

Eliminating Strong/Weak Parameter on T

[Introduction] This paper identifies potential flaws in the labeling theory proposed in Chomsky (2015) and suggests a system that eliminates them. I first point out that the strong/weak parameter on T regarding labeling, which is proposed by Chomsky to capture the parametric difference in EPP/ECP between non-null subjects languages (non-NSLs) and null subject languages (NSLs), is nothing but a stipulation and hence should be eliminated in favor of the optimal design of language. I then propose a new way of $\langle \varphi, \varphi \rangle$ labeling by *Merge of a head H to SPEC*. As a consequence of this proposal, it is shown that the strong/weak parameter on T is eliminable and the spirit of Chomsky's labeling analysis of EPP/ECP is maintained, with significant empirical advantages.

[Chomsky 2015] Dropping the long-standing assumption that EPP is a universal property of language, Chomsky (2015) proposes to parameterize the strength of T regarding labeling: T in English-type non-NSLs, with weak agreement, is too "weak" to serve as a label, hence there must be an overt subject in SPEC-T to label the SPEC-TP as $\langle \varphi, \varphi \rangle$ by the agreeing features, whereas T in Italian-type NSLs, with strong agreement, is "strong" enough to serve as a label, hence there is no need to have an overt subject in SPEC-T to label the SPEC-TP. Given this, he offers a unified account of the parametric difference in ECP: in non-NSLs, T is too weak to serve as a label, hence a *wh*-subject must stay in SPEC-T to keep the SPEC-TP labeled as $\langle \varphi, \varphi \rangle$, but in NSLs, T is strong enough to label the SPEC-TP, hence a *wh*-subject doesn't have to stay in SPEC-T and can raise further. Assuming that memory is phase-level, he also provides an argument for why *wh*-extraction from SPEC-T is allowed in English when the phase head C (i.e., the complementizer *that*) undergoes deletion: if the SPEC-TP has already been labeled as $\langle \varphi, \varphi \rangle$ at a phase level, then the information is available and stays labeled at the phase level, so that even if a *wh*-subject raise further to SPEC-C, no labeling failure occurs. In this particular case, he notes that phasehood of C is inherited by T along with φ -feature inheritance and accordingly activated on T by C(*that*)-deletion. Thus, T functions as a phase head here, and as a result, the interior of T is identified as a transferred domain, observing the PIC. For this reason, when C is deleted, a PIC violation by further *wh*-extraction from SPEC-T can be circumvented, but when C is not deleted, C retains phasehood, so that the interior of C (viz. TP) is transferred, along with a *wh*-subject in SPEC-T, and therefore yielding the *that*-trace effect. In this way, EPP, ECP, and their parametric differences can be unified under the labeling theory once we adopt the strong/weak parameter on T.

[Flaws] These arguments are suggestive, but unfortunately, there is a hurdle to be overcome before we accept the parameter as a principled property. The Minimalist Program tells us that the notion of "strength" runs the risk of stipulation. In fact, Chomsky (1998:127) himself has suggested the following research agenda, pointing out that such a notion should be eliminated in favor of the optimal design of language: "optimal design should eliminate such strange and difficult properties as strength." Given this, the strong/weak parameter on T should be eliminated. In relation to this, one of the conceptual problems of the parameter in question is why it is only T that undergoes such parameterization. In contrast to T, Chomsky (2015) claims that R is universally too weak to label. This asymmetry between T and R with respect to the strength of labeling itself may be an interesting question, but if it is dispensable, it should be abandoned, in favor of the optimal design of language. The parameterization of Italian T as strong also poses a fundamental question for Chomsky's (2013) argument for why the external argument (Subj) must raise from the v*P-internal position: $\{\alpha \text{ Subj}, v^*P\}$, neither a head, is structurally ambiguous and cannot be labeled, and hence Subj must raise. By raising Subj to SPEC-T, what is visible to labeling becomes v* alone, as in $\{t, v^*P\}$, *t* the lower copy of Subj, so that α is labeled as v*. The reason for the invisibility of the lower copy of the subject to labeling, he argues, is that it is part of a discontinuous element, which is created when some element (Subj) in a domain D1 (SPEC-v*) goes to a different domain D2 (SPEC-T). This argument is suggestive, but once we accept the strong/weak parameter on T, unfortunately, the impact of the argument is undermined: if Italian T alone can label the SPEC-TP due to its strength, then it is not necessary for the subject to raise to SPEC-T, allowed to stay in situ. The question is, then, how the predicate-internal subject construction is labeled in such a situation. According to Belletti's (2001, 2004) series of studies, the subject that does not undergo raising to SPEC-T amounts to staying around in its base v*P-internal position; hence still visible to labeling. The construction in question is thus expected to result in a labeling failure, contrary to fact, and the question remains how the predicate-internal subject construction is labeled when the subject does not raise to SPEC-T, or more relevantly, when T is strong.

[Refinement] Alexiadou & Anagnostopoulou (1998) offer an interesting approach to EPP-satisfaction: EPP can be satisfied by either DP-raising to SPEC-T or V-raising to T. They propose that in non-NSLs like English, with weak φ -agreement marking on V, EPP is satisfied by DP-raising to SPEC-T, whereas in NSLs like Italian, with rich φ -agreement marking on V, EPP can be satisfied by V-raising to T. Since it is easy to find morphological evidence for the postulation of such a distinction, I assume their approach to EPP-satisfaction in the labeling theory. Then, following Chomsky's (2013) suggestion that SPEC-Head agreement be mandatory for $\langle \varphi, \varphi \rangle$ labeling, I assume the Head-to-SPEC movement hypothesis (Toyoshima 2000, Matushansky 2006, a.o.), according to which all movement is to a SPEC position, whether the movement is of a head or of a phrase. This hypothesis seems to make sense conceptually, in particular in the context of the simplest Merge-based system: if Merge applies freely, as in $\text{Merge}(\alpha, \beta) \rightarrow \{\alpha, \beta\}$, and if *Merge of a head H to SPEC (Merge H-to-S)* is just an instance of the general application of Merge, then there is nothing to preclude the possibility, as is the case of Merge of a phrase to SPEC: once we accept a free Merge system, anything goes; rather, any stipulation that blocks it carries a burden of proof.

[Analysis] Only with these assumptions can Chomsky's labeling analysis of EPP be maintained: in non-NSLs like English, the SPEC-TP construction is labeled as $\langle \varphi, \varphi \rangle$ by raising DP to SPEC-T, as Chomsky argues, whereas in NSLs like Italian, it is done by raising of V to SPEC-T with Merge H-to-S. Note that raising of V to SPEC-T is not

blocked for non-NSLs like English – in a Merge-based system, what is possible in a language X is also possible in another language Y, and vice versa. Suppose that raising of V to SPEC-T applies in English, an effective way of $\langle\phi, \phi\rangle$ labeling in Italian. Then V raises to SPEC-T. In this case, however, a labeling failure occurs: since V in English has no rich ϕ -agreement, the raised V does not qualify as an element to trigger $\langle\phi, \phi\rangle$ labeling. Thus, in English-type languages, the absence of rich ϕ -agreement marking on V makes $\langle\phi, \phi\rangle$ labeling on T impossible. Thus, in our analysis, the parametric difference in EPP is attributed to richness of ϕ -agreement on V, and Merge per se is universal but only its lexical input varies. One of the important differences between Chomsky’s analysis and ours is that the universality of EPP is not maintained in the former, but in the latter is a consequence of labeling necessity: in our analysis, EPP is just a universal property of T, and, therefore, must be observed in all languages by means of Merge. In our theory, therefore, there is no need to stipulate the strong/weak parameter on T to ensure $\langle\phi, \phi\rangle$ labeling, which would be a welcome development, in favor of the optimal design of language. Given this consideration, the potential problem of asymmetry between T and R with respect to the strength of labeling also disappears: namely, T is inherently weak, just like R; hence no need to be parameterized to begin with. It follows then that the so-called EPP is yet a universal property of language, holding for both phases CP and v^*P in general.

[Consequences] Let’s return to the crucial difference between Chomsky’s analysis and ours, i.e., the way of satisfying EPP in Italian-type NSLs: for Chomsky, strong T has the key, but for ours, it is V with rich ϕ -agreement. Consider the following scheme in (1), t a copy of V: **(1)** [$_{\beta}$ V [T [$_{\alpha}$ DP, t_V]]] (“verra Gianni.” ‘Gianni will come.’) In (1), as noted above, β is labeled as $\langle\phi, \phi\rangle$ by raising V to SPEC-T. Note here that the normal labeling procedure can apply to α as well: by raising V to SPEC-T, what is visible to labeling in α becomes D alone, hence α is labeled D, as required. Thus, in our analysis, the problem of how the predicate-internal subject construction is labeled when the subject stays in situ can be circumvented. One may wonder what would happen if the construction in question is transferred to the interfaces with the label of D, not that of v^* , contrary to the standard assumption. However, there is no conceptual argument against this; hence no reason to reject what seems to be different from the standard. Another consequence of our theory gives a straightforward answer to the observation that the *that*-trace effect is obviated by a presence of the expletive: **(2)** *What_i do you think that [$_{\alpha}$ there is t_i in the box]?* Chomsky (2015) does not provide an explanation for the grammaticality of an example like (2), but once we incorporate Merge H-to-S, it becomes possible to account for it without extra stipulation: since α stays labeled as $\langle\phi, \phi\rangle$ thanks to Merge of *there* to SPEC-T, further *wh*-extraction is allowed. This analysis is compatible with Hornstein & Witkos’s (2003) proposal that the expletive is a head H, and shares the values of its ϕ -features with its associate by forming a constituent underlyingly, as in $\{D_{there}, NP_{associate}\}$. Details aside, Merge H-to-S permits a loophole for further extraction from the criterial position. Significantly, this is reminiscent of Rizzi & Shlonsky’s (2007) “skipping strategies” – that is, XP can raise further, if the criterial position is occupied by another element. Thus, without recourse to the strength of T, we can now reinterpret Chomsky’s labeling analysis of ECP in Italian as follows: the SPEC-TP stays labeled as $\langle\phi, \phi\rangle$ thanks to Merge of V to SPEC-T, hence further *wh*-extraction is allowed. The same analysis is true of the *que-qui* contrast in French: **(3)** *Qui_i crois-tu *que^{pk} qui [$_{\alpha}$ t_i viendra]?* The ungrammaticality with *que* is explained by Chomsky’s theory: *Qui_i* must stay in SPEC-T to keep α labeled as $\langle\phi, \phi\rangle$; otherwise, the derivation results in a labeling failure or a PIC violation. Also, the grammaticality with *qui* can be accounted for in the same way as (2), particularly in collaboration with Taraldsen’s (2001) proposal that French *qui* consists of *que* and expletive *-i*: thanks to Merge of the expletive *-i* to SPEC-T, the SPEC-TP stays labeled as $\langle\phi, \phi\rangle$, and hence further *wh*-extraction is allowed. In this way, once we adopt Merge H-to-S, EPP, ECP, their parametric differences, and the phenomena captured by skipping strategies can be unified under the labeling theory.

[Extension] The reasoning along these lines opens up a new possibility of explaining the difference between overt *wh*-movement languages like English and *wh*-in-situ languages like Japanese (cf. Fukui 1986, Kuroda 1988, a.o.). Compare: **(4)** a. English has *wh*-movement: *What_i did Mary buy t_i ?*

b. Japanese does not have *wh*-movement: *Mary-ga nani-o kai-masi-ta ka?* (meaning (4a))

If *wh*-interrogatives must be labeled as $\langle Q, Q\rangle$ with feature sharing (Chomsky 2013, 2015), then the question is: how the SPEC-CP construction is labeled in (4a) and (4b), respectively. Under the proposed theory, in which heads (such as Italian V, English *there*, or French *-i*) can ensure labeling by Merge H-to-S, the difference can be derived as follows: in English (4a), the SPEC-CP is labeled as $\langle Q, Q\rangle$ by raising the *wh*-phrase *what* to SPEC-C, whereas in Japanese (4b), it is done by raising of the Q-head *-ka* to SPEC-C from a *wh*-phrase via Merge H-to-S, as in [C [Mary-ga nani-o t_i kai-masita] ka_i]; see Hagstrom (1998). Thus, in our analysis, the parametric difference in *wh*-interrogatives is attributed to the very existence of the Q-head *-ka* (cf. Cheng 1991, Hasegawa 2005): thanks to Merge of *-ka* to SPEC-C, the SPEC-CP is appropriately labeled, and a *wh*-phrase can stay in situ. This analysis predicts that in Japanese, unlike English, raising from *wh* criterial position is allowed. This prediction is borne out. Compare: **(5)** a. English obeys the *wh*-criterion: **What_i do you wonder [$_{\alpha}$ t_i [C_Q [John likes t_i]]]?*

b. Japanese does not obey it: *Nani-o, Ken-wa [$_{\alpha}$ t_i [C_Q [Tom-ga t_i t_i kata ka]]] siritagetiteiru no?*
‘What does Ken want to know whether Tom bought?’

In English (5a), since *what* raises from the criterial position, α loses the label of $\langle Q, Q\rangle$, and thus the sentence is ruled out as a labeling failure in the embedded CP. In contrast, in Japanese (5b), since α stays labeled as $\langle Q, Q\rangle$ thanks to Merge of *-ka* to SPEC-C, further *wh*-movement of *nani-o* from the criterial position is allowed without inducing such a failure. In this way, if we invoke simple visibles (heads such as Italian V, English *there*, French *-i*, and Japanese *-ka*) in the simplest Merge-based system, then complex visibles (parametric variations with respect to overt subjects or *wh*-movement) can be explained by interaction of the simple invisibles (viz. Merge and labeling).