This paper considers that-t(race) effects in the Minimalist Program and proposes a principled explanation. It is claimed that that-t effects follow from the fact that the EPP feature in T cannot be satisfied at any point in the derivation. It is shown that the proposed analysis falls out from independently motivated theoretical assumptions and is empirically supported in that it correctly accounts for grammatical and ungrammatical instances of that-t effects from a number of languages, which also reveals that parametric variations with that-t effects are properly accommodated. We also discuss two theoretical implications for the Minimalist Program, both of which resonate with recent arguments in the Minimalist literature.*

Keywords: that-trace effects, the EPP feature, derivational computation, local relations, parametric variations

1. Introduction

Subjects and objects often show asymmetries in a number of empirical domains. One such asymmetry is found in extraction out of embedded clauses. This extraction asymmetry between subjects and objects (so-called that-t(race) effects) is illustrated in (1) and (2):

(1) Who do you think [(that) t will visit Mary]?
(2) Who do you think [(that) John will visit t]?

As the above data show, subject wh-movement is sensitive to an overt com-

* This is a revised and expanded version of my paper, which was originally given orally at the 24th National Conference of the English Linguistic Society of Japan, held at the University of Tokyo and which subsequently published as Mizuguchi (2007). I am indebted to Professor Kinsuke Hasegawa, Professor Hidekazu Suzuki, and anonymous EL reviewers for invaluable comments and helpful suggestions on earlier versions of this paper. Especially, I would like to express my sincere gratitude to Professor Suzuki for helping me at every stage up to the preparation of this paper. I would also like to thank Kazuhiro Yurugi and the participants in the conference for helpful comments. Needless to add, all remaining inadequacies are mine.

© 2008 by the English Linguistic Society of Japan
plicenter in C and a wh-subject can be extracted out of embedded clauses only when that is absent, at least in Standard English. On the other hand, objects can be wh-moved regardless of that in the embedded clause. The observed asymmetry has attracted much attention in the literature and a number of proposals have been put forth to account for this curious empirical fact. These proposals can be said to be descriptively adequate and can account for that-t effects at the level of explanation attained by the general theoretical assumptions under which they are formulated. In terms of the Minimalist Program (MP), however, these proposals are not true explanations. Put differently, that-t effects have so far defied any principled explanation.

Our aim in this paper is to take a fresh look at that-t effects in the MP and to give a principled account to this subject-object asymmetry. We will claim that that-t effects observed in (1) follow as one consequence of the derivational computation envisioned in the MP, hence proposing a derivational solution. We will show that that-t effects are naturally explained in the derivation without any unmotivated, stipulatory assumptions, demonstrating that a principled explanation in the sense of the MP is given to that-t effects.

This paper has the following organization. In section 2, we will briefly review two major previous accounts of that-t effects as our background and see why a new proposal is motivated. In section 3, we will propose to explain that-t effects in terms of derivational computation in the MP and claim that that-t (or more generally, overt Comp-trace) effects come from the unsatisfied EPP feature in T. In section 4, we will validate our proposal by showing that it can explain a number of empirical cases from various languages and that parametric variations with that-t effects are accommodated. In section 5, we will discuss two theoretical implications for the MP. In section 6, a conclusion of the paper will be presented.

1 This qualification is necessary, as there are dialects of English where that-t is not excluded (Pesetsky (1981/1982) and Sobin (1987)). We focus on Standard English. But see footnote 17 for one possible way to deal with the dialectal difference under our proposal.
2. Previous Approaches to That-t Effects

2.1. A Filter-based Approach

Chomsky and Lasnik (1977) propose a filter-based approach and explain the ungrammaticality of (1) with the "that-trace filter." The that-trace filter is stated as follows:

\[(1) \text{Is, that } \{\text{Np e} \ldots \}, \text{unless S' or its trace is in the context: } [\text{NP NP } \text{...}]\] (Chomsky and Lasnik (1977: 451))

Since the embedded clause in (1) has a that-t sequence as defined in (3), the data is ruled out (or "filtered out"). The ungrammaticality of (1) is captured with this filter.

It may be true that the that-trace filter excludes (1). But the filter itself is problematic. Conceptually, a filter is stipulated as such and merely redescribes the problem to be solved. In other words, it itself needs to be explained. Empirically, this filter is too strong and makes wrong predictions. There are languages where the that-trace filter does not apply. One such language is Italian (Rizzi (1982)). Consider (4):

\[(2) \text{Chii credi } [\text{cp the ti ama Sophia Loren}]? \]
who think-you that loves Sophia Loren

'Who do you think loves Sophia Loren?'

( Uriagereka (1988: 245))

Chomsky and Lasnik argue that the that-trace filter is universal. With languages like Italian taken into account, however, the filter is empirically problematic. But suppose that the that-trace filter is not universal but is a language-particular filter to account for languages like Italian. If it is language-specific, however, it adds another conceptually undesirable problem and is all the worse because languages are considered to be uniform in the underlying computation (Chomsky (2001)).

For these conceptual and empirical reasons, it can be concluded that a filter-based approach has problems and is far from a true explanation.

2.2. ECP-based Approaches

Another mode of theoretical explanation of that-t effects, which was much entertained in the Government and Binding (GB) model of grammar, is to use the ECP (Empty Category Principle). A number of ECP-based approaches have so far been proposed to that-t effects (Chomsky (1981, 1986), Browning (1996), Kayne (1984), Lasnik and Saito (1992), Pesetsky (1981/1982) and Rizzi (1990)). The ECP is a principle that warrants well-formedness on traces and the key notion of the ECP (in the explanation of
that-\(t\) effects) is proper government, which is either lexical (or theta) government or antecedent government. The crux of ECP-based accounts of that-\(t\) effects is that a subject trace in [Spec, IP] is prevented from being linked to an intermediate trace in [Spec, CP] and is not licensed when that is present, which violates the ECP (... \([\text{CP} ti [C' that [\text{IP} ti ...]])\). On the other hand, as for an object trace, its well-formedness is locally warranted by the verb that selects an object and no ECP violation is incurred. Hence that does not affect the grammaticality of such examples as (2).

ECP-based approaches to that-\(t\) effects can be said to be more explanatory in that government is regarded as a general mechanism in the GB model. Once they are subject to Minimalist scrutiny, however, we immediately find that they are conceptually problematic, even though they may indeed explain the data in (1) and (2) in a general manner. An important question arises on the status of the theoretical concept that is crucially employed: (proper) "government." Given the Minimalist conception of language design, government is suspect for its arbitrary nature because it is motivated neither by interface conditions nor by efficient computation considerations (the Strong Minimalist Thesis (Chomsky (2000))). In terms of the MP, it is also stipulated and is just a quasi-assumption to fill up the lack of understanding. The ECP is misconceived and explanations based on the ECP are misconceived as well. Then, ECP-based accounts of that-\(t\) effects are not principled explanations. To greater or lesser degrees, ECP-based approaches are no different from a filter-based approach in the perspective of the MP.

To summarize, in this section we have discussed two previous approaches to that-\(t\) effects that have been proposed in the literature and have seen that they are not true explanations. With this background, in the next section, we will rethink that-\(t\) effects in a Minimalist setting and propose a principled explanation along the research guidelines of the MP, which we will see is free from unmotivated assumptions and accommodates variations with that-\(t\) effects.

3. A Minimalist Approach to That-\(t\) Effects

ECP-based accounts of that-\(t\) effects suggest that there is a locality effect between a subject trace in [Spec, IP] and an intermediate trace in [Spec, CP] when that is present. In the GB model, this locality effect is captured representationally with traces and government. In the MP, on the other hand, locality is captured in the process of derivations. Then if that-\(t\) ef-
fects follow derivationally, they are given a principled explanation in the sense of the MP.

In the Minimalist model of grammar, it is assumed that derivations proceed step by step through local computation of formal (uninterpretable) features. Given this mode of syntactic derivations, then a hypothesis that is logically derived in the explanation of that-t effects is that some formal feature cannot be satisfied in the derivational process, which causes the derivation to crash at the interfaces and leads to that-t effects.

In this paper, based on this Minimalist hypothesis, we will claim that the formal feature that sets off crash is the EPP feature in T and propose an EPP analysis of that-t effects. (5) will be advocated in this paper:

(5) That-t effects show up due to the fact that the EPP feature in T cannot be satisfied in the derivation.

In the rest of this section, we will first spell out independently motivated theoretical assumptions that are fundamental to our theoretical proposal. Then we will discuss how our analysis works and show that that-t effects follow as one consequence of derivational computation.

3.1. Preliminary: Some Theoretical Assumptions

In this subsection, we will introduce the following three independently motivated theoretical assumptions as a preliminary to our discussion.

The first assumption is on (the nature of) the EPP feature. We assume that this feature is one property of the faculty of language and is part of its design specifications. Suppose that the content of the EPP feature in T is a D-feature (Chomsky (1995)) and is a selectional feature that triggers Merge/Move of an element with this feature (Chomsky (2000) and Lasnik (2001)). This assumption is reasonable in that a DP is a typical syntactic

---

2 The EPP feature has various names in the recent literature (see Chomsky (2000) and subsequent work). We adopt the “EPP” in this paper because the terminology is sufficiently familiar and because we focus on the EPP feature in T, a traditional instance, even though the Projection Principle no longer exists in the MP and the EPP has been generalized.

There is much recent research that questions the EPP (feature) as a principle in the grammar (Bošković (2002, 2005) and Epstein and Seely (2006) among others). At the moment, however, it is still controversial whether this conclusion is correct, since there is as much argument for the EPP (Lasnik (2001, 2003)); the jury is still out on this. In this paper, we assume that the EPP feature plays a role in syntactic derivations. To the extent that our proposal is justified, it can be said that the EPP is endorsed. I thank an anonymous reviewer for EL for directing me to arguments against the EPP.
category that is Merged/Moved to [Spec, TP] and is identified as a subject.  

The second assumption is that the EPP feature is satisfied as a result of Merge/Move through the formation of local relations with a T head. More specifically, local relations that derivationally fall out from Merge/Move are either Spec-to-head or head-to-head. Alexiadou and Anagnostopoulou (1998) argue that XP-movement to [Spec, TP] (so-called subject raising) is not the only way to satisfy the EPP feature in T. They suggest that the computational system, CHL allows two options to satisfy it: the first is to Merge/Move a phrasal element with a D property (most typically, a DP) to [Spec, TP] as conventionally assumed (= (6a)); the other is to Merge/Move a “head” (or more specifically, verb) to T (= (6b)), which is because rich verbal agreement has exactly the same status as pronouns and has a D-feature as one of its properties, just like DPs.  

Merge/Move can create the following configurations due to the EPP feature in T:

\[
\text{(6) a. } [\text{TP DP}_+ +_D \ T_{\text{EPP}}] [vP \ldots v [vP \ldots]] \text{ (Merge/Move XP)}
\]

\[
\text{b. } [\text{TP} T_{\text{EPP}}-V_{-D}V_{+D}] [vP \ldots t_j [vP t_i \ldots]] \text{ (Merge/Move X}^0\text{)}
\]

Under Bare Phrase Structure Theory (Chomsky (1995)), where bar levels are invisible to computation, Alexiadou and Anagnostopoulou’s suggestion is given a natural explanation. In both cases of (6), XP and X\(^0\) have the same local relation with an EPP-bearing T head because there is no element \(\gamma\) such that XP/X\(^0\) c-commands \(\gamma\) and \(\gamma\) c-commands T due to the invisibility of bar levels (T'): both XP and X\(^0\) c-command T in the same manner. Spec-to-head and head-to-head naturally fall into one and the same local relation. In addition, in both (6a) and (6b), Merge/Move is triggered and structure is built up as a result of the EPP feature. The EPP feature in T can thus be satisfied either by XP-movement (or XP-Merger) to [Spec, TP] or by head movement (or head Merger) to T.

Finally, we assume a theory of derivation proposed in Chomsky (2005, 2007), where the basic unit of derivations is a syntactic object called “phase” projected from C or transitive v (v\(^*\)) and derivations proceed phase by phase. Under this derivational model, computational properties are restricted to phase heads and T, a non-phase head, is devoid of any feature in

\[^3\text{Categories other than DPs that are regarded as “subjects” (or that are in [Spec, TP]) have been shown to be dominated by an abstract DP (Bresnan (1994)) or to be featurally similar or identical to DPs (Matsuyama (2001)).}\]

\[^4\text{The assumption is not unreasonable. A common assumption in the literature is that subject agreement inflections have developed from the reanalysis of subject pronouns (Bresnan and Mchombo (1987) and Hopper and Traugott (1993) among others).}\]
the lexicon. Instead, T’s computational properties (i.e. uninterpretable φ-features and the EPP feature) and tense property belong to C that selects it, which is a phase head, and are inherited to T from C in the derivation when C is Merged to TP (a “feature inheritance theory,” in which features are inherited from phase heads to non-phase heads):

\[
(7) \quad [\text{CP} \ C_{\{\phi/EPP/Tense\}} \ [\text{TP} \ T_{\{\phi/EPP/Tense\}} \ [\text{vP} \ \text{Subj} \ v \ [\text{vP} \ V \ \text{Obj}]]]]
\]

Inheritance of features from C to T

This theoretical assumption is not unreasonable: it has been argued in the literature that it is only when C is present that T manifests its Case/agreement properties and functions as a locus of the Case and agreement system; otherwise, raising/ECM structures come out (Collins (1993), Hiraiwa (2001), Pesetsky and Torrego (2001), Watanabe (1993) among others). This suggests that the locus of Case/agreement properties is not T but C, and T derivationally acquires its feature-checking properties from C as shown in (7).

With the above independently motivated theoretical assumptions in place, first consider (8) as our example and see how its derivation goes, which is summarized in (9) below:

(8) Who likes Mary?

(9)  a. \([\text{vP} \ who \ v \ [\text{vP} \ \text{likes} \ \text{Mary}]]\)
   b. \([\text{TP} \ T \ [\text{vP} \ who \ v \ [\text{vP} \ \text{likes} \ \text{Mary}]]]\)
   c. \([\text{CP} \ C_{\{A/A'\}} \ [\text{TP} \ T \ [\text{vP} \ who \ v \ [\text{vP} \ \text{likes} \ \text{Mary}]]]]\)  \text{Agree}(C, who)
   d. \([\text{CP} \ who; \ C_{\{A/A'\}} \ [\text{TP} \ T_{\{EPP\}} \ [\text{vP} \ t_i \ v \ [\text{vP} \ \text{likes} \ \text{Mary}]]]]\)  \text{Merge}(C, who)
   e. \([\text{CP} \ who; \ C_{\{A/A'\}} \ T_{j\{EPP\}} \ [\text{TP} \ t_j \ [\text{vP} \ t_i \ v \ [\text{vP} \ \text{likes} \ \text{Mary}]]]]\)  \text{Merge}(C, T)

The first derivational stage relevant to computation is when C is Merged to TP (= (9c)). Note that no computation (feature checking) is induced when T is Merged to vP at (9b), for it is featurally empty as we have discussed. At (9c), the Merged C has φ-features and the EPP feature as well as those features responsible for wh-movement. For convenience, we collectively call the former “A”-features and the latter “A’”-features. At this stage, C Agrees with who, simultaneously values these features, and

---

5 In this paper, v is used for both transitive v (v*) and non-transitive v for expository purposes, since our discussion does not depend on the distinction.
(internally) Merges who to [Spec, CP] upon this Agreement. But the EPP feature, since it is a selectional feature, cannot be satisfied via Agree and remains in the derivation. As a result, unlike $\phi$-features, it is inherited to T as shown in (9d). This inherited EPP feature in T, however, cannot be satisfied by internally Merging (Moving) who to its Spec since the wh-phrase has already been Merged to [Spec, CP] as a result of Agreement with the C head. Instead, it is argued that $C_{HL}$ employs the other option to satisfy the EPP feature: that is, a head-to-head local relation (cf. (6b)). Since English does not have rich verbal agreement unlike Italian, verbs do not have a D property that satisfies the EPP feature through a head-to-head relation with T. We suggest that the EPP feature inherited onto T can be satisfied by its movement to C (T-to-C movement) and forming a head-to-head relation with that head, just like V-to-T movement (see (10)). This suggestion is not unreasonable in that C, like D, is a category that has a D property, which is evidenced by the fact that CPs, just like DPs, can raise to [Spec, TP] to sat-

---

6 In the derivation of subject wh-movement like (8), we assume, contrary to Chomsky (2005, 2007), that $\phi$-features are not inherited to T but that they are valued together with A'-features in a long distance manner when C Agrees with a wh-subject. This mode of feature valuation of $\phi$-features without inheritance to T is reasonable since A'-features and $\phi$-features reside in the same head ("Maximize matching effects" (Chomsky (2001: 15))).

One reviewer raises a question on the nature of "Maximize matching effects." We answer it by saying that the feature maximization is considered as one corollary derived from the economy principle, which, in turn, follows from the Strong Minimalist Thesis. We define "economy principle" as follows (Chomsky (1995, 2004)):

(i) No superfluous computation is allowed in narrow syntax.

Given (i), the inheritance of $\phi$-features from C to T in (8) is superfluous and goes against (i) because C Agrees with who, which is the goal of $\phi$-features as well as A'-features, and can value both of these features with a single probing. On the other hand, the EPP feature, since it is a selectional feature, cannot be satisfied via Agree and its inheritance to T is not superfluous and does not violate (i); in fact, it is the only way for the EPP feature to be satisfied.

7 On the assumption that movement is motivated by a generalized form of the EPP feature, it follows that there are two EPP features in C: one is associated with A-features ($\phi$-features), a traditional instance of the EPP, and the other is associated with A'-features. We assume that the movement of who at the stage of (9d) satisfies only one of the two EPP features, as a result of which the other remains unsatisfied in the derivation. Since the EPP feature, as we have argued, is a selectional feature, each selection of the EPP feature triggers Merge/Move.
isfy the EPP feature in T. Consider (11): 8, 9

(10) \[ cp \ C_{\{+D\}} \rightarrow T_{\{EPP\}} \rightarrow t_i \rightarrow [vP \ldots] \] (cf. (6b))

(11) a. \[ TP \rightarrow [cp \ C\rightarrow T_{\{EPP\}} \rightarrow t_i \rightarrow [vP \ldots] \] believed \( t_i \) by everyone.

b. \[ TP \rightarrow [cp \ C\rightarrow T_{\{EPP\}} \rightarrow t_i \rightarrow [vP \ldots] \] seems to appear to be \( t_i \) obvious to Mary.

T-to-C movement (head Merger) satisfies the inherited EPP through a local head-to-head relation and the derivation of (8) is successfully computed as a convergent derivation (= (9e)).

Thus, in the derivation of subject \( wh \)-movement, the EPP feature in T is satisfied not by subject movement (raising) to [Spec, TP] but by T-to-C movement; a \( wh \)-subject moves in one fell swoop from [Spec, vP] to [Spec, CP] without moving to [Spec, TP]. In other words, contrary to the standard assumption, \( A' \)-movement is not from a derived \( A \)-position (that is, [Spec, TP]). Three pieces of empirical evidence for this derivation are considered. The first is another subject-object asymmetry in \( wh \)-movement observed in the following data:

(12) a. Who likes Mary?

b. *Who does like Mary? (non-emphatic)

(13) a. Who does John like?

b. *Who John likes?

In matrix \( wh \)-questions, do-support is required only for a non-subject \( wh \)-phrase. Suppose that affixes are combined with verbs under PF adjacency

---

8 Some researchers (Koster (1978) and Stowell (1981) among others) argue that CPs are not in [Spec, TP], hence are not true subjects. Instead, they argue that such CPs are topics. But this claim is empirically wrong. To see this, consider (i) (Bošković (2005)):

(i) a. *To John, that book, Mary gave.

b. To me, that John likes Mary seems obvious.

cf. c. To me, John seems happy.

(ia) shows that multiple topicalization is impossible. If CP subjects are not true subjects but topics, (ib) would be ruled out just like (ia) for multiple topicalization. Since (ib) is perfect just like (ic) and makes a clear contrast with (ia), it can be concluded that CP subjects are true subjects, not topics. Therefore, they have a \( D \) property and can satisfy the EPP feature in T.

9 One reviewer asks why CPs can be free from Case assignment if \( C \) has a \( D \) property; a common assumption is that a \( D \) property implies Case. We do not have an illuminating answer here but only suggest that Case is associated with a category \( N \), not a category \( D \). This assumption is not implausible given that there are languages like Japanese which lack a functional category \( D \) (Fukui (1986)) but whose NPs arguably have abstract Case.
in English (Lasnik (1995b)). Suppose further that the PF adjacency required for merging verbs with affixes is disrupted by an element iff it has Case (Jaeggli (1980), Bobaljik (1995)). This means that a head of A-chain disrupts PF adjacency and blocks affixal merger in the PF component, which is evidenced by (14). The data illustrate to contraction, which Jaeggli argues requires PF adjacency:

(14) a. *Who do you wanna be outside with you?
    b. *I wanna John be outside with me.

In (14), to contraction is blocked because there is a head of A-chain between want and to in the PF representation (the trace of who in (14a) and John in (14b)).

With this in mind, the asymmetry between (12) and (13) suggests that who does not move to [Spec, TP] in (12); if a subject wh-phrase moved to that position, PF adjacency between C, where affixes reside in matrix questions, and the verb would be disrupted just like (14) because the trace/copy in [Spec, TP] would be a head of A-chain, and do-support would be required to save the stranded affixes in C. The contrast in (12) and (13) shows that there is no head of A-chain in [Spec, TP] in subject wh-movement. In other words, a wh-subject moves to [Spec, CP] from [Spec, vP] in one fell swoop and [Spec, TP] is not created.

The second is an interesting asymmetry between A-movement and A'-movement as regards quantifier float discussed in McCloskey (2000). Consider the following contrast:

(15) a. *They were arrested all last night.
    b. Who was arrested all in Duke Street?

It is well-known that quantifiers stranded in the object position are ungrammatical (Bošković (2001)). However, wh-movement is exceptional in this regard. McCloskey argues that raising to [Spec, TP] (i.e. A-movement) causes the ungrammaticality in (15a) because the NP associated with a quantifier first moves to [Spec, DP] in order to strand it. Since NP-movement within DP is considered as an instance of A'-movement (as [Spec, DP] is an A'-position), further movement to [Spec, TP], an A-position, will be “improper” movement (A-to-A'-to-A), which is illegitimate, as is evidenced by the ungrammaticality of English superraising in (16):

(16) *[TP John, seems [CP ti that it is certain [TP ti to ti win]]].

On the other hand, in (15b), if who moves to [Spec, CP], an A'-position, directly from within vP after undergoing A'-movement to [Spec, DP] to strand a quantifier, the movement will be “proper” movement (A-to-A') and the data is ruled in (see Bošković (2001) for an account of (15b) with heavy
NP shift). Accordingly, the grammatical example (15b), in contrast with the ungrammatical example (15a), suggests that subject wh-movement does not make a stopover at [Spec, TP], contrary to the standard assumption.

The third piece of evidence comes from languages like Yiddish. In Yiddish, an expletive es ‘it’ can be Merged to [Spec, TP] when a subject is wh-moved. Consider the following data:

(17) Ikh veys nit [CP ver [TP es iz gekumen]].

I know not who Expl is come

‘I don’t know who has come.’ (Diesing (1990: 68))

This data suggests that the inherited EPP feature in T is satisfied by Merging es to its Spec. Then it supports our argument that a wh-subject does not move to [Spec, TP] on its way but moves directly to [Spec, CP] from [Spec, vP]; if a trace or copy of a wh-subject were left behind in [Spec, TP] or the inherited EPP feature in T were satisfied by a wh-subject as standardly assumed, the expletive would be blocked from being Merged to the Spec position.

From these three pieces of evidence, it is concluded that a subject wh-phrase moves directly from [Spec, vP] to [Spec, CP] without making a stopover at [Spec, TP].

3.2. A Derivational Explanation of That-t Effects

With our preliminary discussion in 3.1, we now turn to our main argument in this paper and discuss how that-t effects are explained in the derivational framework we are assuming. Relevant examples are repeated below from section 1:

(1) a. *Whoi do you think [that ti will visit Mary]?
   b. Whoi do you think [ti will visit Mary]?
(2) a. Whoi do you think [that John will visit ti]?
   b. Whoi do you think [John will visit ti]?

In the derivation of long distance wh-movement such as (1) and (2), a wh-phrase must move successive-cyclically through phase edges (i.e. [Spec,
CP] and [Spec, vP]) for further access in the next higher phase due to the PIC (Phase Impenetrability Condition). To implement successive cyclic movement, we follow Chomsky (2000) in assuming that it is driven by P(eripheral)-feature with the EPP optionally assigned to phase heads. For convenience of our discussion, we consider P-feature as one instance of A'-features and represent it as “A’” in bracket representations.

Let us first consider (1) and see why that-t effects are observed in (1a). In the embedded clause of (1a), C has A-features (φ, EPP, Tense) as well as P-feature. Thus in light of features, the structure of the embedded clause of (1a) is the same as that of (8), and is represented as (18a). As we have discussed with (8), since the goal of these features is a subject wh-phrase who in [Spec, vP], at this derivational stage, the embedded C Agrees with who in both A-features and P-feature. As a result of Agreement with C in P-feature, the wh-subject is internally Merged to [Spec, CP], which is an intermediate movement for the PIC. On the other hand, the EPP feature, which cannot be satisfied via Agree unlike φ-features, is inherited to the embedded T in (1a) as in (8). This inherited EPP feature, however, cannot be satisfied through forming a local relation with C in this case unlike (8). In the derivation of (1a), T cannot move to C to create a head-to-head relation for the EPP satisfaction because C is lexically filled with an overt complementizer that, which blocks T-to-C movement (= (18b)):

\[
\begin{align*}
\text{(18)} & \quad \text{a. } [\text{CP that-C}_{(A/A')} [\text{TP will-T} [\text{vP who v vP visit Mary}]]] \\
& \quad \quad \text{Agree(C, who) + Merge(C, who)} \\
\text{b. } [\text{CP whoi that-C}_{(A/A')} [\text{TP will-T}_{(EPP)} [\text{vP ti v vP visit Mary}]]] \\
& \quad \quad \text{*T-to-C movement (Merge(C, T))}
\end{align*}
\]

That a lexically realized, overt complementizer blocks head movement to C is suggested by V2 (Verb Second) in German. In this language, verb movement to C, that is, V2 is excluded when daβ ‘that’ is present in C in the complement clause and a verb-final, non-V2 order appears in this case. Consider the following data:

\[
\begin{align*}
\text{(19)} & \quad \text{a. Ich weiß [CP daβ Peter hier wohnt].} \\
& \quad \quad \text{I know that Peter here lives} \\
& \quad \quad \text{‘I know that Peter lives here.’} \\
& \quad \quad \text{(McCray (1981: 205))} \\
\text{b. *Ich weiß [daβ Peter wohnt hier].} \\
\end{align*}
\]

Thus, T-to-C movement cannot be employed to satisfy the inherited EPP feature in (1a). Since English does not have rich verbal agreement, V-to-T

\[\text{11 We will discuss the PIC in section 5. See (52) in that section for its details.}\]
movement cannot be used to satisfy the EPP feature in T, either. Besides, the other option (that is, XP-movement) is also excluded; who has already Moved to [Spec, CP] as a result of P-feature Agreement with C.

The outcome is that the inherited EPP feature in T remains unsatisfied in the derivation in the embedded clause and the derivation crashes when the embedded CP gets Transferred at the matrix vP phase because the unsatisfied EPP feature is an illegitimate object at the interfaces. We claim that the ungrammaticality of (1a) is due to the failure of the inherited EPP feature to be satisfied in the derivation. Thus, that-t effects follow as one consequence of derivational computation (= (5)).

Now consider (1b). With our analysis of that-t effects, the grammaticality of this example is straightforwardly explained since the inherited EPP feature in T can be satisfied successfully through a head-to-head local relation thanks to T-to-C movement (= (20b)). In this instance, there is no overt element in C that blocks this movement unlike (1a). In the embedded clause, the derivation is computed as a convergent derivation. As a result, no that-t effects are observed. Consider the derivation of the embedded clause of (1b) in (20):

(20) a. $[\text{CP} C_{\{A/A'\}} [\text{TP} \text{will-T} [\text{VP} \text{who v} [\text{VP} \text{visit Mary}]]]]$

b. $[\text{CP} \text{who}_i C_{\{A/A'\}} [\text{TP} \text{will-T}_{\{\text{EPP}\}} [\text{VP} \text{t}_i \text{v} [\text{VP} \text{visit Mary}]]]$

c. $[\text{CP} \text{who}_i C_{\{A/A'\}}\text{-will-T}_{\{\text{EPP}\}} [\text{TP} \text{t}_j [\text{VP} \text{t}_i \text{v} [\text{VP} \text{visit Mary}]]]]$

That T-to-C movement is possible when an overt complementizer is absent is also suggested by German data. Unlike (19), when daβ is missing and C is null, a verb can move to C and V2 is in fact observed even in the embedded clause. Consider the following data:

(21) a. Sie glaubte $[\text{CP} \text{das Kind hatte das Brot gegessen}]$.

    she thought the child had the bread eaten

    ‘She thought that the child had eaten the bread.’

b. *Sie glaubte $[\text{CP} \text{das Kind das Brot gegessen hatte}]$.

The feasibility of T-to-C movement, which satisfies the inherited EPP feature in T, explains the grammaticality of (1b), in contrast with (1a), where this movement is impossible.

If, as we proposed, the inherited EPP feature holds the key to that-t effects, their absence in examples like (2), where a wh-object is extracted from the embedded clause, also straightforwardly follows from our proposal. Recall that such data are grammatical irrespective of an overt complementizer. In their embedded clause, C has A-features ($\varphi$, EPP, Tense) as well as P-feature and the embedded structure is the same as that
of (1) in terms of features that drive computation. In cases under consideration, however, the goal of A-features and the goal of P-feature are different. In other words, the embedded C does not Agree with the same single goal in A-features and P-feature. It does Agree with an object wh-phrase at the edge of the embedded [Spec, vP] in P-feature and attract it to its Spec upon this Agreement (= (22a)). On the other hand, A-features as a whole (i.e. not only EPP but also φ), since they cannot be satisfied through C’s Agreement with who, are inherited to the embedded T (= (22b)). Then the inherited EPP feature in T, along with φ-features, is satisfied by the embedded subject John in [Spec, vP]: T Agrees with John in φ-features, as a result of which the subject is internally Merged to [Spec, TP], which satisfies the inherited EPP feature through a Spec-to-head local relation (= (22c)). Unlike a subject wh-phrase, which moves to [Spec, CP] from [Spec, vP] upon its Agreement with the embedded C, a non-wh subject phrase stays in [Spec, vP] when the EPP feature is inherited to T together with φ-features, which means that T-to-C movement is not required for the EPP satisfaction; subject raising to [Spec, TP] satisfies the inherited EPP feature. Consider the following derivation:

\[
(22) \text{a. } [\text{CP (that)-C}_{\{A/A'\}} \text{TP will-T } [\text{vp who}_i \text{[John v [vp visit } t_i])]] \\
\text{Agree(C, who) + Merge(C, who)}
\]

\[
(22) \text{b. } [\text{CP who}_i (that)-C_{\{A\}} \text{TP will-T}_{\{A(\phi, \text{EPP})\}} \text{[vp } t_i \text{[John v [vp visit } t_i])]]
\]

\[
(22) \text{c. } [\text{CP who}_i (that)-C_{\{A\}} \text{TP John}_j \text{will-T}_{\{A(\phi, -\text{EPP})\}} \text{[vp } t_j \text{[v [vp visit } t_i])]] \\
\text{Agree(T, John) + Merge(T, John)}
\]

It follows that the movement of an object wh-phrase is insensitive to a lexically filled complementizer. Therefore, that-t effects do not appear in this case. The subject-object asymmetry is given a derivational explanation under our proposal in light of features that are essential to linguistic computation.

Our derivational explanation of that-t effects through the unsatisfied EPP feature also accounts for their absence in adjunct wh-movement as exemplified in (23) in the same way that it explains the absence in object wh-movement in (2): the inherited EPP feature can be satisfied by the movement of the subject John in [Spec, vP] to [Spec, TP] (subject raising), and T-to-C movement is not required for this purpose.

\[
(23) \text{a. } \text{How}_i \text{ do you think [that John will explain the fact } t_i]? \\
\text{b. } \text{How}_i \text{ do you think [John will explain the fact } t_i]?
\]

To summarize our discussion of that-t effects in (1), we have argued that they come out for the reason that the inherited EPP feature in T cannot be
satisfied at any point in the derivation, claiming that that-t effects are accounted for in terms of derivational computation (= (5)). The proposed analysis explains not only that-t effects in (1) but also why they are not at all observed in other instances of wh-movement (i.e. object and adjunct wh-movement), being specific to subject wh-movement. Since the essence of that-t effects is the unsatisfied EPP feature in T, government, which is central to the ECP, is unnecessary for the explanation of that-t effects.12

3.3. Other Instances of That-t Effects

In the last subsection, it was argued that that-t effects follow as one consequence of derivational computation. In this subsection, we will discuss other instances of that-t effects. It has been noted in the literature that that-t effects are observed not only in cases such as (1), where the embedded C is that; overt complementizers in general induce that-t effects (hence, that-t effects are more generally “overt Comp”-trace effects) (Bresnan (1977), Browning (1996), Culicover (1993) and Pesetsky (1981/1982) among others). Such complementizers include for, whether, if, like and a complex complementizer as if.13 Consider (24) (cf. (1)):14

(24) a. *Whoi were you hoping \([CP \text{ for } t_i \text{ to stay}]\)?
   b. *Whoi are you wondering \([CP \text{ whether/if } t_i \text{ will substitute for Professor Smith}]\)?

12 Thus, our proposal contributes to the elimination of government in the domain of the ECP (that-t effects), resonating with other studies that try to eliminate government from other empirical domains like locality, Case theory, and binding theory, along the Minimalist guidelines. See Appendix for some discussion of other Minimalist approaches to that-t effects.

13 We assume that whether is a head of C (see German data (26) below).

14 A major difference between that-t effects and other overt Comp-trace effects is that the absence of overt complementizers does not improve grammaticality. Consider (i):

   (i) a. *Whoi were you hoping \([t_i \text{ to stay}]\)? (cf. (24a))
   b. *Whoi are you wondering \([t_i \text{ will substitute for Professor Smith}]\)? (cf. (24b))
   Plausibly, these examples are ungrammatical for independent reasons: as for (ia), the Case feature of who will remain unvalued if for is absent and as for (ib), the absence of whether/if will not satisfy the selectional requirement of wonder.

   On the other hand, (iia) is grammatical.

   (ii) a. Whoi does it seem \([t_i \text{ lost the file}]\)? (cf. (24c, d))
   b. *Whoi does it seem \([\text{that } t_i \text{ lost the file}]\)?

   However, this example has different semantics and corresponds to (iib). The pair in (ii) is on a par with that in (1). We can then conclude that even if overt complementizers like and as if are absent, the data remain ungrammatical. Plausibly, the ungrammaticality is explained in the same manner as that of (ib).
c. *Who_1 does it seem [CP like t_1 lost the file]?
d. *Who_1 does it seem [CP as if t_1 lost the file]?
The ungrammaticality of (24) straightforwardly follows from our proposal
and is analyzed on a par with that of (1a). Crucial to our proposal is that
the embedded C is overtly filled, which precludes T-to-C head movement
for the satisfaction of the EPP feature inherited to T from a C head. In
(24), overt complementizers such as for are in C, blocking this movement
just like that in (1a). Furthermore, subject raising to [Spec, TP] cannot
be employed since a wh-subject moves in one fell swoop from [Spec, vP]
to [Spec, CP]. Therefore, the derivation of the embedded clause in (24)
proceeds in the way illustrated in (18). In other words, overt Comp-trace
effects, which include traditional that-t effects, fall into the general derivational
scheme in (25) and their ungrammaticality is attributed to the fact that
the inherited EPP feature in T cannot be satisfied in the derivation (“OC” in
(25) is “overt complementizer”):

(25) a. [CP OC-C_{A/A'} [TP T_{EPP} [vP Wh_{(Subject)} v [vP ... ]]]]

Agree(C, Wh) + Merge(C, Wh)

b. [CP Wh_{(Subject)}} OC-C_{A/A'} [TP T_{EPP} [vP t_1 v [vP ... ]]]

\*T-to-C movement (Merge(C, T))

No additional assumption is required to exclude (24), and overt Comp-trace
effects, including traditional that-t effects, are given a derivational explana-
tion under the proposed EPP analysis.

As in the case of that, German V2 shows that other overt complementiz-
ers also block head movement to C. For instance, consider an embedded
clause introduced by ob `whether.’ In this clause, V2 is excluded and a
non-V2 order is observed (cf. (19)):

(26) a. Ich weiß nicht [CP ob Peter hier wohnt].
I know not whether Peter here lives

‘I don’t know whether Peter lives here.’

b. *Ich weiß nicht [CP ob Peter wohnt hier].

(McCray (1981: 205))

These data show that ob blocks head movement to C and endorse our argu-
ment that the failure of T-to-C movement is the cause of ungrammaticality
of (24), along with (1a): overt Comp-trace effects are due to the unsatisfied
EPP feature inherited to T.

3.4. Relative Clauses and That-t Effects

Before leaving our discussion of that-t effects in wh-movement, we will
discuss that-t in relative clauses. As explicitly stated in the “unless” clause
in Chomsky and Lasnik's that-trace filter, the relativization of a subject does not show that-t effects. Consider (27):

(27) The man [CP OP, *(that) ti came here yesterday] is my brother.

Interestingly, as shown in (27), that must be present in subject relativization. If that in relative clauses is an overt complementizer and a null operator movement is involved in (27) as is commonly assumed, the example would be ruled out under our proposal: since a subject null operator moves in one fell swoop from [Spec, vP] to [Spec, CP] within the relative clause, T-to-C movement will be required for the EPP satisfaction. This movement, however, is blocked by that, just like (1a) and (24).

To explain the absence of that-t effects in (27), we suggest that that in relative clauses is a non-agreeing variant of agreeing relative pronouns such as who and which, instead of a complementizer. This suggestion is not implausible given two facts from English. The first is that in languages like English, that also works as a demonstrative pronoun. Consider (28):

(28) [DP That] is John's car.

Given this use of that in the English lexicon, it is not unreasonable to assume that that employed in relative clauses is a relative pronoun that originates from a demonstrative pronoun that, just like a relative pronoun who originates from an interrogative pronoun who.

The second is that at least in languages like English, that and relative pronouns like who are complementary as shown in (29) (traditionally known as doubly-filled Comp):

(29) *The man [who that came here yesterday] is my brother.

This fact is straightforward if that is a relative pronoun along with who, which and the like in relative clauses and they compete for the same structural position (that is, [Spec, CP]).

Given this not unreasonable assumption that that is a relative pronoun, the absence of that-t effects in relative clauses is not problematic but is in fact expected. Since there is no overt C, T can move to C and the inherited EPP feature in T can be satisfied through a T-to-C head relation (see (30)). Thus, examples like (27) are analyzed on a par with (1b):

---

15 But this suggestion, admittedly, is tentative because there are still important questions to be answered regarding the differences between that and wh-relatives, historical changes, and parametric variations. Here, we leave these discussions for future research.
DERIVATION, MINIMALISM, AND THAT-TRACE EFFECTS

4. Variations with That-t Effects

In the last section, we claimed that that-t effects follow derivationally from the fact that the inherited EPP feature cannot be satisfied in the derivation. An overt complementizer blocks T-to-C movement, which disenables the EPP feature in T from being satisfied by means of a head-to-head relation with C. Given that the culprit for that-t effects is the unsatisfied EPP feature in T in the derivation, we should consider how the proposed derivational EPP analysis explains variations with that-t effects. It has been pointed out in the literature that that-t effects are not always observed and that variations are found across and within languages. Given our proposal, if a language allows a grammatical strategy other than T-to-C movement to satisfy the inherited EPP feature in T and the EPP feature can be satisfied in the derivational process, that-t effects will not be observed even when an overt complementizer is present in C and T cannot move to this C. The purpose of this section is to consider the empirical validity of the proposed analysis of that-t effects by showing that a number of grammatical cases of that-t effects from various languages fall under the analysis.

In fact, we have already seen two instances in which that-t effects are absent: object and adjunct wh-movement in cases like (2) and (23). We have argued that that-t effects are not observed in these cases because the inherited EPP feature in T is satisfied by non-wh subjects (that is, XP-movement to [Spec, TP] or subject raising). In this section, our focus is on those cases in which subject wh-phrases indeed move out of the embedded clause with an overt complementizer but that-t effects do not appear because the EPP feature is satisfied. We will discuss three empirical cases in this section.

4.1. XP Merger

The first empirical case is the one in which an XP other than a moved subject wh-phrase can be independently Merged to [Spec, TP]. In this case, the inherited EPP feature in T is satisfied by this Merger and if our proposal is correct, that-t effects must be absent in those languages that allow this XP Merger. One of such languages is Yiddish. Recall (17):

\[
\text{(30) } \left[\text{DP The man } \left[\text{CP that } \left[C' \ T_h \left(T_{\text{EPP}} \rightarrow \text{TP } t \right) \left[\text{VP } t \text{ came here yesterday} \right] \right] \left[\text{C } A/A^\alpha \right] \right] \right] \text{ is my brother.}
\]

T-to-C movement (Merge(C, T))
(17) Ikh veys nit [CP ver [TP es iz gekumen]].

I know not who Expl is come

‘I don’t know who has come.’ (Diesing (1990: 68))

In Yiddish, an expletive es ‘it’ can be Merged to [Spec, TP] when a subject is wh-moved, which endorses our argument that subject wh-movement proceeds directly from [Spec, vP] to [Spec, CP]. Given that expletive Merger is due to the inherited EPP feature in T, our proposal can explain the absence of that-t effects in a Yiddish example in (31a) below:

(31) a. ?Veri hot er moyre [CP az [TP es vet ti kumen]]?
    who has he fear that Expl will come

b. *Veri hot er moyre [CP az [TP vet ti kumen]]?
    ‘Who does he fear will come?’ (Diesing (1988: 137))

(31a) is well-formed. Since the inherited EPP feature in T is satisfied by the Merger of es, even if az ‘that’ is present in the embedded C and T-to-C movement is blocked, that-t effects do not appear in this example. On the other hand, (31b) is ill-formed for the same reason that English (1a) is ungrammatical: az prevents T from moving to C, which leaves the inherited EPP feature in T unsatisfied. The derivation crashes due to the fact that the inherited EPP feature in T remains unsatisfied in the derivation.

Likewise, that-t effects are absent when Yiddish shows V2 in the embedded clause. Diesing (1990) argues that V2 in Yiddish is at TP level and that a V2ed phrase is in [Spec, TP], with verb movement only up to T. Put differently, [Spec, TP] is an A'-position as well as an A-position and when a non-subject phrase is V2ed, a subject is anchored in [Spec, vP], its base-Merged position. We can interpret V2 in TP by assuming that V2 is formed due to the EPP feature inherited to T from C in such languages as Yiddish. Then even if az is present, the inherited EPP feature in T can be satisfied by a non-subject, V2ed phrase even without T-to-C movement. Hence, that-t effects will not be found in V2 contexts. Consider (32), where the object (ot di bikher) and an adverbial NP (haynt) are V2ed. In these examples, that-t effects are not observed and our proposal explains the grammaticality of (32):

(32) a. Veri hot er nit gevolt [CP az [TP ot di bikheri zol ti leyenen tj]]?
    who has he not wanted that the books should read
    ‘Who did he not want to read the books?’ (Diesing (1990: 75))
b. Veri hot er moyre [CP az [TP haynt vet t kumen]]? who has he fear that today will come  
‘Who is he afraid will come?’ (Diesing (1988: 138))

The data of the same sort are also found in French. It is well-known that French shows a complementizer alternation (so-called que-qui alternation) (Pesetsky (1981/1982) and Rizzi (1990)) when wh-movement applies. As the following examples illustrate, an overt complementizer, which corresponds to English that, must be qui instead of que when a subject is wh-moved. On the other hand, qui cannot appear when an object is wh-moved. Consider the following subject-object extraction pair:

(33) a. Quel livrei crois-tu [CP que/*qui les filles vont acheter ti]? which book think-you that the girls will buy  
‘Which book do you think that the girls will buy?’

b. Quelles fillesi crois-tu [CP *que/qui ti vont acheter ce livre-la ]? which girls think-you that will buy that book-there  
‘Which girls do you think will buy that book there?’ (Taraldsen (2002: 29))

This que-qui alternation in (33), just like Yiddish contrast in (31), is straightforward under our proposal that the inherited EPP feature in T holds the key to that-t effects. Taraldsen (2002) argues that French qui is actually analyzed as que followed by i, which is an expletive Merged to [Spec, TP] just like Yiddish es. The syntactic structure of (33b) is analyzed as follows:

(34) Quelles fillesi ... [CP que [TP i vontj [VP t ti tj acheter ce livre-la]]] (qui = que + i)

Then the grammatical contrast between que and qui in (33) is explained in the same manner as that in Yiddish (31) and straightforwardly follows from our proposal: qui must appear in subject wh-movement so that the inherited EPP feature in T can be satisfied through the Merger of an expletive i. In other words, (33b) with que corresponds to (31b) in Yiddish. On the other hand, qui is naturally ruled out in object wh-movement because the EPP feature in T is satisfied by the subject les filles ‘the girls’ through its movement to [Spec, TP] (see (22)); with no more EPP, the expletive cannot be Merged to [Spec, TP] in this case.

We have discussed Yiddish and French thus far and have argued that
when an XP (e.g. an expletive and a V2ed phrase) can be independently Merged to \([\text{Spec, TP}]\) and the inherited EPP feature in T is satisfied, *that*-t effects do not appear even when an overt complementizer is present. Other languages that are explained by our proposal, hence support the EPP analysis of *that*-t effects, include Swedish and Danish. Consider the following data:

(35) a. Vilken elevi trodde ingen \([\text{CP} \text{ att } [\text{TP} \text{ hani skulle fuska}]])? (Swedish)

   which pupil thought nobody that he would cheat
   ‘Which pupil didn’t anyone think would cheat?’

   b. *Vilken elev trodde ingen \([\text{CP} \text{ att } [\text{TP} \text{ skulle fuska}]])?

   (Engdahl (1982: 166))

(36) a. Vennen \([\text{CP} (\text{som}) \text{ han pastod } [\text{CP at } [\text{TP} \text{ der havde }

   friend-Def (that) he claimed that there had
   lant bogen ]]) \text{ var forsvundet.} \) (Danish)

   borrowed book-Def was disappeared
   ‘The friend that he claimed had borrowed the book had dis-

   appeared.’

   b. *Vennen \([\text{CP} (\text{som}) \text{ han pastod } [\text{CP at } [\text{TP} \text{ havde lant bogen}]])

   \text{ var forsvundet.} \) (Engdahl (1985: 21))

In Swedish (35a), a resumptive pronoun *han* is Merged to \([\text{Spec, TP}]\) in the embedded clause and in Danish (36a), an expletive *der* is Merged there. In both cases, unlike their counterparts in (35b) and (36b), the inherited EPP feature in T is satisfied thanks to the Merger of an XP, even though T-to-C movement cannot be employed for a lexical C. Hence, *that*-t effects are absent even with overt complementizers.

To summarize our first argument in this subsection, those languages we have considered here demonstrate that when the inherited EPP feature is satisfied by XP Merger, *that*-t effects are not observed. Given that XP Merger is due to the inherited EPP feature in T, our proposal, which attributes *that*-t effects to the unsatisfied EPP feature in T, can give a straightforward explanation to the absence of *that*-t effects in the languages in this subsection.

4.2. Layered CP

The second empirical case in point is the one in which an extra head is motivated between C with an overt complementizer and T to which T can move, and the inherited EPP feature can be satisfied by this movement through the formation of a head-to-head relation. Examples relevant to this case are provided by so-called adverb effects (Bresnan (1977), Browning
(1996), Culicover (1993) and Hasegawa (1993/1994) among others). An interesting fact about that-t effects is that they disappear when sentential adverbials come between a that-headed C and a trace/copy in the embedded clause. Consider the following examples:

(37)  
   a. Who did she say [that tomorrow ti would regret his words]?  
   b. Which doctor did you tell me [that during an operation ti had had a heart attack]?  
      (Bresnan (1977: 194, fn. 6))

(38)  
   a. Robin met the man who Leslie said [that *(for all intents and purposes) ti was the mayor of the city].  
   b. This is the tree which I said [that *(just yesterday) ti had resisted my shovel].  
      (Culicover (1993: 98))

Basically following Browning (1996), we assume that sentential adverbials motivate CP recursion in the embedded clause (see also Culicover (1993) and Rizzi (1997) for layered CP): that is in a higher C head and there is another C head that comes between a that-headed C and T as one participant of the clausal architecture, which hosts and licenses sentential adverbials (Cinque (1999)). The embedded clause of (37) and (38) is analyzed as follows:

(39)  
   [CP Spec that-C [CP [S-Adverbial] C [TP Spec T [vP ... ]]]]

With this analysis in mind, note that a lower C head motivated by sentential adverbials is not lexically filled and that T can move to this head for its EPP feature without any problem (T movement to a lower C). Then the inherited EPP feature in T can be satisfied through a head-to-head relation with this C head. Consider the derivation in (40):

(40)  
   a. [CP that-C_{A/A'} [CP tomorrow C [TP would-T [vP who ... ]]]]
   b. [CP who that-C_{A/A'} [CP tomorrow C [TP would-T_{EPP} [vP ti ... ]]]]

   T-to-C movement (Merge(C,T))
   c. [CP who that-C_{A/A'} [CP tomorrow C-would-T_j_{EPP} [TP ti [vP ti ... ]]]]

Given the structure (39), which is independently motivated by sentential adverbials, the inherited EPP feature in T can be successfully satisfied through T-to-C movement even if an overt complementizer that is present. Our proposal explains why data such as (37) and (38) are gram-
matical and why adverbials mitigate *that*-effects.\textsuperscript{16, 17}

In addition to English, the structure (39) is also independently motivated in Icelandic and the language provides us with the same kind of empirical evidence for our proposal. Icelandic is a typical V2 language and V2 in this language, unlike V2 in Yiddish, applies at CP level. Unlike German, V2 is always possible even in the embedded clause and it has been suggested in the literature that V2 in embedded clauses is analyzed in the same way as V2 in matrix clauses: that is, X°-movement to C and XP-movement to its Spec. Since V2 is available in the embedded clause even when *að* ‘that’ is present, it is reasonable to assume that CP is layered, with the overt complementizer being in a higher C and a V2ed phrase being in the Spec of a lower null C head. For instance, consider the following data:

(41) a. ... *að* bókina hafa börnin lesið í gær.
   ‘... that book-the have children-the read yesterday
   ‘... that the children read the book yesterday.’

b. ... *að* í gær hafa börnin lesið bókina. (Vikner (1995: 101))

One empirical indication of this layered CP structure in the embedded clause in Icelandic is that an expletive *pað* ‘there’ can be Merged even in the embedded clause:

(42) ...*að* pað hefur einhver boroða epli.
   ‘... that there has someone eaten apple
   ‘... that someone has eaten an apple.’ (Vikner (1995: 189))

\textsuperscript{16} An alternative analysis of the grammaticality of (37) and (38) under our proposal is that sentential adverbials are Merged in [Spec, TP], which satisfies the inherited EPP feature in T just like expletives in Yiddish and French (see also den Dikken (2006) for this analysis on independent grounds). One piece of empirical evidence which suggests this alternative analysis comes from Yiddish. As the following data shows, when an adverb is fronted, *es* is excluded, which implies that the adverb is in [Spec, TP] and the EPP feature in T is satisfied by the adverb:

(i) Zi iz gekumen zen ver frier (*es) vet kontshen.
   she is come see who earlier Expl will finish
   ‘She has come to see who earlier will finish.’ (Diesing (1988: 134))

But this analysis raises non-trivial questions and we must address them before it is fully viable. We just note this alternative here and leave the investigation of these questions for future. Our discussion is not affected by the decision between the two alternatives.

\textsuperscript{17} It is reported that there are dialects (or speakers) of English where *that*-effects do not appear even without sentential adverbials. One solution to this dialectal difference is that in relevant dialects (but not in Standard English), CP can be layered regardless of sentential adverbials. Then T can move to a lower C irrespective of *that* in a higher C and the inherited EPP feature in T can be satisfied through this head movement as illustrated in (40).
Unlike English *there*, the Icelandic counterpart is not Merged to [Spec, TP] but to [Spec, CP] for the purpose of V2 (Holmberg and Platzack (1995), Maling and Zaenen (1978), Thráinsson (1979)). The data in (42) shows that a layered CP structure is independently motivated in Icelandic, and the clausal architecture of its embedded clause is analyzed in the same way as that of (37) and (38), though for a different reason (that is, V2):

(43) \[ \text{CP Spec C-að [CP [V2ed phrase] C [TP Spec T [VP ... ]]]} \]

With (43), an important aspect of V2 for our discussion is that T moves to this lower C together with v-V for V2. If our proposal is on the right track, the inherited EPP feature in T can be satisfied for this head movement through a head-to-head relation with a lower C head. Then *that-t* effects should not appear even when an overt complementizer *að* is in a higher C head, as in the case of adverb effects in English. This is empirically corroborated. *That-t* effects are not observed in Icelandic and this fact falls under our proposal. Consider (44):

(44) Hveri sagðir þú að tði hefði borðað þetta epli?
    who said you that had eaten this apple
    ‘Who did you say had eaten this apple?’

(Maling and Zaenen (1978: 480))

To sum up our second empirical discussion, English data like (37) and (38) and Icelandic data like (44) demonstrate that when an independent head with a D property is motivated between an overt complementizer-headed C and T to which T can move to satisfy the inherited EPP feature by forming a head-to-head relation there, *that-t* effects are not observed even when *that* or its counterpart in other languages is present. If the source of *that-t* effects is the unsatisfied EPP feature in T in the derivation, as our proposal claims, its disappearance in the English and Icelandic data straightforwardly follows from the proposal.

### 4.3. V-to-T Movement

The final empirical case is the one in which the inherited EPP feature in T can be satisfied by verb movement, not by subject movement to [Spec, TP]. As we briefly introduced in 3.1, Alexiadou and Anagnostopoulou (1998) argue that rich verbal agreement has exactly the same status as pronouns and that the EPP feature in T can be satisfied by verb movement to this head, which leads to the other local relation with it: that is, a head-to-head relation (see (6b) above). They say that so-called null subject languages, one instance of which is Italian, are languages of this type. Then given our proposal, *that-t* effects will not appear in such lan-
guages because the inherited EPP feature in T is satisfied by verb movement to T, even though T-to-C movement is blocked for a lexically filled complementizer. In fact, it has been pointed out that null subject languages do not show _that-t_ effects. According to Alexiadou and Anagnostopoulou, Greek and Romance languages like Italian and Spanish are languages that opt for satisfying the EPP feature in T by verb movement, which is at least confirmed by a V-initial, surface order of these languages (VSO/VOS). _That-t_ effects are absent in these languages as expected. Consider the following data:

(45) Pjοςι nomizis [CP oti ti tilefonise]? (Greek)
    who think-2s that telephoned
    ‘Who do you think called?’  (Roussou (2002: 40))

(46) Chi credi [CP che ti ama Sophia Loren]? (Italian)
    who think-you that loves Sophia Loren

(47) Quién dijiste [CP que ti salió temprano]? (Spanish)
    who said-you that left early
    ‘Who did you say left early?’  (Perlmutter (1979: 103))

Our proposal can account for the absence of _that-t_ effects in null subject languages. Since the EPP feature in T can be independently satisfied by verb movement to T, T-to-C movement is not required for the EPP satisfaction and the presence of overt complementizers does not affect the grammaticality of subject extraction from the embedded clause.

To summarize our discussion in this section, we have considered the validity of our proposal by examining how grammatical cases of _that-t_ effects from various languages are explained under our proposal. Provided that the culprit for _that-t_ effects is the unsatisfied EPP feature in the derivation, as our proposal claims, if the inherited EPP feature is somehow satisfied, _that-t_ effects will be absent even when an overt complementizer is present in C and T-to-C movement is blocked. We have provided three empirical cases from various languages. As we have demonstrated, the languages we have considered here have strategies of EPP satisfaction other than T-to-C movement in embedded clauses and the absence of _that-t_ effects in these languages straightforwardly follows from our proposal. Since our proposal can deal with grammatical as well as ungrammatical instances of _that-t_ effects, hence their cross-linguistic and intra-linguistic variations, with independently motivated mechanisms alone, it can be concluded that our derivational analysis of _that-t_ effects with the unsatisfied EPP feature in T is valid.
Before we leave this section, we would like to discuss an empirical case from Yiddish that endorses our proposal in yet another way. As we have discussed above with (31b), that-t effects are observed in Yiddish. There is an interesting fact about this language that corresponds to the ungrammaticality of (31b). Consider the derivation of long distance wh-movement of an object. In Yiddish, when an object wh-phrase undergoes wh-movement, a subject moves to [Spec, TP] in the embedded clause just like English (2). Consider (48):

\[(48)\ Vosi\ hot\ er\ nit\ gevolt\ [CP\ az\ [TP\ mir\ zoln\ t\ leyenen\ t]]? \]

what has he not wanted that we should read

‘What did he not want us to read?’ \(\text{(Diesing (1990: 71))}\)

The grammaticality of (48) is straightforward and is analyzed on a par with English data like (2): the EPP feature in the embedded T is satisfied by the raising of mir ‘we.’ What is interesting about object wh-movement is that when a subject remains in [Spec, vP], either the expletive es must be Merged to [Spec, TP] as in (31a) (= (49a)) or an inflected verb in T has to move to C (T-to-C movement) (= (49b)). If neither takes place, ungrammaticality results (= (49c)). Consider the following paradigm:

\[(49)\]

\[(49a)\ Vosi\ hot\ er\ nit\ gevolt\ [CP\ az\ [TP\ es\ zoln\ mir\ leyenen\ t]]? \]

what has he not wanted that Expl should we read

‘What did he not want us to read?’ \(\text{(Diesing (1990: 71–72))}\)

\[(49b)\ Vosi\ hot\ er\ nit\ gevolt\ [CP\ zoln\ t\ leyenen\ t]\]

\[(49c)\ *Vosi\ hot\ er\ nit\ gevolt\ [CP\ az\ [TP\ zoln\ mir\ leyenen\ t]]? \]

\(\text{(Diesing (1990: 71–72))}\)

The data crucially relevant to our discussion is (49b). In this example, the subject mir remains in [Spec, vP] but the derivation is grammatical thanks to verb movement to C. Given that (48) and (49a) are grammatical thanks to XP Merger to [Spec, TP], the grammaticality of (49b) endorses our argument that the EPP feature in T can be satisfied through T-to-C movement for the reasons we have discussed. This is also suggested by (49c), where an overt complementizer az in C blocks T-to-C movement and XP Merger is required, which rules out an in-situ subject. It should be noted that under our proposal, the ungrammaticality of (49c) is explained in the same way as that-t effects: the EPP feature in T remains unsatisfied in the derivation. The factor that determines the well-formedness of the above Yiddish paradigm is the EPP feature in T, and our proposal not only explains that-t effects and (49c) in a unified manner but also accounts for why
T-to-C movement is required even in object wh-movement when a subject stays in [Spec, vP] (or more generally, when XP Merger to [Spec, TP] is excluded).

5. Theoretical Implications

In this paper, we have proposed a derivational analysis of that-t effects (= (5)) and have discussed the validity of the proposal with three empirical cases from various languages. In this section, we will discuss two theoretical implications that follow from our proposal, both of which are concerned with main issues in the MP.

5.1. Head Movement

The first and foremost implication is that head movement is a property of narrow syntax. A crucial element of our proposal is that T moves to C in narrow syntax, so that the inherited EPP feature in T is satisfied through a head-to-head relation with a C head. In other words, T-to-C movement has narrow syntactic effects. In the current literature, a more dominant assumption is that head movement is a phonological process and it has been pushed out of narrow syntax. For instance, in the following well-known verb movement contrast between English and French, what triggers verb movement is not a syntactic mechanism like feature checking (as proposed in Chomsky (1995)) but phonological factors: the phonological component of English does not raise verbs to T while the French counterpart in fact does:

\[(50)\] John often kisses Mary. (English)
\[(51)\] Jean embrasse\(t_i\) souvent \(t_i\) Marie. (French)

Jean kisses often Marie

'Jean often kisses Marie.'

If our discussion is on the right track, however, it suggests that this assumption must be rethought. Head movement is not excluded but is operative in narrow syntax as well.\(^{18}\) As we have discussed, head movement has

\(^{18}\) Note that head movement (or local Merge) is not problematic under a kind of economy condition proposed in Chomsky (2000) (so-called No Tampering Condition), which requires that the relations provided by Merge and composition not be destroyed in later derivations, the relevant relations here being sisterhood and c-command (ibid.: 136–137). Head movement (or head adjunction) does not tamper with the basic relations within a projection.
a function of removing the unsatisfied EPP feature in T and can be considered to be an essential property of narrow syntax. In fact, evidence has been pointed out that head movement is subject to locality constraints in narrow syntax just like XP-movement (Johnson (1991)), and there is much research in the recent syntactic literature demonstrating that head movement plays an important role in narrow syntactic derivations (den Dikken (2006), Gallego (2006), Miyagawa (2001) and Pesetsky and Torrego (2001)). Besides, Ogawa (2007) shows that head movement has non-trivial semantic effects, which also suggests that head movement is operative within narrow syntax.

From these considerations, we can conclude that head movement can take place in narrow syntax (at least in some cases), and at the present stage of research, the jury is still out on the status of head movement as a phonological reflex. To the extent that the discussion in this paper is on the right track, our proposal, together with other works cited above, favors the view that head movement is one legitimate operation in narrow syntax.

5.2. Non-attraction Movement

The second implication is that movement can be self-serving or "greedy." Put differently, non-attraction movement is possible in narrow syntax. The T-to-C movement we have considered takes place only when the inherited EPP feature in T cannot be satisfied by XP Merger or by verb movement. Since the head movement is driven for the unsatisfied EPP feature in T and C is irrelevant for this purpose, it is due to the internal property of T, not of C. This then suggests that T-to-C movement is an instance of greedy movement. In fact, Bošković (2005) persuasively argues that some instances of movement (such as successive cyclic movement) are analyzed more as non-attraction movement than as attraction movement due to some uninterpretable feature inherent in a mover. If our proposal is on the right track, it argues for the view that self-serving, non-attraction movement is implemented in $C_{HL}$ along with attraction movement.19

However, this conclusion raises an important conceptual question on greedy movement in terms of economy considerations (local economy). Chomsky (1995) proposes to reformulate movement as "attraction" in terms of the MLC (Minimal Link Condition). Given the MLC, Greed requires global comparison of derivations in order to determine the closest

19 The idea explored here resonates with Lasnik's (1995a) Enlightened Self Interest, which says that some requirement of a mover or of a target drives movement.
element. On the other hand, Attract allows local determination, making the decision locally, without any global comparison. In addition, Greed requires global comparison to decide whether or not a certain instance of greedy movement will indeed satisfy some property of a mover, which is not the case under attraction, where such a property, if any, is satisfied as a side effect of attraction. In other words, greedy movement allows "look-ahead," which is unfavorable to computational complexity. Then to the extent that local economy is favored over global economy in terms of computational complexity, greedy movement should not be an option of CHL. Furthermore, empirical discussions have suggested that those cases that used to be treated under Greed are equally dealt with locally under Attract (Collins (1997)). Then these conceptual and empirical arguments may argue against Greed, which is unfavorable to our proposal that T-to-C movement is greedy due to the unsatisfied EPP feature in T.

We would like to argue that non-attraction movement is not problematic in light of local economy if local economy in C_{HL} is understood within a phase-based theory of computation. It may be true that local economy is superior to global economy if computational complexity is seriously considered, which is one element of the MP. At the same time, however, look-ahead cannot be excluded completely from C_{HL} to the extent that narrow syntax works in such a way as to achieve convergence (see Chomsky (2000) for the optional EPP assignment in long distance movement) and local economy in C_{HL} should not be the one proposed in Collins (1997) but rather the one that allows minimum look-ahead or local comparison of derivations.20 Given this conception of local economy and phase-based computation proposed in Chomsky (2000), we then suggest that comparison of derivations is possible locally within a phase, which keeps computational complexity to a minimum. Phase-based derivations restrict the domain of computation to the current phase, ignoring previous phases by imposing the PIC (Phase Impenetrability Condition). The PIC is stated as follows:21

(52) The domain of H is not accessible to operations at the next phase; only H and its edge are accessible to such operations.

(Chomsky (2001, 2004))

---

20 For much related discussion in this subsection, see Fujita (1996). We basically follow Fujita (1996) in the conception of local economy.

21 H is a phase head (either C or transitive v). "Edge" is the residue outside of H', either specifiers or elements adjoined to HP.
Then to the extent that look-ahead is locally computed within the current phase being subject to computation and is restricted to this phase, computational complexity caused by look-ahead is substantially reduced and is kept to a minimum.

With phase-based derivations in mind, recall that in our analysis, T moves to the immediately higher head (most typically, C that selects it) within a CP phase. We should note that the decision to move T to C for its unsatisfied EPP feature is locally made in this CP phase, which means that comparison of derivations is made not globally but locally, phase by phase. Furthermore, the target of movement is also determined locally and automatically within this phase, since it is the next higher head (in our instance, C).

To the extent that comparison of derivations takes place at each stage of derivation (i.e. at each phase), which is the case for the proposed T-to-C movement, it is not globally but locally performed, conforming to local economy considerations. Step-by-step comparisons are made possible in CIL so long as local economy in narrow syntax is understood in terms of phases. The problem of global computation is circumvented. Therefore, no problem arises with our proposal in terms of local economy.

In summary, we have considered two theoretical implications for the MP that follow from our proposal: head movement and non-attraction movement. We have also discussed how local economy should be understood if non-attraction movement is operative in syntax.

6. Conclusion

Our aim in this paper has been to give a principled account to that-t (more generally, overt Comp-trace) effects along the research guidelines of the MP. The MP assumes that derivations proceed step by step through local computation of formal (uninterpretable) features. We have proposed a derivational analysis of that-t effects. We have argued that that-t effects are due to the unsatisfied EPP feature in T: the inherited EPP feature in T cannot be satisfied at any point in the derivation and crashes the derivation when it is Transferred. Thus, under our proposal, that-t effects fall out as one consequence of derivational computation. We have demonstrated that a number of data from various languages can be successfully accounted for by our EPP analysis and have shown that parametric variations with that-t effects are accommodated. Finally, we have discussed two theoretical implications for the MP, both of which are independently advocated in recent
Appendix: Recent Approaches to That-t Effects

In this appendix, we will discuss two recent accounts of that-t effects in the MP and show that the proposed analysis is conceptually and empirically superior to these accounts.

Ishii (2004) claims that that-t effects are ruled out by the PIC. Adopting the Vacuous Movement Hypothesis (VMH) (Chomsky (1986)), he says that wh-subjects target [Spec, TP], not [Spec, CP] and move only up to [Spec, TP] in the embedded clause for successive cyclic movement. This has the effect that wh-subjects are inaccessible to the matrix v probe (the P-feature and the EPP feature) in later derivations because they are not at the edge of the embedded CP. This inaccessibility of wh-subjects for the PIC, Ishii argues, leads to that-t effects.

Ishii's analysis, however, faces problems. Empirically, the VMH is quite questionable. There is much evidence suggesting that wh-subjects do move to [Spec, CP]. One piece of such evidence is the distribution of the hell attached to wh-phrases. Pesetsky (1987) notes that wh the hell is allowed only when it has moved to [Spec, CP] overtly. As (53) shows, the hell is possible with wh-subjects, which suggests that they have moved to [Spec, CP]:

(53)  a. Who the hell bought what?
    b. Who wonders who the hell bought this book?
    c. *What did Sue give to whom the hell?

(Pesetsky and Torrego (2001: 405–406, fn. 9))

Conceptually, Ishii's account requires mechanisms unfavorable to the MP. For instance, Ishii introduces complicated definitions (see his (28)–(30)) so that the movement to [Spec, TP] by wh-subjects can simultaneously satisfy the EPP feature in the embedded C, a feature optionally assigned for successive cyclic movement. Since these complications do not follow from the Strong Minimalist Thesis, they are undesirable. Furthermore, since the EPP feature is a selectional feature as persuasively argued in Lasnik (2001), it is unlikely that the EPP feature in C can be satisfied by wh-subjects in [Spec, TP] without movement; the satisfaction of the EPP feature requires structure building (i.e. Merge/Move).

In addition to these problems, Ishii needs a stipulation that the embedded CP is not projected in that-less clauses in (1b); otherwise, wh-subjects in the embedded [Spec, TP] would also be inaccessible to the matrix v. But
as we have already argued in section 3, it has been suggested that the Case/agreement properties of T are a reflex of C-T relations. CP must then be projected even in (1b). Finally, if TP is not propositional, hence not a phase, operative complexity would increase in (1b) because a phase is not present in that-less embedded clauses. The other account is found in Pesetsky and Torrego (2001) (P&T). They offer an economy-based account of that-t effects. They argue that C has an uninterpretable Tense (uT) feature with the EPP property. They further argue that a wh-subject can satisfy uT in C when it moves to the embedded [Spec, CP] for uWh in C, a feature which triggers successive cyclic movement, on their assumption that nominative case is uT in D. Since that in C is a reflex of T-to-C movement and this movement has a function of satisfying uT in C just like subject wh-movement to [Spec, CP] under P&T's analysis, that-t effects are ruled out by economy: there is simply no need to rely on T-to-C movement to satisfy uT in C in subject wh-movement. It seems that there are problems with P&T's proposal, however. The assumption crucial to their analysis is that uT (nominative Case) is also present in C, in addition to T. However, it is unclear how this assumption is motivated conceptually and empirically, except for the purpose of explaining that-t effects and related matters. Furthermore, under their assumption, this uT in C is deleted (checked/valued) by a wh-subject in (1b). But this goes against an activation condition on Agree, which requires both a probe and a goal to have an undeleted uninterpretable feature for the operation to take place. This is because uT of the subject, which is supposed to be a goal of a probe uT in C, has already been deleted through φ-feature Agreement with T. P&T's analysis also requires that uninterpretable features be deleted against each other. However, uninterpretable features are deleted only through their Agreement (Matching) with their interpretable counterparts. Finally, P&T argue that that is not in C but in T, and moves to C for the deletion of uT in C (...) [CP [T that], [TP t ...]]). This immediately raises a question of where that is base-generated in cases where modal auxiliaries are present in T, since they should compete for the same structural position (that is, T).

Now let us turn to an empirical problem. P&T's analysis predicts that that-t effects are universal, for subject wh-movement to intermediate [Spec, CP] always precludes T-to-C movement for the deletion of uT in C due to

22 See also An (2007) for arguments against the VMH.
economy. As we have already discussed, however, there are languages like Italian in which \textit{that}-\textit{t} effects are absent. Given that a subject \textit{wh}-phrase undergoes movement to [Spec, CP] due to the PIC in these languages as well, then it will be expected that \textit{that} (or its counterparts), which is assumed to be a reflex of T-to-C movement, never appears and that \textit{that}-\textit{t} is universally excluded, which is empirically wrong.

In this Appendix, we have considered two recent Minimalist accounts of \textit{that}-\textit{t} effects and have shown that they face non-trivial conceptual and empirical difficulties. On the other hand, as we have shown in sections 3 and 4, our proposal does not require any special assumptions or complications and is empirically desirable: it only follows from independently motivated theoretical assumptions, and can properly explain both grammatical and ungrammatical examples of \textit{that}-\textit{t} effects. This demonstrates that our proposal is conceptually and empirically more favorable than those we have considered in this Appendix.\textsuperscript{23}

\textbf{REFERENCES}


\textsuperscript{23} Roussou (2002) is another recent study which proposes an account of \textit{that}-\textit{t} effects in a slightly different framework. She argues that \textit{that}-\textit{t} effects are due to the fact that an overt complementizer blocks T-to-C movement, which is required for the lexicalization of Agr-feature in T, which she says is a PF requirement. But this assumption is questionable. It is unclear why a phonologically \textit{null} complementizer can lexicalize Agr-feature, if a PF condition requires this feature to be phonologically substantiated. Our T-to-C movement, on the other hand, is required for narrow syntax and this problem does not arise. Moreover, she assumes that a \textit{wh}-phrase does not move but is base-generated in [Spec, CP], and Agrees with Agr-feature, a variable, in the embedded C for identification. This Agreement relation, however, has to cross a number of phase boundaries and is blocked by the PIC, which Roussou has in mind.
DERIVATION, MINIMALISM, AND THAT-TRACE EFFECTS


Department of English
Dokkyo University
1-1 Gakuen-cho, Soka-shi
Saitama 340-0042
e-mail: mizuguchi@dokkyo.ac.jp