On the Derivation of Wh-Subject Questions in English

Suenaga, Kodai
Graduate School of Humanities, Kyushu University : Master's Program

https://doi.org/10.15017/4793634

出版情報：九大英文学. 64, pp.85-110, 2022-03-31. 九州大学大学院英語学・英文学研究会
バージョン：
権利関係：

On the Derivation of *Wh*-Subject Questions in English*

Kodai Suenaga

1. Introduction

With respect to the derivation of *wh*-subject questions in English, two analyses have been adopted in numerous previous studies. One is the Vacuous Movement Hypothesis (henceforth VMH), in which *wh*-objects/adjuncts move to SPEC-CP but *wh*-subjects do not (George (1980), Chomsky (1986), and among others). The other is the unified analysis, in which, by analogy with *wh*-objects/adjuncts, *wh*-subjects also move to SPEC-CP (Bošković (2019), Messick (2020), and among others).

Recently, Chomsky (2013, 2015) implied the significance of the VMH by assuming that English T is too weak to serve as a label, which is also the key to account for the EPP and ECP uniformly. Following Chomsky's labeling theory, *wh*-subjects cannot move to SPEC-CP due to the weakness of T, so that the agreement between the overt subject and T has to be established within the SPEC-TP, supporting the VMH. However, I will point out empirical and theoretical problems with the VMH analysis supported by the labeling theory. On the other hand, I will argue that serious problems arise in the case where *wh*-subjects move to SPEC-CP as well. Thus, I provide a convergent derivation that allows *wh*-subjects to move to SPEC-CP in a conceptually desirable way, which overcomes these problems. In fact, I assume that the derivation of *wh*-subject questions in English involves T-to-C raising by Internal pair-Merge, where Affix Hopping and nominative Case assignment will be the important factors.

I organize this paper as follows. First, in section 2, I will review the derivation of *wh*-subject questions in English from the view point of the labeling theory. In section 3, I will point out two empirical problems with the VMH analysis, and then

discuss the theoretical contradiction in the labeling framework proposed by Chomsky (2015). Section 4 poses serious problems resulting from *wh*-subjects moving to SPEC-CP. Then, I will propose in section 5 that the derivation of *wh*-subject questions in English involves T-to-C raising by Internal pair-Merge, which can resolve the problems with the two competing analyses. Section 6 extends my analysis to the *that*-trace effect, providing (non-)convergent derivations without resorting to any technical notions that have been proposed so far. Section 7 concludes this paper.

2. The Structures of *Wh*-Subject Questions in English

In this section, I will review the derivation of *wh*-subject questions in English from the view-point of the labeling theory proposed by Chomsky (2013, 2015).

2.1 *Wh*-Movement in Question

In the literature, two competing analyses have been proposed for *wh*-subject questions. One is the VHM analysis, in which *wh*-subjects occupy SPEC-TP as the final landing site. The other is the unified analysis, in which, by analogy with *wh*-objects/adjuncts, *wh*-subjects end up in SPEC-CP, as illustrated in (1b, c), respectively.

(1) a. Who loves Mary?
b. [$_{CP}$ C [$_{TP}$ Who loves Mary]]?
c. [$_{CP}$ Who C [$_{TP}$ loves Mary]]?

In (1b), the *wh*-subject *who* stays in SPEC-TP. In (1c), on the other hand, it undergoes further movement to SPEC-CP. In any case, both derivations yield the same linear ordering. Now the question is, following the VHM analysis, why do *wh*-subjects not move to SPEC-CP?

The VMH has been adopted in numerous previous studies since George (1980). Recently, Chomsky (2013, 2015) implied the significance of the VMH by postulating the inheritance of the functional properties from C to T, which is also the key to account for the *that*-trace effect. Based on the framework of Chomsky (2013, 2015), Tanigawa (2017) argues that all features of C should be inherited by T and that C undergoes deletion in the derivation of *wh*-subject questions, defending the VMH.

Before looking at his analysis in some detail, I will review Chomsky's (2013, 2015) labeling framework, focusing on both the notion of the Labeling Algorithm and an explanation of the *that*-trace effect.

2.2. Labeling Theory (Chomsky 2013, 2015)

Chomsky (2013, 2015) holds that Merge is just a set-formation operation (cf. Seely (2006)) and is applied freely, and that any set formed by Merge must have a label for interpretations at the conceptual-intentional and sensorimotor interfaces. In order to determine the label, Chomsky introduces the Labeling Algorithm by Minimal Search. Minimal Search works in a top-down fashion, and the first head it finds serves as the label of the set. Chomsky argues that there are two ways to label the set. To begin with, consider the set produced by merging a head H and a phrase XP:

(2) $\{_{\alpha}$ H, XP$\}$ α=H

The Labeling Algorithm is straightforward in this case. Minimal Search sees the set and finds the first head H, which is chosen as the label. A more complicated case is the set produced by merging two phrases XP and YP:

(3) $\{_{\alpha}$ XP, YP$\}$ α=?

In XP-YP situations like (3), Minimal Search cannot identify the label, because it finds two heads X and Y simultaneously. Chomsky argues that there are two strategies to circumvent this situation. The first one is the application of Internal Merge, which makes a copy of the moved element invisible to Minimal Search.

(4) $\{$ XP $\{_{\alpha}$ ~~XP~~, YP$\}$ α=Y

Assuming that XP is moved here, the label of the set can be determined as Y like (4), where the head Y of YP is the label, because Minimal Search cannot see the copy of XP left by Internal Merge (marked with strikethrough). The second strategy is that agreeing features provide the label, as in (5).

(5) $\{_{\alpha}$ XP$_{[F]}$, YP$_{[uF]}\}$ α=<F, F>

In this case, X carries a valued feature [F] and Y has the unvalued feature [*u*F], then the label is determined based on the shared features as <F, F>.

In addition to this algorithm, Chomsky proposes the following assumption,

which not only deduces the EPP but provides a new insight into the *that*-trace effect.

(6) T is too "weak" to serve as a label (in English). With overt subject, the SPEC-TP construction is labeled by the agreeing features.

(Chomsky (2015: 9))

Consider the relevant structure in (7), where the *phi* feature agreement between DP and TP is the point.

(7) a. {$_{\beta}$ T {$_{\alpha}$ DP vP}}

b. {C {$_{\gamma}$ DP$_{[\varphi]}$ {$_{\beta}$ T$_{[u\varphi]}$ {$_{\alpha}$ ~~DP~~ vP}}}} α=vP, β=T, γ=<φ, φ>

In (7a), the DP subject is base-generated in SPEC-vP, producing the DP-vP configuration. After Merge of T, in order to determine the label of α, DP moves to SPEC-TP by Internal Merge, while producing the DP-TP configuration in (7b). In this case, their heads both share *phi* features and they hold an agreement relation. Now T can serve as the label of β because it is strengthened by the <φ, φ> label. If the DP subject does not move to SPEC-TP here, T cannot serve as the label, which leads the derivation to crash at the interfaces. That is, the EPP can be deduced from labeling.

With respect to the *that*-trace effect, it is well-known that the extraction of *wh*-subjects from inside subordinate clauses with overt *that* is banned like (8):

(8) a. *Who do you think that loves Mary?

b. {$_{\beta}$ who that {$_{\alpha}$ ~~who~~ T {$_{vP}$ ~~who~~ loves Mary}}} α = ?

In (8b), the *wh*-subject *who* moves to lower SPEC-CP, escaping from the Transfer domain (marked with shading). Recall that T cannot serve as the label unless it undergoes <φ, φ> labeling. The relevant labeling is impossible here, because Minimal Search cannot see the copy of *who* in SPEC-TP. That is, T cannot serve as the label if *wh*-subjects move to SPEC-CP, which leads the derivation to crash at the interfaces.

Things change if *that* is null as in (9):

(9) a. Who do you think loves Mary?

b. {$_{\beta}$ that {$_{\alpha}$ who T {$_{vP}$ ~~who~~ loves Mary}}}

c. {$_{\beta}$ $\emptyset$ {$_{\alpha}$ who$_{[\varphi]}$ T$_{[u\varphi]}$ {$_{vP}$ ~~who~~ loves Mary}}} α = <φ, φ>

The structure at the point where the phase head *that* merges is shown in (9b). In this case, *who* moves up to SPEC-TP and all the functional features of *that* (*phi* feature,

tense, Q) including its phasehood are inherited by T. Chomsky assumes further that *that*-deletion applies before Transfer (marked with Ø), and then the Transfer domain shifts to the complement of T as in (9c). Therefore, *who* is still available in the next phase. Moreover, thanks to <φ, φ> labeling, T can serve as the label here.

2.3. The Derivation of *Wh*-Subject Questions

In compliance with Chomsky's assumptions, as argued by Tanigawa (2017), *wh*-subjects cannot move to SPEC-CP in a sentence like (10), since T cannot serve as the label if *wh*-subjects move to SPEC-CP, which leads the derivation to crash.

(10) a. Who loves Mary?

b. {$_{\beta}$ who C {$_{\alpha}$ ~~who~~ T {$_{vP}$ ~~who~~ loves Mary}}} α = ?

In fact Tanigawa (2017) argues that, also in the derivation of *wh*-subject questions, all features of C are inherited by T and that C undergoes deletion. Following his analysis, the relevant derivation proceeds as follows.

(11) a. Who loves Mary?

b. {$_{\beta}$ C {$_{\alpha}$ who$_{[uQ, \varphi]}$ T {$_{vP}$ ~~who~~ loves Mary}}}

c. {$_{\beta}$ Ø {$_{\alpha}$ who$_{[uQ, \varphi]}$ T$_{[Q, u\varphi]}$ {$_{vP}$ ~~who~~ loves Mary}}} α = <Q, Q>, <φ, φ>

In this case, *who* is base-generated within SPEC-vP and then ends up being in matrix SEPC-TP as in (11b), where T inherits Q and *phi* features of C. Then C undergoes deletion as in (11c). Therefore, T is strengthened by <φ, φ> labeling, thereby T can serve as the label here. In short, the VMH is supported in terms of the labeling theory.

To summarize, I have reviewed the derivation of *wh*-subject questions in English in terms of the labeling theory. In fact, following Tanigawa's (2017) analysis, I have shown that the reason why *wh*-subjects do not move to SPEC-CP can be attributed to the weakness of T: with the overt subject, the SPEC-TP construction has to be labeled by the agreeing features, supporting the VMH.

3. Arguments against *Wh*-Subjects Staying in SPEC-TP

In this section, I will first discuss two empirical problems with the VMH analysis and then point out the theoretical contradiction in the labeling theory.

3.1. Problems with Sluicing and Scopal Contrast

As discussed in the previous section, in terms of the labeling theory, *wh*-subjects cannot move to SPEC-CP due to the weakness of T, so that the agreement relationship between the overt subject and T has to be established within the SPEC-TP, supporting the VMH. As will be seen below, however, the VMH confronts two problems.

The first argument comes from the elliptical phenomenon known as *sluicing*. As is usually assumed, sluicing involves *wh*-movement to SPEC-CP followed by TP-deletion, as illustrated in (12):

(12) a. John solved something, but I don't know what.

b. …I don't know [$_{CP}$ what C [$_{TP}$ John T [$_{vP}$ solved t_{what}]]].

As shown in (12b), the *wh*-object *what* moves to the embedded SPEC-CP and then TP is deleted (marked with shading here). If the same applies to subject questions, as Bošković (2019) and Messick (2020) say, *wh*-subjects cannot be located within SPEC-TP. Consider the relevant elliptical sentence in (14):

(14) A: Someone solved the problem.

B: Who?

In this case, if the *wh*-subject *who* remains in SPEC-TP, it should be included within TP, the sluiced site. On the other hand, if *who* moves to SPEC-CP, it can escape from the sluiced site, as illustrated in (14b, c), respectively.

(14) a. [$_{CP}$ C [$_{TP}$ who T [$_{vP}$ solved the problem]]]?

b. [$_{CP}$ who C [$_{TP}$ t_{who} T [$_{vP}$ solved the problem]]]?

In order for this argument to go through, however, another possible alternative must be ruled out.[1] One might assume that the vP-deletion applies here like (15):

(15) [$_{CP}$ C [$_{TP}$ who T [$_{vP}$ solved the problem]]]?

In (15), even if *who* remains in SPEC-TP, it can escape from the deleted site, vP. As (16) shows, however, *do*-support should be applied in that case:

(16) A: Someone solved the problem.

B: Who did? (Merchant (2008))

As can be seen from the dummy auxiliary *did*, Merchant (2008) observes that *do*-support takes place in (16B) if vP-deletion is applied (see also Lasnik (1999a)).

The second argument against the VMH is the scopal contrast, as argued by Mizuguchi (2014) and Bošković (2019). Consider the relevant sentences in (17):

(17) a. Someone loves everyone. (s.o. > everyone; everyone > s.o.)
b. Who loves everyone? (who > everyone; *everyone > who)
(Mizugichi (2014), Bošković (2019))

In (17a), the object quantifier *everyone* can take scope over the subject *someone* (abbreviated as s.o. in parenthesis). In (17b), on the other hand, *everyone* cannot take scope over the *wh*-subject *who*. As Bošković (2019) says, this contrast is unexpected if the *wh*-subject *who* could occupy SPEC-TP just like the subject *someone*.

3.2. Theoretical Contradiction of Weak T

Chomsky (2015) assumes that T is too weak to serve as a label, and that <φ, φ> labeling is necessary to strengthen the weak T in English. Chomsky extends this assumption to another head, the lexical root R in all languages, providing a new insight into the Exceptional Case Marking (ECM): the subject DP of the lower infinitival clause is assigned accusative Case by the matrix verb like *believe*. Consider the relevant structure in (18), where the movement of *John* is the point:

(18) a. I believe John to win.
b. $\{_{\delta}$ v* $\{_{\gamma}$ John$_{[\varphi]}$ $\{_{\beta}$ R$_{[u\varphi]}$ $\{_{\alpha}$ ~~John~~ to win$\}\}\}\}$ β=R, γ=<φ, φ>
c. $\{_{\delta}$ R-v* $\{_{\gamma}$ John$_{[\varphi]}$ $\{_{\beta}$ R$_{[u\varphi]}$ $\{_{\alpha}$ ~~John~~ to win$\}\}\}\}$ δ=R-v*

The structure at the point where the phase head v* merges is shown in (18b). In this case, the DP *John* moves to SPEC-RP from the infinitival SPEC-TP, and then the unvalued *phi* feature of v* is inherited by R (*believe*), so that the agreement relationship between *John* and R is established. Therefore, thanks to <φ, φ> labeling by Minimal Search, R can serve as the label here. Chomsky assumes further that, after Minimal Search, R-to-v* raising is applied by Internal pair-Merge, which makes v* invisible to syntax, so that the Transfer domain shifts to the complement of the R copy and the R-v* amalgam becomes the label of the set δ as in (18c).

Following these assumptions, it is easy to show that the extraction of the ECM *wh*-subject from within the infinitival SPEC-TP can be licit like (19):

(19) a. Who do you expect to win?

b. {$_{\delta}$ v* {$_{\gamma}$ who$_{[\varphi]}$ {$_{\beta}$ R$_{[u\varphi]}$ {$_{\alpha}$ ~~who~~ to win}}}} β=<φ, φ>

c. {$_{\delta}$ R-v* {$_{\gamma}$ who$_{[\varphi]}$ {$_{\beta}$ R$_{[u\varphi]}$ {$_{\alpha}$ ~~who~~ to win}}}} δ=R-v*

In (19b), *who* moves to SPEC-RP from the infinitival SPEC-TP, and then the unvalued *phi* feature of v* is inherited by R, so that R can serve as the label due to <φ, φ> labeling here as well. Next in (19c), R-to-v* raising is applied and then the Transfer domain shifts to the complement of the R copy. Therefore, *who* is still available in the next phase. However, consider the inside of the transferred set α. Recall in Chomsky (2015) that T is too weak to serve as a label and that it needs to be strengthened by <φ, φ> labeling in English, which is also the crucial key to vindicate the VMH. If so, as Mizuguchi (2017) says, the infinitival T cannot serve as the label, since it is impossible to establish the agreement relationship between *who* and T (*to*) within the set α, which would lead the derivation to crash.

In short, I have pointed out two empirical problems with the VMH analysis, concerning *sluicing* and the scopal contrast, and then discussed the theoretical contradiction in the labeling framework proposed by Chomsky (2015), where it is not clear how the set involving infinitival T is labeled. These facts suggest the possibility that *wh*-subjects move to SPEC-CP in the derivation of *wh*-subject questions.

4. Problems with *Wh*-Subjects Moving to SPEC-CP

This section poses serious problems resulting from *wh*-subjects moving to SPEC-CP in English. In fact, Affix Hopping and nominative Case assignment will be the important factors in this discussion.

4.1. Problems with Affix Hopping

In the previous section, I pointed out the empirical and theoretical problems with the VMH analysis derived from the labeling theory. However, two serious problems arise in the case where *wh*-subjects move to SPEC-CP.

The first problem comes from Affix Hopping, which is a traditional morphological operation by which an unattached affix in the T position is lowered

onto a verb. Hayashi (2020) explains the reason why subjects cannot move to SPEC-CP in English, without resorting to "weak T." In fact, contrary to Chomsky (2015), he argues that all heads are strong and qualify to serve as labels. Before looking into the relevant problem, I will review Hayashi's (2020) proposals in some detail.

Hayashi (2020) assumes that, following Epstein, Kitahara and Seely (EKS) (2017), agreement (feature valuation) is established at the interfaces based on <F, F> labeling by Minimal Search. Due to this view, there is no need to care about the simultaneous occurrence of feature valuation and Transfer, so that Feature Inheritance can be the optional operation. Furthermore, this optional system gives a solution to the problem with the notion of weak heads in Chomsky (2015).

Richards (2007) argues that Feature Inheritance is a required operation, in that unvalued features have to be transferred simultaneously with their valuation (see Richards (2007) for a detailed discussion). However, recall the discussion in the previous section concerning ECM constructions. Following Chomsky (2015), owing to R-to-v* raising by Internal pair-Merge, the unvalued *phi* feature on R (inherited from the phase head v*) is not transferred simultaneously with its valuation. Accordingly, Richards's argument for Feature Inheritance is no longer tenable in terms of the labeling theory, so that Hayashi (2020) states that Feature Inheritance can be applied freely and only convergent derivations survive.

Indeed, the optional Feature Inheritance helps resolve the theoretical problem arising from the notion of "weak R" in the labeling framework. Recall that in Chomsky (2015), the agreement relationship between R and its complement has to be established, whereby R can serve as the label via <φ, φ> labeling (cf. ECM constructions). Although this notion works well with the label R with a nominal complement bearing *phi* feature, there is a serious problem concerning the label R without it. Consider the relevant structure in (20), where the derivation of bridge verbs like *think* is the point.

(20) a. I think that John loves Mary.

b. $\{_{\gamma}$ v*$_{[u\varphi]}$ $\{_{\beta}$ R $\{_{\alpha}$ that John loves Mary$\}\}\}\}$ α=that, β＝R

c. $\{_{\gamma}$ R-v*$_{[u\varphi]}$ $\{_{\beta}$ ~~R~~ $\{_{\alpha}$ that John loves Mary$\}\}\}\}$

(20b) shows the stage where the matrix phase head v* is introduced into the derivation. In this case, Hayashi proposes that Feature Inheritance should not be applied and then Minimal Search provides the labels α and β. Next in (21c), R-to-v* raising occurs with Internal pair-Merge and thereby v* with [*u*φ] becomes invisible to syntax, so that the phase is cancelled and Transfer does not take place, which is based on the view that unvalued features mark phases (Chomsky (2015)). In fact, if the inheritance of [*u*φ] from v* to R has to be applied in this case, it is unclear how [*u*φ] on R can be valued since the CP complement does not have *phi* feature (see also EKS (2016)).

Before going into details about the issue with Affix Hopping, reconsider the *that*-trace effect at this point.

(21) a. *Who do you think that loves Mary?

b. {$_{\beta}$ who$_{[\varphi]}$ that {$_{\alpha}$ ~~who~~ T$_{[u\varphi]}$ {$_{R\text{-}v^*}$ ~~who~~ loves Mary}}}} α = T

The derivation reaches the stage shown in (21b). Recall that, as stated already, *who* cannot move to the lower SPEC-CP because of the notion of "weak T." Given the infinitival T, however, such a notion must be contradictory. Hayashi (2020) gives a different solution from Chomsky (2015), without resorting to the weak T. In this case, the culprit is not the label of T but the [*u*φ] on T. If *who* moves to SPEC-CP as in (21b), Minimal Search cannot assign the <φ, φ> label since it cannot see the copy of *who* in SPEC-TP, which leads the [*u*φ] on T to remain unvalued at the interfaces, causing the derivation to crash. That is why, even if T is not weak, *wh*-subjects cannot move away from within SPEC-TP in English.

Crucially, Hayashi argues that Affix Hopping is interrupted if subjects move to SPEC-CP in English. Following his argument, phase heads can keep the unvalued *phi* feature via optional Feature Inheritance. If so, another potential alternative can ensue. Consider what happens if subjects move to SPEC-CP and [*u*φ] is not inherited.

(22) {$_{\beta}$ Subject$_{[\varphi]}$ C$_{[v\varphi]}$ {$_{\alpha}$ T {$_{R\text{-}v^*}$ ~~Subject~~ love Mary}}}} α = T, β=<φ, φ>

In (22), the subject ends up being in the matrix SPEC-CP. In this case, the phase head C keeps the unvalued *phi* feature and then Minimal Search assigns the label <φ, φ> to the set β, thereby [*u*φ] on C is valued at the interfaces. Here comes the problem with Affix Hopping. With respect to the case like (22), Hayashi assumes that the valued

phi feature on C must be attached to a verb V (R-v*) via Affix Hopping at the sensorimotor interface. In this case, however, Affix Hopping from C to V should be interrupted by T lying halfway between the two heads, since affixes can only hop from one head to the head immediately below it. Therefore, even if [*u*φ] on C can be valued, the derivation where subjects move to SPEC-CP must be ruled out in English.

4.2. Problems with Nominative Case Assignment

The second problem concerns with the locus of *wh*-subject movement in the derivation. As for the derivation of *wh*-subject questions (under the unified approach), the most intuitive analysis would be that *wh*-subjects move to SPEC-CP through SPEC-TP, as shown in (23).

(23) a. Who loves Mary?

b. [$_{CP}$ who C [$_{TP}$ ~~who~~ T [$_{R\text{-}v^*}$ ~~who~~ loves Mary]]]?

In (23b), *who* is base-generated in the argument position inside the set R-v*, and then moves to SEPC-CP by way of SPEC-TP. However, a number of researchers have pointed out (explicitly or implicitly) that *wh*-subjects move directly from their argument position to SPEC-CP without stopping over SPEC-TP. One recent study adopting this view is Messick (2020), who argues that the movement of *wh*-subjects proceeds in a one-fell-swoop to SPEC-CP like (24):

(24) a. Who loves Mary?

b. [$_{CP}$ who C [$_{TP}$ T [$_{R\text{-}v^*}$ ~~who~~ loves Mary]]]?

Messick (2020) assumes that T is not deficient in English, so that T can serve as a label even if *who* does not move to SPEC-TP here. In fact, as Bošković (2019) and Messick (2020) say, this one-fell-swoop view has been attested in a lot of languages such as some English dialects (McCloskey (2000); den Dikken and Griffiths (2018)), Italian dialects ((Rizzi (1982), Rizzi and Shlonsky (2007)), Lubukusu (Diercks (2010)), Kaqchikel (Erlewine (2016)), Kinande (Bošković (2016)), and so on.

Given the conventional EPP, the one-fell-swoop view might be less of an issue with languages that do not require overt subjects for SPEC-TP. In English, however, the finite T generally asks for an overt subject and nominative Case is assigned via

the agreement relationship between them. For instance, EKS (2012) propose that T inherently bears a tense feature in the lexicon and that nominative Case is assigned by the combination of *phi* agreement with the tense property of T in English. If *wh*-subjects move directly from their argument position to SPEC-CP, it brings up a problem with nominative Case assignment. That is, following the one-fell-swoop view, there is a need to explain how nominative Case can be assigned without agreement between *wh*-subjects and T.

As for the issue above, one may assume that the relevant Case assignment is possible via the agreement relationship between the *wh*-subject and C, if C inherently bears a tense feature in the lexicon and then keeps the unvalued *phi* feature as its own. However, this assumption raises questions: What exactly is the T without tense property? Can such a T serve as the label at the interfaces? Furthermore, as Hayashi (2020) argues, the problem with Affix Hopping remains as well.

To sum up, I have pointed out the two problems resulting from *wh*-subjects moving to SPEC-CP. In fact, following the one-fell-swoop view, there are some glitches to be ruled out before it is vindicated, where Affix Hopping and nominative Case assignment constitute the defective factors.

5. Proposal and Analysis

In this section, based on the analyses of EKS (2012, 2017) and Hayashi (2020), I will reconsider the derivation of *wh*-subject questions in English and attempt to provide a convergent derivation that allows *wh*-subjects to move to SPEC-CP in a conceptually desirable way, overcoming the thorny problems encountered in previous sections.

5.1. Assumptions

As discussed so far, I pointed out empirical and theoretical problems concerning the VMH analysis derived from the labeling framework proposed by Chomsky (2015). However, serious problems result from *wh*-subjects moving to SPEC-CP as well. In order to solve these problems, I make the following assumptions.

(25) a. All heads are strong to serve as labels, and Feature Inheritance can be applied freely. (cf. Hayashi (2020))

b. Nominative Case is assigned according to the combination of the <φ, φ> label with an inherent tense feature of T. (cf. EKS (2012))

c. In the matrix CP, Transfer and Minimal Search apply to the entire phase. (cf. Obata (2010))

d. T-to-C raising occurs in Narrow Syntax. (cf. Landau (2020))

e. The order of phase-level operations proceeds as follows:

1. (optional) Feature Inheritance → 2. Minimal Search →

3. Internal pair-Merge → 4. (entire) Transfer. (cf. Chomsky (2015))

The assumption in (25a) resolves the issues discussed in the previous sections. Since all heads are strong, the infinitival T and the bridge verb R can serve as labels without <φ, φ> labeling (see also Hayashi (2020)). The assumption in (25b) contributes to the solution of the problem with Case assignment as will be seen later. Next in (25c), I assume with Obata (2010) that the entire phase is transferred in the case of the matrix CP. Since Chomsky (2000), it has long been assumed that Transfer applies to the complement of a phase head. If so, however, the elements in the edge of matrix CPs will never be included in any of the transferred domains. That is questionable because they should be pronounced and interpreted. Even more questionably, if Minimal Search has to see the domains to be transferred, how can it determine the labels involving the relevant elements? That is why Transfer and Minimal Search should be applied to the entire phase in the case of matrix CPs. In (25d), I follow the view that T-to-C raising takes place in syntax, which can be deduced from demonstrable scope and polarity effects (see Landau (2020) in detail). Finally, the order of phase-level operations in (25e) looks almost similar to the one in Chomsky (2015), but it adopts the new options of optional Feature Inheritance and entire Transfer.

5.2. Affix Hopping and the Case Assignment

Before considering the derivation of *wh*-subject questions, this section briefly discusses the issues of Affix Hopping and Case assignment. Following Hayashi's

(2020) analysis, Affix Hopping is disrupted if *wh*-subjects move to SPEC-CP. EKS (2012) states that nominative Case is assigned according to the combination of *phi* agreement with the inherent tense property of T. So, I will attempt to give the solution to the two problems, while still allowing *wh*-subjects to move to SPEC-CP during the derivation.

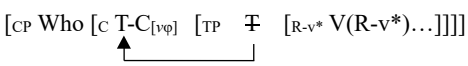
I propose that the derivation of *wh*-subject questions in English involves T-to-C raising by Internal pair-Merge, which establishes the local relationship between the two heads, so that Affix Hopping works without interruption. This is based on the assumption that Affix Hopping is preceded by the delete operation. Consider the elliptical constructions in (26):

(26) a. John likes dogs, and Mary does ~~like dogs~~, too.

b. * John likes dogs, and Mary ~~likes dogs~~, too.

In English vP-ellipsis constructions without overt auxiliary verbs, *do*-support must take place as in (26a), where Affix Hopping is not applied. As Lasnik (1999b) argues, if Affix Hopping is applied prior to vP-deletion, *do*-support cannot occur, which makes the vP-ellipsis construction fail as in (26b). As for elliptical phenomena, I follow Landau's (2020) analysis in assuming that ellipsis and copy deletion recruit the same silencing operation, in which the silencing feature is hosted on the head of the elided constituent (including copy), which instructs the sensorimotor interface not to pronounce a deletion site. Following these arguments, even if *wh*-subjects move to SPEC-CP in the derivation, the local requirement for Affix Hopping can be satisfied by T-to-C raising, as illustrated in (27):

(27) *T to C raising*



In this simplified structure, the *wh*-subject ends up in SPEC-CP, and then T-to-C raising is applied by Internal pair-Merge. As is the case with vP-ellipsis, since the deletion of the T copy is applied prior to Affix Hopping (of the affixial [*v*φ] here), the local relationship between C (T-C) and V (R-v*) is established, so that the affixial [*v*φ] can hop to the head V, as shown below.[2, 3]

(29) *Affix Hopping*

$[_{CP}$ Who $[_{C}$ T-C$_{[v\varphi]}$ $[_{TP}$ < T > $[_{R\text{-}v^*}$ V…]]]] (CP = <Q, Q>, <φ, φ>)

In fact, thanks to T-to-C raising, nominative Case can be assigned to the *wh*-subject occupying SPEC-CP at the sensorimotor interface, which is based on the combination of the <φ, φ> label with the tense feature of T undergoing Internal pair-Merge to C.

5.3. Derivation

Under the present proposal, the derivation of matrix *wh*-subject questions proceeds as follows.

(29) a. Who loves Mary?

b. {$_{\gamma}$ C$_{[Q,\, u\varphi]}$ {$_{\beta}$ T {$_{\alpha}$ who$_{[uQ,\, \varphi]}$ {$_{R\text{-}v^*}$ … }}}} (EM of T and C)

c. {$_{\delta}$ who$_{[uQ,\, \varphi]}$ {$_{\gamma}$ C$_{[Q,\, u\varphi]}$ {$_{\beta}$ T {$_{\alpha}$ *t*who {$_{R\text{-}v^*}$ … }}}}} (IM of *who*)

d. {$_{\delta}$ who$_{[uQ,\, \varphi]}$ {$_{\gamma}$ C$_{[Q,\, u\varphi]}$ {$_{\beta}$ T {$_{\alpha}$ *t*who {$_{R\text{-}v^*}$ … }}}}}

(α = R-v*, β = T, γ = C, δ = <Q, Q>, <φ, φ>)

e. {$_{\delta}$ who$_{[uQ,\, \varphi]}$ {$_{\gamma}$ T-C$_{[Q,\, u\varphi]}$ {$_{\beta}$ ~~T~~ {$_{\alpha}$ *t*who {$_{R\text{-}v^*}$ … }}}}} (T to C raising)

f. {$_{\delta}$ who$_{[uQ,\, \varphi]}$ {$_{\gamma}$ T-C$_{[Q,\, u\varphi]}$ {$_{\beta}$ ~~T~~ {$_{\alpha}$ *t*who {$_{R\text{-}v^*}$ … }}}}} (entire Transfer)

In (29b), T externally merges to α and C to β in that order, where the *wh*-subject *who* with [*u*Q] and [φ] features is base-generated in the argument position. Next in (29c), *who* internally merges to γ. In (29d), Feature Inheritance does not occur and then Minimal Search applies to the entire set δ. Additionally, in (29e), T-to-C raising is applied by Internal pair-Merge. Finally, the entire CP phase is transferred as in (29f). Accordingly, based on the combination of the <φ, φ> label with the tense feature of T undergoing Internal pair-Merge to C, [*u*φ] on C is valued and nominative Case is assigned to the *wh*-subject *who* at the sensorimotor interface. In fact, since the deletion of the lower T copy is applied prior to Affix Hopping of [*v*φ] on C, the local relationship between C and V (R-v*) is established, so that the affixial [*v*φ] can be attached to the head V at the sensorimotor interface.

The derivation above provides a clue to the problems concerning sluicing and the scopal contrast. Consider the relevant elliptical structure pair in (30b-c):

(30) a. A: Someone solved the problem. B: Who?

b. [$_{CP}$ C [$_{TP}$ who T [$_{R\text{-}v^*}$ solve the problem]]]?

c. [$_{CP}$ who T-C [$_{TP}$ ~~T~~ [$_{R\text{-}v^*}$ solve the problem]]]?

In this case, if the *wh*-subject *who* remains within SPEC-TP as in (30b), it should be included within the sluiced site. On the other hand, since *who* ends up in SPEC-CP under the proposal, it can escape from the sluiced site as in (30c).[4]

As mentioned previously, the object quantifier *everyone* can take scope over the subject *someone* within SPEC-TP, but *everyone* cannot take scope over the *wh*-subject *who* as in (31).

(31) a. Someone loves everyone. (everyone > someone)

b. Who loves everyone? (*everyone > who)

As argued by Bošković (2019), this contrast is unexpected if *who* could occupy SPEC-TP just like the subject *someone*. However, it can be expected under the proposal. Consider the relevant structure pair in (32):

(32) a. [$_{CP}$ C [$_{TP}$ Someone loves everyone]] (everyone > someone)

b. [$_{CP}$ Who C [$_{TP}$ loves everyone]]? (*everyone > who)

Since the *wh*-subject *who* ends up in SPEC-CP in (32b), the final landing sites of the two subjects differ from each other, which gives rise to the relevant scopal difference.[5]

To summarize, I have provided the convergent derivation that allows *wh*-subjects to move to SPEC-CP, which can resolve the problems with the two competing analyses. In fact, I have proposed that the derivation of *wh*-subject questions in English involves T-to-C raising by Internal pair-Merge, which not only satisfies the local requirement for Affix Hopping but also enables nominative Case assignment.

6. Extension to the *That*-Trace Effect

In this section, following the proposed analysis, I will attempt to provide some new insights into the *that*-trace effect, without resorting to any of the technical notions such as the ECP, EPP or weak T. In fact, Affix Hopping and nominative Case assignment will be the crucial factors in this discussion as well.

6.1. Derivation

As mentioned earlier, *wh*-subjects cannot be extracted from within subordinate clauses introduced by overt *that* as in (37a):

(37) a. * Who do you think that loves Mary?
b. Who do you think loves Mary?

As (37b) shows, in Standard English, the extraction of *wh*-subjects is sensitive to the overt C head *that* and it becomes possible only if *that* is null. In order to elucidate this intriguing effect, various studies have been made relying on various technical notions, such as the ECP (Chomsky (1986)), the EPP (Mizuguchi (2008)), and weak T under the labeling theory (Chomsky (2015)). These kinds of approaches to the *that*-trace effect can seemingly be suggestive and explanatory in that ungrammatical data like (37a) can be captured with the relevant notions. On the other hand, it is also true that they have faced a lot of conceptual and theoretical criticisms so far. In this section, I will reconsider the *that*-trace effect under the present proposal, while making no appeal to any of them. Consider the relevant derivation in (38) (I omit irrelevant derivation for expository reasons):

(38) a. *Who do you think that loves Mary?
b. $\{_{\gamma}$ that$_{[u\varphi]}$ $\{_{\beta}$ T $\{_{\alpha}$ who$_{[u\mathrm{Q},\,\varphi]}$ $\{_{\text{R-v*}}$ … $\}\}\}\}$ (EM of *that*)
c. $\{_{\delta}$ who$_{[u\mathrm{Q},\,\varphi]}$ $\{_{\gamma}$ that$_{[u\varphi]}$ $\{_{\beta}$ T $\{_{\alpha}$ *t*who $\{_{\text{R-v*}}$ … $\}\}\}\}\}$ (IM of *who*)

(38b) shows the stage where the phase head *that* with the [*u*φ] feature externally merges to β. Next, in (38c), the *wh*-subject *who* with [*u*Q] and [φ] features internally merges to γ from the argument position. Now, the derivation reaches the phase level while simultaneously arriving at a problematic situation. In this case, it does not matter whether Future Inheritance occurs or not. What really matters here is that T-to-C raising cannot be applied, which is the crucial key to account for the *that*-trace effect. Recall in the previous section that T-to-C raising plays an essential role both in satisfying the local requirement for Affix Hopping and in assigning nominative Case to *wh*-subjects. This infeasibility of T-to-C raising is based on the natural assumption that an element moves into an unoccupied position by an overt element. Consider the relevant examples in (39):

(39) a. If you should change your mind, no one would blame you.

b. Should you change your mind, no one would blame you.

c. * If should you change your mind, no one would blame you.

In formal English, the subject and the auxiliary can be inverted in conditional clauses with *if* omitted. In (39b), for instance, the subject *you* and the auxiliary *should* are inverted. As (39c) shows, however, subject-auxiliary inversion cannot be applied when *if* is not omitted. Assuming that the head C position of CP in a conditional clause is occupied by *if*, and that the subject-auxiliary inversion involves T-to-C raising, the auxiliary cannot raise to the filled position, so that the subject-auxiliary inversion cannot be applied in (39c). Bearing these arguments in mind, consider the relevant structure in (38), repeated here as (40).

(40) a. *Who do you think that loves Mary?

b. $[_{\delta}\ \text{who}_{[u\text{Q},\ \varphi]}\ [_{\gamma}\ \text{that}_{[u\varphi]}\ [_{\beta}\ \text{T}\ [_{\alpha}\ t_{\text{who}}\ [_{\text{R-v*}}\ \dots\ ]]]]]$

In this case, even if [*u*φ] on *that* can be valued, T-to-C raising cannot be applied since the head C position of the subordinate clause in (40) is occupied by overt *that*, so that Affix Hopping and nominative Case assignment will never be feasible. Therefore, the derivation in (38) is doomed to fail, which derives the *that*-trace effect.[6]

Things change if *that* is null as shown in (41), where the feasibility of Affix Hopping and nominative Case assignment will be the point.

(41) a. Who do you think loves Mary?

b. $\{_{\gamma}\ \text{C}_{[u\varphi]}\ \{_{\beta}\ \text{T}\ \{_{\alpha}\ \text{who}_{[u\text{Q},\ \varphi]}\ \{_{\text{R-v*}}\ \dots\ \}\}\}\}$ (EM of C)

c. $\{_{\delta}\ \text{who}_{[u\text{Q},\ \varphi]}\ \{_{\gamma}\ \text{C}_{[u\varphi]}\ \{_{\beta}\ \text{T}\ \{_{\alpha}\ t_{\text{who}}\ \{_{\text{R-v*}}\ \dots\ \}\}\}\}\}$ (IM of *who*)

d. $\{_{\delta}\ \text{who}_{[u\text{Q},\ \varphi]}\ \{_{\gamma}\ \text{C}\ \{_{\beta}\ \text{T}_{[u\varphi]}\ \{_{\alpha}\ t_{\text{who}}\ \{_{\text{R-v*}}\ \dots\ \}\}\}\}\}$

(Inheritance of [*u*φ], α = R-v*, β = T)

e. $\{_{\delta}\ \text{who}_{[u\text{Q},\ \varphi]}\ \{_{\gamma}\ \text{T}_{[u\varphi]}\text{-C}\ \{_{\beta}\ \cancel{\text{T}}\ \{_{\alpha}\ t_{\text{who}}\ \{_{\text{R-v*}}\ \dots\ \}\}\}\}\}$ (T-to-C raising)

(41b) shows the stage where the phase head null C with [*u*φ] externally merges to β. Next in (41c), *who* with [*u*Q] and [φ] features internally merges to γ. In (41d), [*u*φ] of C is inherited by T, and then T-to-C raising is applied by Internal pair-Merge as in (41e), which makes C invisible to syntax, thereby the Transfer domain shifts to the complement of the T copy (cf. Chomsky (2015)).

The next phasal operation proceeds as follows.

(42) a. $\{_{\zeta}$ v*$_{[u\varphi]}$ $\{_{\varepsilon}$ R $\{_{\delta}$ who$_{[uQ,\,\varphi]}$ $\{_{\gamma}$ T$_{[u\varphi]}$-C $\{_{T}$ T $\{_{\text{R-v*}}$ t_{who} …$\}\}\}\}\}$

(EM of R and v*)

b. $\{_{\zeta}$ v*$_{[u\varphi]}$ $\{_{\varepsilon}$ R $\{_{\delta}$ who$_{[uQ,\,\varphi]}$ $\{_{\gamma}$ T$_{[u\varphi]}$-C $\{_{T}$ T $\{_{\text{R-v*}}$ t_{who} …$\}\}\}\}\}$

(γ=T-C, δ=<φ, φ>, ε＝R)

c. $\{_{\zeta}$ R-v*$_{[u\varphi]}$ $\{_{\varepsilon}$ R $\{_{\delta}$ who$_{[uQ,\,\varphi]}$ $\{_{\gamma}$ T$_{[u\varphi]}$-C $\{_{T}$ T $\{_{\text{R-v*}}$ t_{who} …$\}\}\}\}\}$

(R-to-v* raising)

In (42a), matrix R (*think*) and v* are introduced into the derivation. In (42b), Feature Inheritance does not occur and Minimal Search provides the labels under ε. Then in (42c), R-to-v* raising is applied. Following Hayashi (2020), since v* with [*u*φ] becomes invisible to syntax, the phase is cancelled and Transfer does not take place in this case, so that *who* is still available in the next phase level. Accordingly, based on the combination of the label <φ, φ> with the tense feature of T, [*u*φ] on T is valued and nominative Case is assigned to the *wh*-subject *who* at the sensorimotor interface.[7] Furthermore, even if *wh*-subjects move to the lower SPEC-CP, thanks to T-to-C raising, the local requirement for Affix Hopping can be satisfied as well.[8, 9]

6.2. Non-Subject Extractions

Before closing this section, I would like to investigate the derivation of non-subject extractions. Unlike the case of *wh*-subjects, interestingly, the extraction of *wh*-objects/adjuncts from inside subordinate clauses with overt *that* is licit, like in (43).

(43) a. What do you think that John loves?

b. Why do you think that John loves Mary?

In order to diagnose this contrast, consider the *wh*-object extraction in (44) (I focus solely on the extraction of *wh*-objects here, but the analysis carries over to the extraction of *wh*-adjuncts):

(44) a. What do you think that John loves?

b. $\{_{\delta}$ what$_{[uQ,\,\varphi]}$ $\{_{\gamma}$ that$_{[u\varphi]}$ $\{_{\beta}$ John$_{[\varphi]}$ $\{_{\alpha}$ T $\{_{\text{R-v*}}$ … $\}\}\}\}\}$ (IM of *what*)

c. $\{_{\delta}$ what$_{[uQ,\,\varphi]}$ $\{_{\gamma}$ that $\{_{\beta}$ John$_{[\varphi]}$ $\{_{\alpha}$ T$_{[u\varphi]}$ $\{_{\text{R-v*}}$ … $\}\}\}\}\}$

(Inheritance of [*u*φ], α = T, β = <φ, φ>)

(44b) shows the stage where the *wh*-object *what* with the [*u*Q] and [φ] features internally merges to γ. Next in (44c), [*u*φ] of the phase head *that* is inherited by T and then Minimal Search provides the labels α and β, which undergo Transfer. The next phasal operation proceeds as in (45):

(45) a. {ζ v*$_{[u\varphi]}$ {ε R {δ what$_{[uQ,\,\varphi]}$ {γ that {$_{<\varphi,\,\varphi>}$ John$_{[\varphi]}$ T… }}}}}

(EM of R and v*)

b. {ζ v*$_{[u\varphi]}$ {ε R {δ what$_{[uQ,\,\varphi]}$ {γ that {$_{<\varphi,\,\varphi>}$ John$_{[\varphi]}$ T… }}}}}

(γ=δ=that, ε＝R)

d. {ζ R-v*$_{[u\varphi]}$ {ε ~~R~~ {δ what$_{[uQ,\,\varphi]}$ {γ that {$_{<\varphi,\,\varphi>}$ John$_{[\varphi]}$ T… }}}}}

(R-to-v* raising)

In (45a), matrix R (*think*) and the phase head v* are introduced into the derivation. Next, in (46b), Feature Inheritance does not occur and then Minimal Search provides the labels under ε. Then in (46c), R-to-v* raising is applied by Internal pair-Merge, which cancels the phase. Thereafter, the *wh*-object *what* will move to the matric SPEC-CP, providing the <Q, Q> label. What I would like to emphasize is the fact that T-to-C raising is unnecessary in this case, since Affix Hopping and nominative Case assignment will be feasible based on the combination of the <φ, φ> label with the tense feature of T within the set <φ, φ>. In fact, (46) suggests that non-subject *wh*-phrases can be extracted from within embedded clauses headed by C, whether it be overt or null. For the same reason they are convergent without T-to-C raising.[10]

(46) a. What do you think John loves?

b. Why do you think John loves Mary?

In sum, extending my analysis to the *that*-trace effect, I have considered (non-)convergent derivations, without resorting to any technical notions such as the ECP, EPP, or weak T. In fact, I have shown that the intriguing effect can be derived from the infeasibility of Affix Hopping and nominative Case assignment. There are still remaining issues that must be discussed within the proposed analysis: for instance, subject raising in declarative clauses, *wh*-island phenomena, cross-linguistic differences, and so on. For lack of space, however, I have to leave these intriguing issues for future research.

7. Conclusion

In this paper, I have reconsidered the derivation of *wh*-subject questions in English. Following Tanigawa's (2017) analysis based on the framework of Chomsky (2013, 2015), I have shown that *wh*-subjects cannot move to SPEC-CP due to the notion of weak T, which supports the VMH. However, I pointed out that the VMH analysis faces empirical problems concerning *sluicing* and the scopal contrast. On the other hand, I argued that serious problems arise in the case where *wh*-subjects move to SPEC-CP as well. In fact, Affix Hopping and nominative Case assignment will never be feasible if *wh*-subjects move to SPEC-CP. Then, I provided the convergent derivation that allows *wh*-subjects to move to SPEC-CP in English, which can resolve problems with the two competing analyses. Specifically, I proposed that the derivation of *wh*-subject questions in English involves T-to-C raising by Internal pair-Merge, which not only satisfies the local requirement for Affix Hopping but also enables nominative Case assignment. Furthermore, I have shown that this analysis helps derive the *that*-trace effect, without resorting to any technical notions that have faced various difficulties.

Notes

*I wish to express my gratitude to Nobuaki Nishioka for insightful suggestions and comments. I would also like to thank Edmundo Luna for suggesting stylistic improvements. Remaining inadequacies are purely my own responsibility.

[1.] Messick (2020) empirically shows that this kind of response is not a case of pseudo-sluicing.

[2.] As will be discussed later, Transfer applies to entire matrix CP phases so that C and V (R-v*) will be included in the same transferred domain. Otherwise, assuming that the sensorimotor interface needs to see the whole derivation at the end, it can be said that Affix Hopping occurs after the whole derivation has been completed.

[3.] Hein (2018) states that ellipsis and copy deletion apply at the same time, but an enormous number of cross-linguistic studies have shown that the timing of ellipsis is distributed according to a wide variety of environments. Although such worthwhile topics remain, I shall not go into the arguments concerning elliptical phenomena.

[4.] In this case, TP will be marked as a deletion site since T has the silencing feature (or Merchant's (2001) [E] feature). Furthermore, I assume that deletion is also conducted for the head T itself (pair-merged by C) at the sensorimotor interface, because it has the silencing feature of its own. Therefore, overt elements will never appear on T in the sluiced construction.

[5.] Traditionally, it has been proposed that a quantifier like *everyone* is adjoined to IP/TP via covert LF movement. Given the simplest Merge, however, covert movement is no longer tenable. Furthermore, adjunction has long been under debate since the GB era, and the issue has not been settled under minimalist approaches either. Although such worthwhile issues remain, I will leave them for future research.

[6.] It should be noted that even if C is occupied by overt *that* subject-auxiliary inversion occurs in a certain situation as in (i) (the data comes from Bruening (2015)):

(i) She made it clear that under no circumstances would she cancel the trip.

Interestingly, however, the *that*-trace effect can be ameliorated under such a circumstance as in (ii) (the data comes from Culicover (1993)):

(ii) Leslie is the person who$_i$ I said that under no circumstances would t_i run for any public office.

I am not sure how this correlation can be explained. For a different analysis, see Ishii (2004).

[7.] One may well ask how the *wh*-subject *who* can get nominative Case after moving away the edge position of the T-C amalgam in syntax. I have no definitive answer to the question, but it is worthy noting that θ-role assignment shares the same concern for this matter. As for θ-role assignment, Chomsky (2021) assumes that a θ-assigner τ assigns a θ-role to structural position P(τ), and that X is θ-linked to P(τ) if a copy of X occupies P(τ). Hence, both X and its copy can be linked to their θ-assigner. If the same applies to Case assignment at the sensorimotor interface level, the highest copy of *who* can be linked to its Case-assigner since the lower copy of *who* occupies the relevant structural position: the edge of the T-C amalgam, in which *who* is assigned nominative Case via the combination of the <φ, φ> label with the tense feature of T. Since this topic goes beyond the domain of this paper, however, I will leave it for future research.

[8.] In English interrogative sentences without overt auxiliaries, *do*-support must take place and the dummy verb is followed by a subject in order to actually be interpreted as interrogative sentences. Importantly, the environment in which *do*-support can occur is highly restricted in

Standard English. In fact, it is available only in matrix clauses. It is usually assumed that this constraint comes from language specific property of Standard English. For instance in Belfast English, *do*-support takes place in embedded clauses and subjects follow *do* as in (i):

(i) Who did John claim did he see? (Henry (1995: 114))

If *do*-support only applies to matrix clauses in Standard English and if the inverted dummy verb arises from T-to-C raising, it is surprising that T-to-C raising occurs in the lower CP of (42), like in Belfast English. As of now, I am not sure how the parametric variation is fleshed out in a substantive way. Assuming that relevant prosodic effects (e.g. *do*-support, rising intonation, to name a few) are limited to matrix clauses by the language-specific morpho-phonological constraint in Standard English, however, it can be worthwhile to state that the relevant interrogative sentence in (42) does not violate the constraint, since even if *wh*-subjects move to SPEC-CP, the local requirement for Affix Hopping can be satisfied via T-to-C raising, thereby *do*-support will not occur.

[9.] Note in passing that the present proposal can be compatible with the suggestion of Pesetsky (2021) that subject extraction requires the reduction of CP clause, since the *wh*-subject is extracted from within a <φ, φ> labeled set, not from within a CP clause (headed by over *that*).

[10.] It should be noted that Belfast English allows T-to-C raising in embedded clauses if there is no overt *that* as in (i), where the *wh*-object moves to the matrix SPEC-CP.

(i) a. *Who did John claim that did he see?
b. Who did John claim did he see?
c. Who did John claim that he saw? (Henry (1995: 114))

Although this phenomenon is worthwhile, I focus on Standard English in this paper and leave this topic for future research. I would just like to say that this infeasibility of subject-auxiliary inversion can be compatible with the suggestion of the present proposal that T-to-C raising cannot be applied if the head C position is occupied by overt *that*.

References

Bošković, Željko (2016) "On the Timing of Labeling: Deducing Comp-Trace Effects, the Subject Condition, the Adjunct Condition and Tucking in from Labeling," *The Linguistic Review: Special Issue on Labeling* 33, 17–66.

Bošković, Željko (2019) “Generalized Asymmetry,” ms., University of Connecticut.

Bruening, Benjamin (2015) “Subject Auxiliary Inversion,” ms. University of Delaware.

Chomsky, Noam (1986) *Barriers*, MIT Press, Cambridge, MA.

Chomsky, Noam (2013) “Problems of Projection,” *Lingua* 130, 33–49.

Chomsky, Noam (2015) “Problems of Projection: Extensions,” *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann and Simon Matteini, 3–16, John Benjamins, Amsterdam.

Chomsky, Noam (2021) “Minimalism: Where Are We Now, and Where Can We Hope to Go,” *Genko Kenkyu* 160, 1-41.

Culicover, Peter (1993) “Evidence against ECP: Accounts of the That-t Effects,” *Linguistic Inquiry* 24, 557–561.

den Dikken, Marcel and James Griffiths (2018) “Ellipsis and Spec-Head Agreement,” ms., Hungarian Academy of the Seciences and Lieden.

Diercks, Michael (2010) *Agreement with Subjects in Lubukusu*, Doctoral dissertation, Georgetown University dissertation, Washington, D.C.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2012) “Structure Building That Can’t Be!” *Ways of Structure Building*, ed. by Myriam Uribe-Etxebarria and Vidal Valmala, 253–270, Oxford University Press, Oxford.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2016) “Phase Cancellation by External Pair-Merge of Heads,” *The Linguistic Review* 33, 87–102.

Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely (2017) “Successive Cyclic Wh-movement without Successive Cyclic Crashing,” *Conference paper, 10th International Spring Forum of the English Linguistic Society of Japan*, Meiji Gakuin University, Tokyo, April 23.

Erlewine, Michael (2016) “Anti-Locality and Optimality in Kaqchikel Agent Focus,” *Natural Language and Linguistic Theory* 34, 429–479.

George, Leland (1980) *Analogical Generalizations of Natural Language Syntax*, Doctoral dissertation, MIT.

Hayashi, Norimasa (2020) “Labeling without Weak Heads,” *Syntax* 23, 275―294.

Hein, Johannes (2018) “On the Interaction of Head Movement, Ellipsis and Copy Deletion: The

Case of Mainland Scandinavian," Handout of a Talk Given at the 33rd Meeting of the Comparative Germanic Syntax Workshop, Georg-August Universität Göttingen.

Henry, Alison (1995) *Belfast English and Standard English*, Oxford University Press, Oxford.

Ishii, Toru (2004) "The Phase Impenetrability Condition, the Vacuous Movement Hypothesis, and *That*-t Effects," *Lingua* 114, 183-215.

Landau, Idan (2020) "Constraining Head-Stranding Ellipsis," *Linguistic Inquiry* 51, 281-318.

Lasnik, Howard (1999a) "On Feature Strength: Three Minimalist Approaches to Overt Movement," *Linguistic Inquiry* 30, 197–217.

Lasnik, Howard (1999b) "Verbal Morphology: Syntactic Structures Meets the Minimalist Program," *Minimalist Analysis*, ed. by Howard Lasnik and Mamoru Saito, 97–199, Blackwell, Malden MA.

McCloskey, James (2000) "Quantifier Float and Wh-Movement in an Irish English," *Linguistic Inquiry* 31, 57-84.

Merchant, Jason (2001) *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*, Oxford University Press, Oxford.

Merchant, Jason (2008) "Variable Island Repair under Ellipsis," *Topics in Ellipsis*, ed. by Kyle Johnson, 132-153, Cambridge University Press, Cambridge.

Messick, Troy (2020) "The Derivation of Highest Subject Questions and the Nature of the EPP," *Glossa* 5, 1–12.

Mizuguchi, Manabu (2008) "Derivation, Minimalism, and *That*-Trace Effects," *English Linguistics* 25, 56-92.

Mizuguchi, Manabu (2014) "Phases, Labeling, and Wh-movement of the Subject," Presented at the 32nd Conference of the English Linguistic Society of Japan, Gakushuin University, Tokyo, November 8th.

Mizuguchi, Manabu (2017) "Labelability and Interpretability," *Studies in Generative Grammar* 27, 327-365.

Obata, Miki (2010) *Root, Successive-Cyclic and Feature-Splitting Internal Merge: Implications for Feature-Inheritance and Transfer*, Doctoral dissertation, University of Michigan, Ann Arbor.

Pesetsky, David (2021) "Exfoliation: Towards a Derivational Theory of Clause Size," *Lingbuzz*,

Available at: https://lingbuzz.net/lingbuzz/004440.

Richards, Mark (2007) “On Feature Inheritance: An Argument from the Phase Impenetrability Condition,” *Linguistic Inquiry* 38, 563–572.

Rizzi, Luigi (1982) *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, Luigi and Ur Shlonsky (2007) “Strategies of Subject Extraction,” *Interfaces + Recursion =Language?*, ed. by Uli Sauerland and Hans-Martin Gärtner, 115-160, Mouton de Gruyter, Berlin.

Seely, T. Daniel (2006) “Merge, Derivational C-Command, and Subcategorization in a Label-Free Syntax,” *Minimalist Essays*, ed. by Cedric Boeckx, 182–217, John Benjamins, Amsterdam.

Tanigawa, Shin-ichi (2017) “In Defense of the Vacuous Movement Hypothesis for Wh-Subjects: Perspectives from the Framework of Chomsky (2013, 2015),” *Studies in English Literature English* 58, 57-76.