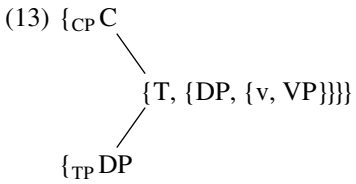
The basic idea of the EKS 2012b approach is this. Suppose we have built up the CP in (11).

(11) {C, {T, {DP, {v, VP}}}}

Following Chomsky (2007, 2008) and Richards (2007), in EKS 2012b we assume that C transmits its (inherently borne) ϕ-features to T, and T then probes for matching ϕ-features of DP. Rejecting countercyclic ''replacement/infixing'' that requires an enrichment of Merge,[26] and maintaining instead simplest Merge, we argue that subject raising to Spec,T is achieved by Merge but yields the following structure, which we first represent in its technical, set-theoretic form:

(12) {DP, {T, {DP, {v, VP}}}}

In purely formal terms, merging DP to {T, {DP, {v, VP}}} embedded in (11) yields the output (12). The relationship between (11) and (12) can be difficult to fully appreciate when examining just the sets. In somewhat informal (though visually more perspicuous) quasi-graph-theoretic terms, what we have is this:

(13) 

Note that in (13) T′ (i.e., the Spec-less T′ = {T, {DP, {v, VP}}}) has both a C sister and a DP sister and thus is immediately dominated by both the CP and the newly created TP projection; that is, the operation Merge creates the new set (12) without destroying any properties of the input set (11).

In EKS 2012b, we proceed to argue on empirical grounds against the Chomsky-Richards timing-of-feature-valuation explanation of Transfer (see Chomsky 2007, 2008, Richards 2007), and we suggest instead that intersecting sets such as (11) and (12), represented informally as in (13), though generable by (simplest) Merge, either halt the NS derivation (since there is no unique accessible root for Merge to apply to) or else constitute insufficient instructions to the CI systems. The situation must be remedied, and the remedy is our explanation for the existence, timing, and target of Transfer. That is, Transfer resolves the illicit intersecting set representation(s) by transferring TP and leaving the ''edge'' of CP for further operations (the alternative—namely, transferring CP and leaving TP—results in nonconvergence).

[26] See Freidin 1999 for independent important arguments against ''replacement'' on the grounds that it is a ''complex'' multifaceted operation not executable given a desirable theory allowing only transformational rules that perform basic primitive operations.

Notice, however, that the EKS 2012b analysis still makes crucial use of LAs/SAs and a form of ''Merge over Move,'' which are the very constructs we are attempting to eliminate here under the conjecture that given the labeling analyses discussed above, these postulates—and the phase-defined derivation incorporating them—are no longer necessary. In the EKS 2012b analysis, C must be merged *before* IM of DP to Spec-less TP (i.e., raising of the Spec of vP to Spec,TP must be ''countercyclic'' IM). But this ordering is a form of ''Merge over Move,''[27] and it is needed in the EKS 2012b system in order to create the ''double peak'' structure (13), which in turn induces the desired application of Transfer.

So the question arises: is it possible to modify the EKS 2012b analysis and still *force* the ''double peak'' structure (and thereby induce Transfer) but without appealing to LAs/SAs and ''Merge over Move''? We believe the answer is yes; in fact, there are at least two ways to derive cyclic Transfer within the EKS 2012b analysis. We briefly sketch these alternative deductions of Transfer below (leaving detailed reanalyses, and their consequences, for future research).

4.1 Deriving the ''Double Peak'' Structure through ''Chain Invisibility''

Suppose LAs/SAs are eliminated entirely; rather than drawing from an LA/SA, NS draws directly from the lexicon.[28] With Chomsky (2007, 2008), assume that there are dedicated phase heads (i.e., that C and v inherently bear ϕ-features). Finally, assume as in the EKS 2012b system that it is the complex of features [ϕ + tense] (for nominative) and [ϕ + transitive property] (for accusative) that checks/values structural Case (under probe-goal agreement).[29] With the above assumptions in mind, suppose a derivation has built up to the Spec-less TP in (14).

(14) {T, {{the, man}, {v, VP}}}

Given free Merge as in Chomsky 2013 (and assuming that there is no ''Merge (EM) over Move (IM)''), we could in principle merge the DP *the man* in Spec,v to the Spec-less TP; that is, we could ''raise'' the subject from Spec,v to Spec,T, *before* merging C, yielding (15).

(15) {{the, man}, {T, {{the, man}, {v, VP}}}}

Next, we could merge C, yielding (16).

(16) {C, {{the, man}, {T, {{the, man}, {v, VP}}}}}

[27] Actually, in EKS 2012b we attempt to deduce ''Merge over Move'' from efficiency considerations; the idea is that Merge involving no search of merged items is preferred over any application of Merge involving search of merged items. Since EM involves no search, all instances of EM must be exhausted (relative to a given LA/SA) before any instance of IM.

[28] Note that there are both empirical and conceptual arguments for LAs/SAs. Above, we have conjectured that the empirical evidence for LAs/SAs can be accounted for without LAs/SAs. But we do not address the conceptual argument in favor of LAs/SAs, based on efficiency considerations.

[29] In the EKS 2012b analysis, these feature complexes (e.g., [ϕ + tense]) do not exist on any category in the lexicon; rather, they are created in the course of a derivation via feature transmission. We focus here just on nominative Case.

Note that in the creation of (16), with (15) as input, no ''countercyclic'' operations have applied; specifically, there was no ''countercyclic'' merger to Spec,TP (this Spec,TP had already been created before the merger of C).

At the derivational point represented by (16), the ϕ-features of C must be transmitted to T; recall that in the EKS 2012b system, Case of *the man* is valued by the feature complex [ϕ + tense]. Since C inherently bears ϕ-features, and since T inherently bears the tense feature, the required feature complex must be created derivationally via feature transmission from C to T.[30]

After feature transmission, T bears the Case-valuing features [ϕ + tense] and hence T can probe (i.e., minimally search for the first matching goal). Now, at first glance, it would appear that Case/agreement feature valuation (between T and *the man*) could take place, and since there was no ''countercyclic'' movement to Spec,T, there is no ''double peak''—with the undesired result that there now is no required Transfer, triggered in order to rectify the ''double peak.'' In other words, it would appear at first glance that we have just lost the ''double peak'' deduction of cyclic Transfer. However, under Chomsky's (2013) conjecture, it follows that in fact, T (now bearing [ϕ + tense]) cannot ''see'' or successfully probe *the man*, since *the man* has moved and its lower copy is invisible to T and hence is not Case-valuable by T; technically, *the man* is no longer in the minimal search ''checking'' domain of T since there is an occurrence of *the man* (in Spec,T) that is not in the minimal search domain of T. Stated informally, T does not c-command all occurrences of *the man* (see Chomsky 1995b for the definition of *occurrences*) and hence T cannot Case-value *the man*; or, in other words, T does not c-command *the man* since T does not c-command the entire ''chain'' of *the man*. Notice further that C cannot Case-value *the man* since C (after feature transmission) does not have all the features necessary for Case checking (namely, ϕ-features *and* the tense feature).[31]

Overall, then, this (noncounter)cyclic derivation, where the subject DP is raised to Spec,T *before* C is merged, fails to yield a convergent derivation. If this analysis is viable, then we do not need to appeal to ''Merge over Move'' in order to induce the ''countercyclic,'' double-peak-forming, Transfer-inducing derivation we desire.

The desired, convergent, ''countercyclic'' derivation is as follows. It starts with (14), repeated here.

(14) {T, {{the, man}, {v, VP}}}

Rather than ''raise'' the subject *the man* to Spec,T before merging C (a derivation that crashes, as we've just shown), suppose we first merge C to Spec-less TP (i.e., merge C to (14)), yielding (17).

[30] We assume that the ϕ-features of C move down to T to create [ϕ + tense] (downward feature transmission of ϕ-features from C to T); the alternative is for the tense feature of T to move up to C, which also creates [ϕ + tense] (upward feature transmission of the tense feature from T to C). As we will show in a moment, the latter feature transmission option must be prohibited.

[31] Note that if the tense feature of T were to ''raise'' to C, then C (now bearing the feature complex [ϕ + tense]) could in principle Case-value *the man* since *the man* is in fact in the domain of C (all occurrences of *the man* are c-commanded by C). It is crucial, then, that feature transmission is downward.

(17) {C, {T, {{the, man}, {v, VP}}}}

Suppose next that there is ϕ-feature transmission from C to T. T now has all features necessary for nominative Case checking. Thus, T probes, and finds the local DP *the man*. And in fact *the man* is in T's domain since all occurrences of *the man* (there is only one) are c-commanded by T. Feature checking/valuation can occur. But the derivation is not done yet. Assuming there is a need to create a Spec,TP position (see our earlier discussion of Chomsky's (2013) analysis of the EPP),[32] the DP *the man* must raise to Spec,TP. But note that such raising of *the man* is now necessarily ''countercyclic.'' And it is precisely such ''countercyclic'' raising that will create the ''double peak'' structure (13) that in turn induces cyclic Transfer. This is just the desired result. So, we have shown that there is an independently motivated modification of the EKS 2012b system that will result in the ''double peak'' deduction of cyclic Spell-Out (PIC) without any appeal to LAs/SAs or to any form of (the currently unstatable) ''Merge over Move.''

4.2 Another Deduction of Cyclic Transfer without Appeal to LAs/SAs or to ''Merge over Move''

There is an alternative deduction of cyclic Transfer requiring neither LAs/SAs nor ''Merge over Move.'' We outline this alternative below and briefly explore implementations of it.

Suppose we have (18); that is, we have built TP and have not yet merged C.

(18) {T, {DP, {v, VP}}}

Next, suppose that DP raises to Spec,T (given the EPP and specifically Chomsky's (2013) account of the EPP effects based on the need to create a labeled syntactic object; but see footnote 32). That is, suppose we raise DP to Spec,T before we merge C. Such raising will produce (19).

(19) {DP, {T, {DP, {v, VP}}}}

Interestingly, for Chomsky (2007, 2008), at the very point that DP merges via IM to Spec-less TP, Transfer applies; that is, there is cyclic Transfer. Why? Chomsky's argument (as we understand it) is as follows: If Transfer waits until later in the derivation to apply, then it can't be determined

[32] As we showed earlier, it follows from Chomsky 2013 that Spec,v must raise from vP (since vP has no label without such raising). Interestingly, the landing site of the raised DP does not follow from Chomsky's account; that is, it does not follow that DP must raise *to* Spec,T, but only that DP must raise *from* Spec,v (so that vP can become labeled). Perhaps relevant to this issue, and supporting this analysis, is the derivation of sentences like *How many men do you think that there are* ____ *outside?* (adapted from Chomsky 1991:445), in which it appears that the moved DP has exited Spec,v but has not moved to Spec,T of the embedded clause. Chomsky's (2013) analysis predicts that the subject ultimately lands where it will not induce a labeling failure. As a result, Spec,T is fine, but Spec,C is problematic if C lacks ϕ-features. The subject can move to a potentially problematic place, as in *wh*-movement to intermediate C, as long as it moves on further. Under free Merge, IM can move the subject anywhere in principle, but if a labeling failure occurs, then the CI representation pays the price. So, labeling indirectly determines the departure site as well as the landing site of the subject. We leave the details of such derivations to further research (see EKS 2013 for relevant discussion).

whether (19) involves two separate DPs (as in, say, *He thinks he is smart*, where there are two separate DPs *he*) or just one DP (as in, say, *He was arrested t(he)*, where there is exactly one DP *he*).[33] There is a critical identity issue that the derivation must keep track of. Only simultaneously with DP-raising can Transfer ''see'' that a single DP was the input to, and output of, Merge and hence that it is a single DP, as opposed to two separate DPs. We assume with Epstein and Seely (2002) that Transfer can see the structural description and the structural change of (i.e., the input to and the output of) one and only one rule application; this much lookback is necessary and, by hypothesis, sufficient. Assuming then that Transfer can look back *only* from output to input of a single rule, once the next instance of Merge applies to (19), it is too late for Transfer to ''know'' that there was just one DP; in short, Transfer must apply internal to the application of Merge that produces (19). But transferring in (19) is problematic: it requires TP transfer immediately after TP creation by IM, incorrectly predicting that no C projections are ever created by Merge of C with a TP that was created by subject raising, under the assumption that Transfer renders an object inaccessible to Merge. But if Transfer waits until after Merge(DP, Spec-less TP), then Transfer will not have the information it needs to tell if one DP or two DPs are present, under Chomsky's proposal.[34] Thus, the cyclic derivation (i.e., build TP; IM Spec,v to Spec,T; merge C; transfer TP) is not generable.

What about the ''countercyclic'' derivation? That is, suppose C is merged before DP raises to Spec,T. Recall (11).

(11) {C, {T, {DP, {v, VP}}}}

Next, suppose DP raises to Spec,T. Recall (13).

[33] It thus appears to us that Chomsky's (2007, 2008) analysis assumes the following: (a) a pair ⟨X, X⟩ is identified as copies created by IM, only immediately after IM; and (b) if ⟨X, X⟩ is not identified as two copies created by IM, then it must be identified as two separate arguments. A less stipulative algorithm would be this: freely assume ⟨X, X⟩ to be two copies or assume them to be two separate arguments. This would arguably suffice empirically; that is, if *He was arrested t(he)* is taken to contain two arguments, there is a θ/Full Interpretation/gibberish problem at CI. If *He thinks he is smart* is taken to involve two copies created by IM, then again, there is a θ/Full Interpretation/gibberish problem at CI. This alternative, ''free interpretation'' analysis seems worthy of detailed exploration but lies beyond the scope of the present article.

It should be noted that the identity problem is much more general. It goes beyond arguments: for example, in *The boy likes the girl*, it arises concerning *the.* Chomsky (2007, 2008) assumes the ''default'' is to treat the two identical elements as repetitions (two separate morphemes), and then he adds exceptions: if they are formed by IM, then they are not repetitions, but copies (of a single entity).

[34] This is similar to the Chomsky-Richards analysis seeking to explain cyclic Spell-Out by appeal to the nature of feature valuation. Chomsky and Richards argue that spelling out a feature before valuation causes a crash (unvalued features appear at the interface). But spelling out a feature after that feature is valued is too late, since Transfer cannot tell whether the feature was inherently lexically valued (in which case it is interpretable and so is not to be transferred) or underwent syntactic valuation (in which case it is uninterpretable and must therefore be spelled out in order to avoid a crash). The solution (Epstein and Seely 2002) is to allow one-operation lookback; that is, Transfer can see the input to and output of a feature valuation operation and so it can tell whether a valued feature underwent syntactic valuation or was inherently valued. The analysis here is similar: only internal to IM can it be determined that ⟨X, X⟩ is two copies. Thereafter, it's impossible to determine if it consists of one DP or two DPs since a ''chain'' of DPs is indistinguishable from two separate DPs. See Epstein and Seely 2002 for detailed discussion.

(13) {$_{CP}$ C
{T, {DP, {v, VP}}}}
{$_{TP}$ DP

Notice that under the EKS 2012b ''double peak'' analysis, cyclic transfer of TP is in effect required/forced (as this analysis seeks to deduce the PIC). However, interestingly, under Chomsky's (2013) identity approach just considered, under which, if IM applies, it must be the last operation in a phase, Transfer also takes place obligatorily at this point—immediately after IM—in order to ''keep track of identity.'' But notice that now there is a redundancy in that ''double peak'' and ''identity'' each force cyclic TP Transfer. That is, assuming the ''identity'' analysis renders the ''double peak'' analysis of transfer unnecessary.[35]

To sum up this exploratory section, we conjecture that given the extension of Chomsky's (2013) ''labeling'' analysis to A-chains, as explored here, there may no longer be any need to appeal to the standard notion of ''phase'' defined in terms of LAs/SAs and ''Merge over Move.''[36] In fact, if Chomsky (2013) is right that the only operation is unified simplest Merge, it is impossible to state a principle of the form ''Merge (EM) over Move (IM).'' We have explored two approaches here, which, if they are on the right track, allow us to deduce the timing of Transfer, as well as the category undergoing Transfer, but without independently stipulating either the PIC or which categories undergo Transfer.

References

Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a postsyntactic operation. In *Phi theory: Phi-features across modules and interfaces*, ed. by Daniel Harbour, David Adger, and Susana Béjar, 295–328. Oxford: Oxford University Press.

Boeckx, Cedric. 2006. *Linguistic Minimalism: Origins, concepts, methods, and aims.* Oxford: Oxford University Press.

Boeckx, Cedric. 2010. A tale of two Minimalisms: Reflections on the plausibility of crash-proof syntax, and its free-Merge alternative. In *Exploring crash-proof grammars*, ed. by Michael T. Putnam, 89–104. Amsterdam: John Benjamins.

[35] Chomsky's (2013) ''identity'' analysis of Transfer has another interesting consequence: it follows from the analysis that feature inheritance operates downward. We noted in section 4.1 that the ϕ-features of C (and analogously the ϕ-features of v) are transmitted to T (and from v to V) and that it is the complex of features [ϕ + tense] that checks (nominative) Case. We noted further (see footnote 31) that it is crucial that the tense feature of T *not raise* to C (to create the [ϕ + tense] complex). If the tense feature *is* transmitted up to C bearing ϕ-features, then a possible landing site for Subj is Spec,C, not Spec,T, because it is C that has ϕ-features (for labeling). So, if Subj has a reason to move out of Spec,v, it will end up in Spec,C. Now, given the ''identity'' analysis, the newly created CP containing Subj in its Spec must undergo Transfer. As a result, there is no way to continue this derivation. Chomsky's (2013) analysis therefore provides a ''labeling + identity'' account for why features must be inherited downward.

[36] This conclusion is reached on independent grounds by Frampton and Gutmann (2002), as well as by Epstein and Seely (2002, 2006).

Bošković, Željko. 1994. D-Structure, Theta-Criterion, and movement into theta-positions. *Linguistic Analysis* 24:247–286.

Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167–218.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.

Bošković, Željko. 2008. On the operator freezing effect. *Natural Language and Linguistic Theory* 26: 249–287.

Castillo, Juan Carlos, John Drury, and Kleanthes K. Grohmann. 1999. Merge over Move and the Extended Projection Principle. In *University of Maryland working papers in linguistics 8*, ed. by Sachiko Aoshima, John Drury, and Tuomo Neuvonen, 63–103. College Park: University of Maryland, Department of Linguistics.

Castillo, Juan Carlos, John Drury, and Kleanthes K. Grohmann. 2009. Merge over Move and the Extended Projection Principle: MOM and the EPP revisited. *Iberia: An International Journal of Theoretical Linguistics* 1:53–114.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and parameters in comparative grammar*, ed. by Robert Freidin, 417–454. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995a. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program: Principles and parameters in syntactic theory*, ed. by Gert Webelhuth, 385–439. Oxford: Blackwell.

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

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures*, vol. 3, ed. by Adriana Belletti, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36:1–22.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language?*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 42–64. Oxford: Blackwell.

Epstein, Samuel David. 1992. Derivational constraints on Ā-chain formation. *Linguistic Inquiry* 23:235–259.

Epstein, Samuel David. 2007. On i(internalist)-functional explanation in Minimalism. *Linguistic Analysis* 33:20–53.

Epstein, Samuel David. 2013. Syntactic representations reduced to 3rd factor primitives of set theory? Ms., University of Michigan, Ann Arbor.

Epstein, Samuel David, Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations.* Oxford: Oxford University Press.

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. 2012a. Exploring phase based implications regarding clausal architecture. A case study: Why structural Case cannot precede theta. In *Phases: Developing the framework*, ed. by Ángel J. Gallego, 103–124. Berlin: Mouton de Gruyter.

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. 2012b. Structure building that can't be! In *Ways of structure building*, ed. by Myriam Uribe-Etxebarria and Vidal Valmala, 253–270. Oxford: Oxford University Press.

Epstein, Samuel David, Hisatsugu Kitahara, and T. Daniel Seely. 2013. Criterial freezing: Transformational cessation vs. CI interpretation. Ms., University of Michigan, Ann Arbor, Keio University, and Eastern Michigan University, Ypsilanti.

Epstein, Samuel David, Acrisio Pires, and T. Daniel Seely. 2005. EPP in T: More controversial subjects. *Syntax* 8:65–80.

Epstein, Samuel David, and T. Daniel Seely. 1999. SPEC,ifying the GF ''subject'': Eliminating A-chains and the EPP within a derivational model. Ms., University of Michigan, Ann Arbor, and Eastern Michigan University, Ypsilanti.

Epstein, Samuel David, and T. Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 65–90. Oxford: Blackwell.

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

Frampton, John, and Sam Gutmann. 2002. Crash proof syntax. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 90–105. Oxford: Blackwell.

Freidin, Robert. 1999. Cyclicity and Minimalism. In *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 95–126. Cambridge, MA: MIT Press.

Goto, Nobu. 2013. Labeling and scrambling in Japanese. *Tohoku: Essays and Studies in English Language and Literature* 46:39–73.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30:69–96.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar.* Cambridge: Cambridge University Press.

Jayaseelan, J. K. 2008. Bare phrase structure and specifier-less syntax. *Biolinguistics* 2:87–106.

Jónsson, Jóhannes Gísli. 1996. Clausal architecture and case in Icelandic. Doctoral dissertation, University of Massachusetts, Amherst.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations.* Cambridge, MA: MIT Press.

Lasnik, Howard. 1992. Case and expletives: Notes toward a parametric account. *Linguistic Inquiry* 23: 381–405.

Lasnik, Howard. 1999. Chains of arguments. In *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 189–215. Cambridge, MA: MIT Press.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235–289.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α: Conditions on its application and output.* Cambridge, MA: MIT Press.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32:303–344.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement.* Cambridge, MA: MIT Press.

O'Neil, John. 1997. Means of control: Deriving the properties of PRO in the Minimalist Program. Doctoral dissertation, Harvard University, Cambridge, MA.

Ott, Dennis. 2010. Grammaticality, interfaces, and UG. In *Exploring crash-proof grammars*, ed. by Michael T. Putnam, 89–104. Amsterdam: John Benjamins.

Pires, Acrisio, and Heather L. Taylor. 2007. The syntax of *wh*-in-situ and common ground. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society 43*, 2:201–215. Chicago: University of Chicago, Chicago Linguistic Society.

Putnam, Michael T., ed. 2010. *Exploring crash-proof grammars.* Amsterdam: John Benjamins.

Richards, Marc D. 2007. On feature inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38:563–572.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar: Handbook of generative syntax*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, Agree, or Merge? *Nanzan Linguistics* 8:109–127.
Saito, Mamoru. 2013a. Case and labeling in a language without phi-feature agreement. Ms., Nanzan University.
Saito, Mamoru. 2013b. Conditions on Japanese phrase structure: From morphology to pragmatics. *Nanzan Linguistics* 9:119–145.
Seely, T. Daniel. 2006. Merge, derivational c-command, and subcategorization in a label-free syntax. In *Minimalist essays*, ed. by Cedric Boeckx, 182–217. Amsterdam: John Benjamins.
Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut, Storrs.
Vlachos, Christos. 2012. *Wh*-constructions and the division of labour between syntax and the interfaces. Doctoral dissertation, University of Patras.

(Epstein)
University of Michigan
Department of Linguistics
440 Lorch Hall
611 Tappan Street
Ann Arbor, MI 48109-1220

sepstein@umich.edu

(Kitahara)
Institute of Cultural and Linguistic Studies
Keio University
2-15-45 Mita, Minato-ku
Tokyo 108-8345
Japan

kitahara@icl.keio.ac.jp

(Seely)
Department of English
Program in Linguistics
612-D Pray-Harrold
Eastern Michigan University
Ypsilanti, MI 48197

tseely@emich.edu