1. Introduction

Let us consider (1) in terms of Chomsky’s (2013; hereafter, POP) labeling algorithm (LA):

(1) [γ There is likely [β to be [α a man in the room]]]

In (1), α is of the form {XP, YP}, constituting an associate-predicate structure. Under LA, such a syntactic object (SO) can be labeled either by raising one of the two elements or by sharing the most prominent features of the two elements (POP), but given that α is not in such a configuration, it is not clear how α is labeled. Likewise, β is of the form {H, XP}, constituting an infinitival T-structure. Under LA, such an SO is labeled by minimal search to H (POP: 43), but given that “T is too ‘weak’ to serve as a label” (Chomsky 2015; hereafter POP+), it is not clear how β is labeled. Furthermore, γ is of the form {there, TP}, constituting an expletive-predicate structure. Under LA, the subject-predicate construction {NP, TP} is labeled by sharing ϕ-features of the two elements and creating a criterial configuration under so-called “SPEC-Head agreement,” “a stronger relation,” or “actual agreement” (POP), but given the fact that the expletive there, in contrast with other full-fledged NPs, cannot enter into an ordinary agreement relation with T, it is not clear how γ is labeled.

How should we interpret this situation? Do all SOs need to be labeled? Or, don’t all SOs need to be labeled? With respect to these issues, there are, of course, some possible approaches, but in this paper, taking seriously “the general principle that all SOs that reach the interfaces must be labeled” (POP), I would like to consider how SOs in there-constructions are labeled to satisfy the general principle. More specifically, it will be shown that not only the theoretical problems of labeling raised in (1) but also a range of interesting phenomena that remain unclear under Epstein, Kitahara and Seely’s (EKS) (2014) labeling analysis of Merge-over-Move under LA are explained by the interaction of Chomsky’s POP+ labeling theory with Abe’s (to appear) Move-approach to there-constructions.

This paper is organized as follows: Section 2 summarizes EKS’s labeling analysis of Merge-over-Move, Chomsky’s POP+ labeling theory, and Abe’s (to appear) Move-approach to there-constructions. Section 3 demonstrates that the problems of labeling mentioned above are resolved by integrating the Move-approach with labeling theory. Section 4 provides a unified explanation for complicated properties of the constructions in terms of labeling. Section 5 concludes the paper.

2. Labeling Theory

2.1. Chomsky (2013) and EKS (2014)

In POP, assuming that “a label is required for interpretation at the interfaces, and that labels are assigned by a minimal search algorithm LA applying to an SO”, Chomsky (2013) argues that such SOs as (2a), (2b), and (2c) are visible to LA, get labeled H, Y, and F, respectively, and converge at the interfaces. But such an
SO as (2d) is invisible to LA, remains unlabeled, and crashes at the interfaces:

(2)  a.  [H, XP] = H
    b.  [XP, · · · [t, YP]] → [t, YP] = Y
    c.  [XP_t, YP_t] = F
        (F, a feature shared by XP and YP)
    d.  [XP, YP] = ?

Under this LA, EKS (2014) proposes an interesting analysis of the contrast between (3a) and (3b), which was accounted for by a UG-stipulative notion called Merge-over-Move (Chomsky 1995, 2000):

(3)  a.  There is likely [a man in the room]
       (= (1))
    b.  *There is likely [a man] [to be a man in the room]

According to EKS, α in (3a) is visible to LA, gets labeled (T), so it converges at the interfaces, but α in (3b) is invisible to LA, remains unlabeled, and hence crashes at the interfaces. This labeling analysis of the contrast is theoretically welcome because it simplifies UG by eliminating the stipulation on Merge (see also Goto 2013 for the same argument).

This analysis is, however, not free from problems. As pointed out in (1), the theoretical problems of labeling remain even with the convergent derivation (3a). Also, crucially, note that if the crashing derivation (3b) proceeds, forming β by raising the associate a man to an intermediate position, as follows:

(4)  There is [a man] likely [a man to be a man in the room]

then the derivation converges. In EKS’s view, it is not clear how β in (4) is labeled. According to their analysis, α in (4) will be labeled (T) in conjunction with the raising strategy (2b) (in the same way as (3b)), but it is not clear how β in (4) is labeled, because it does not fall under any of the labelable configurations listed in (2a)–(2c). Unfortunately, the same problem is expected to occur in examples like the following, particularly around α:

(5)  a.  There is likely to be [a building demolished]
      (cf. (3a))
    b.  There is [a building likely to be demolished]
      (cf. (4)) (Lasnik 1995)

(6)  We proved [there to be a thief among us]
      (cf. (3b)) (Abe to appear)

For the same reason as above, it is not clear how α in (5) and (6) is labeled under EKS’s labeling analysis of Merge-over-Move.

Thus, as a first approximation, EKS’s approach to (3) seems to be on the right track in terms of LA. To cover all the relevant constructions, however, we need a more refined theory of labeling and a more precise analysis of there-constructions.

2.2. Chomsky (2015)

In POP+, assuming the general framework of POP, Chomsky (2015) makes further refinements with respect to how LA is computed in phase theory (among others, see Chomsky 2007, 2008, for essential ideas on phase theory). The relevant ideas about labeling developed in POP+ are quoted in (7):

(7)  a.  “Operations can be free, with the outcome evaluated at the phase level for transfer and interpretation at the interfaces.”
    b.  “The basic principle is that memory is phase-level.”
    c.  “T is similar to roots: T is too ‘weak’ to serve as a label” and “with overt subject, the SPEC-TP construction is labeled ⟨ϕ, ϕ⟩ by the agreeing features.”
    d.  “In terms of labeling theory, Italian T, with rich agreement, can label TP and also {SPEC, TP}.”
    e.  “Since R is universally too weak to label, it follows that the analogue of EPP for v*P holds for both English and Italian, but here EPP refers

2 In fact, Chomsky (2013) argues that the XP-YP structure itself is visible to LA, but it is structurally ambiguous, so that the label of the SO is not determined unambiguously. Since the notion of ambiguity is irrelevant to the following discussion, I will ignore the detailed implementation of the process, ensuring the unlabeability of the SO by assuming that it is invisible to LA (see Goto (2016) for relevant discussion).
to raising-to-object. Just as English T can label TP after strengthening by SPEC-T, so R can label RP after object-raising."
f. “[P]hasehood is inherited by T along with all other inflectional/functional properties of C (ϕ-features, tense, Q), and is activated on T when C is deleted” and “we preserve the computational simplifications of PIC.”
g. “R raises to v* forming R with v* affixed, hence invisible, so phasehood is activated on the copy of R, and DP (which can be a wh-phrase) remains in situ, at the edge.”

With these assumptions, Chomsky (2015) argues that the Extended Projection Principle (EPP) and the Empty Category Principle (ECP) can be unified under the labeling theory (see his paper for the details).

In the following, I will show that the labeling theory developed in POP+ can pave the way for providing a solution to the theoretical problems of labeling in the there-constructions mentioned above, significantly in collaboration with Abe’s (to appear) Move-approach.

2.3. Abe (to appear)

Based on a careful examination of the basic issues of there-constructions, i.e., how an associate of the expletive there obtains Case, and how T and the associate establish an agreement relation, Abe proposes an interesting analysis for the derivation of there-constructions in terms of Chomsky’s (2008) phase theory, which incorporates the feature-inheritance-based probe-goal system that allows parallel derivation upon the introduction of the phase heads C and v* into the derivation. Abe’s core assumptions and proposals are summarized in (8), with the relevant structure given in (9) (his (57)).


b. The partitive Case of the associate (Belletti 1988) is checked by V with ϕ-features inherited from v*, which constitutes the strong phase v*P that is subject to the PIC (Chomsky 2000, 2001).

c. The expletive there has unvalued ϕ-features that may or may not be valued by its associate (Abe’s core proposal).³

(9) (Abe’s (57))⁴

About this structure, Abe claims: “the D-N agreement is optional here, so that the value of the ϕ-features of there can be [iϕ] or remain [uϕ]” (where [iϕ] indicates when the expletive there shares ϕ-values of the associate).

Under these assumptions, Abe proposes to derive (10) as follows:⁵

(10) There is someone in the garden.

a. [DP there someone]

b. [ψ- v* [v’ be [SC [DP there someone] in the garden]]]

c. [C: C [v’ T [vP there [ψ- v* + be [VP someone [v’ hbe [SC [DP tthere tsomeone] in the garden]]]]]]]

d. [CP [TP there [T’ T [vP tthere [ψ- v* + be [VP someone [v’ hbe [SC [DP tthere tsomeone] in the garden]]]]]]]

³ Following Pesetsky and Torrego’s (2007) characterization of formal features, Abe assumes that only unvalued features can act as probes and proposes that the expletive there carries both uninterpretable and unvalued ϕ-features. In this paper, I will ignore the uninterpretability of there, for ease of application. That said, the argument that is developed below is intact even if we take it into consideration.

⁴ A reviewer kindly reminds me of other interesting properties of there-constructions: the appearance of definite NPs in a certain context (cf. Rando and Napoli 1978), and binding/scope relations between associates and locatives (cf. Takano 1996, Kuno 1973). Since it is beyond the scope of this paper to discuss how these properties can be accommodated by the present approach to there-constructions, I leave them for future research.

⁵ For convenience, the following abbreviations are used throughout: EM(X, Y) = external Merge applies to X and Y, yielding the set {[X, Y] (order irrelevant); IM(X, Y) = internal Merge applies to X and Y, merging X to SPEC of Y (where X is structurally higher than Y); Agree(X, Y) = Agree applies between X and Y, valuing unvalued features of X and Y, such as ϕ-features and a Case-feature (where X is structurally higher than Y); and the shaded part is a Transfer domain.
First, EM(there, someone) applies to yield (10a). At this stage, the optional \( \psi \)-feature sharing may take place. Suppose here that it takes place and the expletive \( \text{there} \) shares the \( \psi \)-values of the associate \( \text{someone} \). After that, iterated EMs apply to yield (10b). At this \( \text{v}^*\text{P} \) phase, the following operations apply in parallel to yield (10c): \( \text{v}^* \)-to-\( \text{be} \) \( \psi \)-feature-inheritance, Agree (be, someone), IM (someone, be), IM(there, \( \text{v}^* \)), V-to-\( \text{v}^* \)-raising, and VP-Transfer. Then, at the CP phase (10d), the following operations apply in parallel to complete the derivation: C-to-T \( \psi \)-feature-inheritance, Agree(T, there) and IM (there, T). At this final stage, all the relevant unvalued features of the involved elements are valued, and the expletive \( \text{there} \) has the \( \psi \)-values of the associate \( \text{someone} \), and hence the fact that the matrix T agrees with the associate follows from the indirect agreement relation with the expletive \( \text{there} \) that has raised to the matrix SPEC-T, passing through the embedded SPEC-\( \psi \-star. That is why Abe calls this the Move-approach.6

For the present purposes, it is interesting to notice that the Move-approach can independently account for the contrast in (3), for which EKS provide the labeling analysis. On the Move-approach, the contrast is accounted for in terms of the PIC: as shown in (10), the SPEC-\( \psi \)-star that takes the expletive \( \text{there} \) remains accessible to the next phase, but the SPEC-V that takes the associate \text{a man} becomes inaccessible to the next phase; hence (3a) is derived but (3b) is not. In this way, the Move-approach can also account for the contrast in (3).

However, it may be worth pointing out that the PIC-based analysis of (3) can be safely abandoned in favor of labeling theory. Since “phasehood is activated on the copy of R” (\( (7g) \)), it follows that although the complement of V, i.e., the SC domain in (10) (“R-complement” in Chomsky’s terms) becomes inaccessible to the next phase, the SPEC-V that takes the associate \text{a man} becomes inaccessible to the next phase; hence, we cannot accept the PIC-based analysis of (3) as it is in labeling theory. In the next section, reinterpreting Abe’s Move-approach in terms of Chomsky’s POP+ labeling theory, I would like to show that the problems of labeling that remain unclear in EKS’s views are resolved and the contrast in (3) receives a principled account.

3. Analyses

Now let us first examine how (1) is labeled. On the present assumptions, we have the following derivation at the lower \( \text{v}^*\text{P} \) phase:

\[
\begin{align*}
(11) & \; [\text{there} \; [\text{be} + \text{v}^* [\text{a man}_\psi]]_{\text{T}_\text{Rh}} \; [\alpha \; [\text{there}_\psi \; \text{a man}]]_{\text{[in the room]}]}]
\end{align*}
\]

Here, EM(there, \text{a man}) and EM([there, a man], [in the room]) are applied to yield \( \alpha \). This SO is initially invisible to LA, but will be modified by parallel IMs of (there, \( \text{v}^* \)) and (a man, R), so it eventually becomes visible to LA, as in \( [t, [\text{in the room}]] \), and will be labeled, as required. Notice that IM(a man, R) also appears to yield an SO that is initially invisible to LA, but R can label RP after associate-raising; therefore it also becomes visible to LA, as in \( [\phi, [\beta, [\phi]]] \), and will be labeled, as required. Also, IM(there, \( \text{v}^* \)) appears to yield an invisible SO in the SPEC-\( \psi \)-star, but the SO undergoes further applications of IM of \text{there}, as shown in (12) below, so that it also becomes visible to LA, as in \( [\text{there}, \text{v}^*\text{P}] \), and will be labeled, as required. At this stage, all the relevant SOs are labelable; hence, even if phasehood is activated on the lower copy of R, Transfer of R-complement (\( \alpha \)) leads to the convergence of the derivation.

Next, at the intermediate position, we have the following derivation:

\[
\begin{align*}
(12) & \; [\text{likely} \; [\beta \; \text{there} \; [T(\text{to})]_{\text{there}} \; [\text{be} + \text{v}^* [\text{a man}_\psi]]_{\text{TP} \; [\text{R}]_{\beta \; [\cdot \cdot \cdot ]}}]]]
\end{align*}
\]

Here, EM(T, \( \text{v}^*\text{P} \)) applies to yield a \( [H, \text{XP}] \) structure. Since this SO is visible to LA, it is straightforwardly labeled by minimal search to T. Then, given that “IM is successive-cyclic leading to a criterial position, and is forced to ensure labeling” (POP+), it is expected that the expletive \( \text{there} \) should raise from the lower SPEC-\( \psi \)-star to the intermediate SPEC-T to yield \( \beta \). As discussed just above, although such an SO is initially invisible to LA, it is forced to undergo further applications of IM of \text{there} for labeling at the matrix SPEC-T, as shown in (13) below; hence, it also becomes visible to LA, as in \( [\text{there}, \text{TP}] \), and will be labeled, as required. Furthermore, as the

---

6 In the above, we have illustrated the case in which \( \psi \)-feature sharing takes place. But given (8c), it may not take place and the \( \psi \)-features of \( \text{there} \) may remain unvalued. For empirical support of this claim, see Abe (to appear).
derivation proceeds, EM (likely(A), TP) applies to yield {H, XP} that is straightforwardly labeled by minimal search to A. Again, at this stage, all the relevant SOs are labelable; hence, even if they undergo Transfer at the next phase, the derivation converges.

Finally, we have the following derivation at the matrix CP phase:

(13) \[ [C \gamma \text{there}_\gamma [\text{be} + v^*_\gamma \text{t}_\gamma] [\text{there}_\gamma \text{T}(\text{to}) [\text{there}_\gamma \text{be} + v^* \{\text{a man}_\beta\} [\text{t}_\gamma \text{R} [\text{there}_\gamma \text{be} + v^* \{\text{a man}_\beta\} [R_\gamma \text{be} \text{R}\text{be}]]]]]]]]

Here, through Agree(T, there), IM(there, T) is applied to yield \( \gamma \), and there raises from the intermediate SPEC-T to the matrix SPEC-T in conformity to the assumption that “IM is successive-cyclic, driven by labeling failures, continuing until a criterial position is reached.” (POP\( ^+ \)) Now it is very important to notice that there bears \( \phi \)-features, by hypothesis; therefore, with this overt expletive, \( \gamma \) becomes visible to LA, and will be labeled \( \langle \phi, \phi \rangle \), as required.\(^7\) Again, at this final stage, all the relevant SOs are labelable; hence Transfer of T-complement \( \gamma \) also leads to the convergence of the derivation. In this way, under Chomsky’s POP\( ^+ \) labeling theory and Abe’s Move-approach, we can label the there-construction in conformity to the general principle of labeling.

The proposed labeling analysis of the there-construction keeps EKS’s (2014) labeling analysis of the contrast almost intact. As is clear from the above illustration, since our analysis is based on LA, we can preserve the insight of EKS’s labeling analysis: \( \alpha \) in (3a) is visible to LA, labeled \( \langle \lambda \rangle \), so it converges at the interfaces, but \( \alpha \) in (3b) is invisible to LA, remains unlabeled, and hence crashes at the interfaces. Of particular importance to our analysis is that since an associate is required to raise to SPEC-R to enter into an Agree relation with a verb (V) for partitive Case checking, it is not expected to appear around such elements as likely (A) and to (T), as in (3b), which are not eligible for partitive Case checking. Therefore, as EKS argue, (3b) is still explained as a case of unlabelable \( \{XP, YP\} \) structures that are ruled out at the interfaces.

\(^7\) Given that the expletive there is a head D (cf. (9)) and \( \gamma \) is labeled \( \langle \phi, \phi \rangle \) by the agreeing features of D and T, the relevant labeling of \( \gamma \) should be reduced to minimal search to a \{H, H\} structure with prominent \( \phi \)-features, a variant of (2c).

Specifically with respect to (3a), it is not clear how EKS’s views are compatible with Chomsky’s view that “T is too ‘weak’ to serve as a label,” but under our labeling analysis, it is possible to provide a uniform characterization of labeling of English T, as follows: in English, independent of the finiteness of T, “SPEC of T must be visible when LA applies” (POP\( ^+ \)), and more specifically, English T always needs to “be strengthened by SPEC-T” to be labeled (POP\( ^+ \)), irrespective of whether it appears in the course of the derivation (in intermediate positions) or in the final stage of the derivation (in criterial positions). Hence, in (3a), the infinitival-T is strengthened by the expletive there in SPEC-T, and as a result, \( \alpha \) becomes visible when LA applies, and will be labeled \( \langle \lambda \rangle \) in collaboration with the raising strategy (2b). In other words, thanks to the copy of the expletive there in SPEC-T, which is created by successive-cyclic IM for labeling, the intermediate T’s structural status is identified in the course of the derivation and allowed to take part in LA. In this way, our analysis can provide a principled account for the contrast in (3), elaborating on EKS’s labeling analysis of Merge-over-Move.

In our analysis, labeling of \( \beta \) in (4) also falls into place. In (4), given that partitive Case is in principle checked by any type of verb (Belletti 1988; see (8b)), it follows that in such a complex there-construction as (4), the associate a man can, in principle, check its partitive Case with either the upper be or the lower be. Thus, if we take the upper be to be involved here for partitive Case checking of a man, then labeling of \( \beta \) is straightforward: \( \beta \) is labeled \( \langle \phi, \phi \rangle \) through the usual operations in the matrix v\( ^*P \) phase in this case, such as v\( ^* \)-to-R \( \phi \)-feature-inheritance, Agree(R, a man) and IM(a man, R), a step-by-step analogue to labeling of \( \alpha \) at the lower v\( ^*P \) phase level in (11). The relevant derivation of (4) is given below:

(14) There [\text{be} + v^*_\alpha \{\text{a man}_\beta\} [\text{t}_\gamma \text{man} \text{R} [\text{there}_\gamma \text{be} + v^* \{\text{a man}_\beta\} [R_\gamma \text{be} \text{R}\text{be}]]]]]]]]

In (14), \( \alpha \) is labeled \( \langle \lambda \rangle \) in conjunction with the raising strategy (2b), and \( \beta \) is labeled \( \langle \phi, \phi \rangle \) by the agreeing features of the raised associate a man and the matrix R; hence the derivation converges at the interfaces.

In fact, the contrast in (5) is captured in the same way.
For the same reason as above, in terms of the flexibility in the partitive Case checking of an associate, we can reasonably analyze α in (5a, b) as being labeled \( \langle \phi, \psi \rangle \) appropriately, as follows:

(15) a. There is likely to \([\text{be-v}^* \ [\alpha \ [\text{a building}_\phi] \ [R(demolished)_\psi \ [t_{\text{there}, \ t_a \ \text{building}]]]\]

b. There \([\text{is-v}^* \ [\alpha \ [\text{a building}_\phi] \ [\text{R(demolished)}_\psi \ [t_{\text{there}, \ t_a \ \text{building}}]]]\]

In (15a), since the lower R is involved in partitive Case checking of a building, it follows that α is labeled \( \langle \psi, \phi \rangle \) by the agreeing features of a building and the lower R. On the other hand, in (15b), since the upper R is involved in partitive Case checking of a building, it follows that α is labeled \( \langle \phi, \psi \rangle \) by the agreeing features of a building and the upper R. In this way, given the flexibility in the partitive Case checking of an associate, the labeling of complex there-constructions can also be accommodated in labeling theory.

Significantly, in our analysis, SPEC-RP is labeled \( \langle \phi, \psi \rangle \) via associate-raising, so it is predicted that if it does not take place, the derivation crashes at the interfaces. This prediction is borne out by the following paradigm:

(16) a. *There’ve been \([\alpha \ [\text{a thief}_\psi \ [\text{among us}_\phi]]\]

b. *There have been \([\alpha \ [\text{some books}_\psi \ [\text{on the table}]]\]

c. There’ve been \([\alpha \ [\text{some men}_\psi \ [R(\text{arrested})_\phi \ [t_{\text{there}, \ t_a \ \text{men}]}]]\]

d. There have been \([\alpha \ [\text{some books}_\psi \ [R(\text{put})_\phi \ [t_{\text{there}, \ t_a \ \text{books} \ [\text{on the table}]]\]]\]

Thus, the minimal pairs of there-constructions in (17) also follow from the application of the EPP at the vP phase level, in accord with the general principle of labeling.

In our analysis, labeling of \( \alpha \) in (6) is also straightforward. In (6), given that the expletive there bears \( \psi \)-features (which may or may not be valued by its associate), we can reasonably analyze \( \alpha \) as being labeled \( \langle \phi, \psi \rangle \) in terms of the agreeing features of the raised there and the matrix R, as follows:

(18) We [prove \(+v^* \ [\alpha \ [R\_\psi \ [t_{\text{there}, \ T(\text{to})} \ [\text{be-v}^* \ [\text{a thief}_\psi \ [\text{among us}_\phi ]]]]\]

In (18), since there has raised to the matrix SPEC-R, \( \alpha \) becomes visible to LA and will be labeled \( \langle \phi, \psi \rangle \) by the agreeing features of the raised there and the matrix R. In this way, by appealing to \( \psi \)-features inherently borne by the expletive there, we can accommodate the ECM construction as well.

Under our analysis, therefore, the theoretical problems of labeling are resolved and the distribution of the expletive there and the associate is correctly predicted.

4. Consequences

In this section, to see whether any interesting implication is obtained, let us consider two specific empirical puzzles that have resisted a satisfactory explanation in the literature on there-constructions: Lasnik’s (1995) minimal pair and Takano’s (1998) ECM paradigm.

First, consider the following contrast, which is puzzling in comparison to (5):

(19) Lasnik’s (1995) minimal pair

a. How↓ is there likely to be \([\text{a building demolished}_\psi \ [t\_↓]]\]

b. How↓ is there a building likely to be\([\text{demolished}_\psi \ [t\_↓]]\)?
Under our analysis, (19a) should have the following derivation before IM applies:

(20) \[ \gamma \ C_\text{Q} \ [\beta \ T_\delta \ [\nu \ R \ [\text{likely} \ [T \ (\text{to})] \ [\text{be} + \nu^* \ [\alpha \ R \ (\text{demolished})_\delta \ [\text{there, a building}] \ \text{how}]])]] ]

In (20), how is adjoined to the lower R to modify demolish, and the lower v*P constitutes a strong phase for partitive Case checking of a building (cf. (15a)), and through inheritance of ϕ-features and phasehood from v* to R, the lower R-complement, i.e., the SO = [there, a building], is identified as a transferred domain. Here, to label α, β, and γ, as required, the lower R, the matrix T, and the matrix C need to get strengthened by SPEC-R, SPEC-T, and SPEC-C, respectively. Then, as we have assumed so far, given that IM is successive-cyclic leading to a criterial position, we can derive (21) from (20) by applying IM(how, Q), IM(there, T), and IM(a building, R), successive-cyclically:

(21) \[ \gamma \ C_\text{Q} \ [\beta \ T_\delta \ [\nu \ R \ [\text{likely} \ [T \ (\text{to})] \ [\text{be} + \nu^* \ [\alpha \ R \ (\text{demolished})_\delta \ [\text{there, a building}] \ \text{how}]])]] ]

In (21), a building has raised from the base position (t_a building) to the lower SPEC-R in α, and there has raised from the base position (r^1_{there}) to the matrix SPEC-T in β, passing through the lower SPEC-ν* (r^2_{there}) and the intermediate SPEC-T' (r^3_{there}). Remarkably, how has raised from the base position (n_{how}) directly to the matrix SPEC-C in γ, without passing through any intermediate positions. This is crucially because phasehood is activated on the lower R, and how is not affected by the PIC at the lower v*P phase, and can remain in situ until the derivation reaches the matrix CP. As is clear from (21), since each application of IM does not violate any conditions in the course of the derivation, it follows that the lower R, the matrix T, and the matrix C get strengthened by SPEC-R, SPEC-T, and SPEC-C, respectively. As a result, α, β, and γ become visible to LA, labeled \( \langle \phi, \phi \rangle \), \( \langle \phi, \phi \rangle \), and \( \langle Q, Q \rangle \), respectively, and the derivation converges at the interfaces.

Let us turn to (19b). Under our analysis, (19b) should have the following derivation before IM applies:

(22) \[ \gamma \ C_\text{Q} \ [\beta \ T_\delta \ [\nu^* \ [\alpha \ R_\delta \ [\text{likely} \ [T \ (\text{to})] \ [\text{be} + \nu \ (\text{demolished} \ [\text{there, a building}] \ \text{how}]])]] ]

In (22), in contrast to (21), the upper v*P constitutes a strong phase for partitive Case checking of a building (cf. (15b)), and through the usual inheritance of ϕ-features and phasehood from v* to R, the upper R-complement (δ) is identified as a transferred domain. For the same reason as above, to label α, β, and γ, as required, the upper R, the matrix T, and the matrix C need to get strengthened by SPEC-R, SPEC-T, and SPEC-C through IM(how, Q), IM(there, T), and IM(a building, R), respectively.

Here, before examining the derivation of (22) further, it is important to notice that each application of IM needs to apply at the same time at the matrix phase level, because how is inside the transferred domain along with there and a building. In fact, considering the situation where the required IMs need to apply at once to the elements in the same transferred domain, it is conjectured that (22) will follow a more complex derivation than (20). Thus, to examine the derivation of (22) correctly, it is necessary to describe the implementation of successive-cyclic IM for labeling more precisely. Then, reinterpreting Chomsky’s (1993) Shortest Move or Chomsky and Lasnik’s (1993) Minimize Chain Links in the labeling theory, I would like to assume that each application of IM cannot pass a possible landing site.

With this assumption in mind, let us consider the following derivation to see how each application of IM to there, a building, and how is applied:

(23) \[ \gamma \ C_\text{Q} \ [\beta \ T_\delta \ [\nu^* \ [\alpha \ R_\delta \ [\text{likely} \ [T \ (\text{to})] \ [\text{be} + \nu \ (\text{demolished} \ [\text{there, a building}] \ \text{how}]])]] ]

First, for there to reach SPEC-T to label β, it is required to pass through the intermediate SPEC-T, an A-position (\( \text{(1)} \)), since the final destination of there is the matrix SPEC-T, an A-position. On the other hand, it would also be reasonable to conjecture that there passes through the matrix SPEC-ν* (\( \text{(2)} \)), as well, although it is an A’-position. As it has been argued in the literature, given that efficient computation requires successive-cyclic IM to SPEC-C and SPEC-ν*, i.e., SPEC of a phasehood-bearer, it seems to be natural to assume that there has raised to the matrix SPEC-ν*, as well, for reasons of computational efficiency. Thus, for there to reach the
SPEC-T to label $\beta$, it is expected that there goes through the following steps of IM: IM of there $\Rightarrow$ the intermediate SPEC-T (1) $\Rightarrow$ the matrix SPEC-v* (2) $\Rightarrow$ the matrix SPEC-T in $\beta$. Then, how about the process of IM of a building? For a building to reach the matrix SPEC-R to label $\alpha$, it arguably needs to pass through the intermediate SPEC-T, an A-position (1); since the final destination of a building is the matrix SPEC-R, an A-position, it is natural to regard such IM as an instance of A-movement as well. Hence, when there reaches SPEC-T to label $\beta$, the following steps should be expected for a building: IM of a building $\Rightarrow$ the intermediate SPEC-T (1) $\Rightarrow$ the matrix SPEC-R in $\alpha$. Finally, how about how? For how to reach the matrix SPEC-C to label $\gamma$, it should be required to pass through the matrix SPEC-v*, an A’-position (2), since the final destination of how is the matrix SPEC-C, an A’-position. Therefore, the following steps should be expected for how: IM of how $\Rightarrow$ the matrix SPEC-v* (2) $\Rightarrow$ the matrix SPEC-C in $\gamma$.

Here, it is crucial to notice that IM of there and IM of a building compete with each other for the one SPEC of the intermediate SPEC-T (1), and, at the same time, IM of there and IM of how compete with each other for the one SPEC of the matrix v* (2). As things now stand, the question is whether such a competition is indeed allowed by computation in the course of the derivation. As indicated by the grammaticality of a sentence like *(Who did John ask [why Mary was waiting for twho twhy]?) (where why and who arguably compete with each other for the one SPEC of the embedded C), given that such a competition is, in fact, excluded as part of computational procedures, it should be the case that SPEC cannot be used by more than one element at a time (see Fukui and Speas 1986 and Hiraiwa 2010 for relevant discussion, +

In passing, it is important to recall that in the convergent derivation (21), how is not required to pass through the lower SPEC-v*, competing with IM of there. This is crucially because phasehood there is activated on the lower R, and how is not affected by the PIC at the lower v-P level and can remain in situ, circumventing the competing derivation that causes (19b) to crash at the interfaces. Thus, if the above argument is on the right track, we can provide independent support for Chomsky’s POP+ labeling theory, since the exclusion of how from the transferred domain actually results from the idea that “phasehood is activated on the copy of R, and DP (which can be a wh-phrase) remains in situ, at the edge.”

Finally, let us consider the following paradigm which is puzzling in comparison to (6) to see whether any interesting implication is obtained:

24) Takano’s (1998) ECM paradigm

\[
\begin{align*}
\text{a.} & \quad \text{We proved } [\gamma \text{ Smith to be the thief}] \\
\text{b.} & \quad \text{We proved } [\beta \text{ Smith } [\alpha \text{ to the authorities}] \text{ be the thief}] \\
\text{c.} & \quad \text{We proved } [\beta \text{ to the authorities}] [\gamma \text{ there to be a thief among us}] \\
\text{d.} & \quad \text{We proved } [\beta \text{ there } [\alpha \text{ to the authorities}] \text{ be a thief among us}] 
\end{align*}
\]

(24a) and (24c) can be easily accommodated in the proposed labeling analysis; consider the following relevant derivation (where, for the sake of simplicity, I omit the derivation of the associate):

25) *We [prove+v* [\gamma to the authorities] [\alpha \text{ Smith/there } [T(to) [be [\cdots]]]]]]

In (25), through inheritance of $\phi$-features and phasehood from v* to R, the copy of R is activated as a phase head, be attributed to the competition between there and how rather than that of there and a building. Accordingly, I suggest that in both (15b) and (19b) there and a building do not compete with each other, and, in fact, they pass through the SPEC-T together, as in [there, a building], as the same underlying constituent until they reach and leave SPEC-T. Given this, the ungrammaticality of (19b) can reasonably be attributed to the competition between there and how.
and as a result, R-complement \((\alpha)\) is forced to undergo Transfer. However, notice that \(\alpha\), being of the form \(\{XP, YP\}\), is invisible to LA and remains unlabeled; hence the derivation is correctly expected to crash at the interfaces. Incidentally, one might wonder how \(\beta\) is labeled and why it does not induce a labeling failure as a \(\{XP, YP\}\) structure, but it is important to notice that in (25), the transfer of R-complement \((\alpha)\) actually reduces the relevant SO to a simplex one as follows (for such a manner of labeling, see Fukui 2011 and Narita 2014; also Goto 2013):

\[(26) \ [\beta \ [\text{to the authorities}] \ R] (PP, R)\]

Here, by transferring R-complement \((\alpha)\) in (25), it follows, in effect, that only the matrix R and the experiencer phrase \(\text{to the authorities}\) (PP) at the edge remain in the derivational space. Hence, \(\beta\) is labeled RP by unambiguous minimal search to R.\(^9\)

On the other hand, (24b) does not crash because it circumvents such a labeling failure; consider the following derivation:

\[(27) \ We \ [\text{prove} + v^* \ [\beta \ Smith_{\phi} \ [\text{to the authorities}] [t_{R_{\phi}} \ [\alpha \ Smith \ T(to) \ [\text{be} \ [\ldots]]]]]]\]

First, since \(\alpha\) in (27) is visible to LA, thanks to the raising of \(Smith\) to the matrix SPEC-R, \(\alpha\) does not induce a labeling failure here. Then, since an Agree relation is established between \(Smith\) and the matrix R (cf. (18)), \(\beta\) also becomes visible to LA, and will be labeled \(\langle \phi, \phi \rangle\) by the agreeing features of \(Smith\) and the matrix R. Now, all the relevant SOs are labelable, hence the derivation converges at the interfaces.

That said, in a situation where the experiencer phrase \(\text{to the authorities}\) (PP) is sandwiched between the raised \(Smith\) and the matrix R, one might wonder why it does not block the relevant Agree relation between them, and why it does not break the “stronger relation” that is required for the labeling of such a \(\{XP, YP\}\) structure (POP; see also footnote 1 above). I would like to suggest that the required Agree relation for labeling is ensured by implementing what we may call an asymmetrical double probe, as follows:

\[(28) \ We \ [\text{prove} + v^* \ [\beta \ Smith_{\phi} \ [\text{to the authorities} \]
\[t_{R_{\phi}} \ [\alpha \ Smith \ T(to) \ [\text{be} \ [\ldots]]]]]]\]

In this derivation, the intention is to create a stronger relation for labeling asymmetrically by a combination of (i) in-situ \(\phi\)-probe and (ii) ex-situ Case-probe. Thus, in (28), Agree(R, Smith) is first applied for \(\phi\)-valuation in situ, triggered by the unvalued \(\phi\)-features of R, and then Agree(Smith, R) for Case-valuation is applied, triggered by the unvalued Case-feature of \(Smith\), after \(Smith\) raises to the SPEC-R. Of particular importance to this analysis is that the Case-feature of NPs can serve as a probe independently of \(\phi\)-features and does not have to be valued as the reflection of \(\phi\)-agreement, contra Chomsky’s (2000, 2001) original probe-goal theory of agreement. In this paper, therefore, I conclude that in collaboration with the \(\phi\)-probe, the Case-feature of NPs, by itself, plays a crucial role as a probe to create a stronger relation for labeling.

Under these considerations, the PP in (28) is no longer an obstacle to the labeling of \(\beta\); since it has no ability to assign Case-value to \(Smith\), the PP does not block the Case-probing process; therefore, the required Agree relation for labeling, or more specifically, the so-called stronger relation for the labeling of \(\beta\) is established between the raised \(Smith\) and the matrix R through the asymmetrical double probe, as in (28). Thus, given that the stronger relation that is required for the labeling of such a \(\{XP, YP\}\) structure is established by an asymmetrical double probe, we can circumvent a potential labeling problem that may arise in such a derivation as (28).\(^{10}\)

Why does (24d) get worse even though it has the same derivation as (24b)? Consider the following relevant derivation as (24b)? Consider the following relevant
derivation of (24d):

\[(29) \text{"We [prove + v" } [\gamma \text{ there} \}_{T(to) [\text{be } [\ddots]]}] \]

Comparing it to the convergent derivation (27), it is significant to notice that although the labeling of \(\alpha\) can be done in the same way as (27), the labeling of \(\beta\) cannot. That is, just like (27), \(\alpha\) in (29) can be labeled in conjunction with the raising strategy, but unlike (27), \(\beta\) in (29) cannot, because it does not bear the Case-feature necessary for creating a stronger relation for labeling (cf. (28)). Specifically, in (29), Agree(R, there) for \(\phi\)-valuation may be applied in situ, but Agree(there, R) for Case-valuation is not invoked due to the lack of a Case-feature on there, so that the stronger relation that is necessary for labeling is not established between the raised there and the matrix R. Hence, \(\beta\) remains unlabeled and the derivation crashes at the interfaces.

However, in the cases that we have examined so far, where there is no intervening element between the raised there and the relevant T/R (cf. (13) and (18)), we have argued that the relevant SO can be labeled \(\langle \phi, \phi \rangle\) by the agreeing features of the raised expletive and T/R, without committing to any Case-feature. Therefore we need to clarify why the same does not hold for the labeling of \(\beta\) in (29). To overcome this problem, it is important to recall that the expletive there is in fact a head D (cf. (9)), and labeling of the SO is ultimately reduced to minimal search to a head-head (H, H) structure (cf. footnote 7). If so, I would like to suggest that to label such a head-head structure, a strictly local head-head relation is necessary to be created for ensuring a stronger relation for labeling. Under these considerations, we notice that the PP in (29) is now an obstacle to the labeling of \(\beta\): the PP breaks a strictly local relation between the raised there (H) and the matrix R (H); hence, \(\beta\) becomes invisible to LA and remains unlabeled. Consequently, the derivation is correctly expected to crash at the interfaces. Thus, given that the stronger relation that is required for the labeling of such a [H, H] structure is established by a strictly local head-head relation, we can make a clear-cut distinction between the good examples and the bad ones.\(^{11}\)

4. Conclusion

In this paper, taking seriously Chomsky’s conservative approach to the necessity of labeling – “all SOs that reach the interfaces must be labeled” (POP) –, I have considered how SOs in there-constructions are labeled to satisfy the general principle. First, I identified some potential problems concerning labeling in there-constructions that remain unclear under EKS’s labeling analysis of Merge-over-Move, and then argued that they can be resolved by the interaction of Chomsky’s labeling theory developed in POP\(^+\) and Abe’s Move-approach to them. I have shown that our labeling analysis of there-constructions can not only overcome the theoretical worries but also make the correct empirical predictions about the distribution of the expletive there and the associate. As a consequence of the proposal, I have considered two specific empirical puzzles that have resisted a satisfactory explanation in the literature on there-constructions: Lasnik’s minimal pair and Takano’s ECM paradigm. Maintaining much of the spirit of the minimalist program, I have demonstrated that they can be unified under labeling theory.

I hope that this study has succeeded in reconciling EKS’s labeling analysis of Merge-over-Move and Abe’s Move-approach to there-constructions with Chomsky’s POP\(^+\) labeling theory.

Works Cited


(i) a. *There, for me at least, was nothing left over.
   b. *There, outside on the street, has just been a little girl.
   c. For me at least, there was nothing left over.
   d. Outside on the street, there has just been a little girl.

I would like to thank one of the reviewers for bringing this point to my attention.

\(^{11}\) This analysis is corroborated by the fact that the same effect is observed in finite clauses as well (cf. Hannay 1985).


