RECONSIDERING PHASE-INTERNAL DERIVATIONS: ARE THEY EXCEPTIONAL OR NOT?

MANABU MIZUGUCHI
Dokkyo Medical University

This paper explores phase-internal derivations and considers their consequences for Minimalist theorizing. With the discussion of extraction from subjects, we claim that phase-internal derivations are step-by-step processes, and argue against simultaneous, parallel applications of operations at the phase level. We argue that our claim is theoretically favorable, and show that it can naturally explain extraction phenomena as well as grammatical behaviors of multiple wh-questions. We also take a fresh look at “well-formed” extraction from subjects and subject wh-movement. We propose new analyses and provide empirical evidence in their favor. The discussions in this paper strengthen the argument that step-by-step procedures explain language.*

* This article is a revised and extended version of the paper I presented at the 29th Conference of the English Linguistic Society of Japan (held at Niigata University, November 2011) and which I subsequently published in JELS 29 (Mizuguchi (2012)). I am indebted to the participants in the conference, especially Etsuro Shima, for their helpful comments and questions. I would also like to express my deepest gratitude to three anonymous EL reviewers for their invaluable comments and helpful questions on earlier versions of this paper. As always, all remaining errors and inadequacies are mine. The research reported in this paper was supported in part by Grant-in-Aid for Young Scientists (B) (grant number 24720199) from Japan Society for the Promotion of Science.

1. Introduction

The notion of “cycle” has been considered one of the hallmarks of derivations and has played an important role in the development of syntactic theory. In the Minimalist Program starting with Chomsky (2000), cyclicity of derivations has been captured through “phases” and it has been assumed that derivations are cyclically constructed phase by phase. Phases

Keywords: phases, feature-inheritance, phase-internal derivations, step-by-step procedures, wh-movement
restrict syntactic domains subjected to computation, keeping computational complexity to a minimum. Phase heads are considered to be C and \( \bar{v}^* / v \) (as well as D), and play prominent roles in syntactic computation. In a series of recent papers, Chomsky (2007, 2008), proposing more strengthened phase syntax, argues that phase-internal operations take place simultaneously at each phase level, with no step-by-step applications of operations inside phases.

This paper explores phase-internal derivations under phase syntax currently being investigated. We agree with Chomsky (2000 et seq.) that phase-based computation can make syntactic processes more restrictive, keeping computational complexity to a minimum. We argue, however, that syntax does not implement simultaneous, wholesale applications of operations within phases. Considering subject condition effects under phase syntax with feature-inheritance, we claim, pace Chomsky (2007, 2008), that phase-internal derivations also proceed step by step, eliminating representational properties of syntactic computation from Minimalist syntax.

This paper is organized as follows: in section 2, we first discuss phase syntax as a framework of our discussions and consider two theoretical assumptions central to this syntax. In section 3, we spell out our claim by considering subject condition effects, showing its theoretical and empirical advantages over simultaneous, parallel derivations within phases. In section 4, we consider the validity of our claim with wh-movement phenomena. In section 5, as consequences of our claim, we reconsider licit extraction from subjects and subject wh-movement, proposing analyses which are compatible with our claim. In section 6, we present the conclusion of this paper.

2. Theoretical Background

2.1. Feature-Inheritance from Phase Heads

Phases are basic units of derivations, and derivations are successively built up with phase heads as their computational pillars. Under phase syntax currently being explored, it has been assumed that uninterpretable (or unvalued) features are exclusive properties of phase heads. One major consequence of this assumption is that unlike in previous models, non-phase heads (T, V, N) do not have uninterpretable features of their own (say, Agree (\( \phi \))-features); instead, these features are inherited by non-phase heads in the course of derivations when phase heads are Merged with non-phases. Thus, phase syntax incorporates feature-inheritance as one of its mechanisms. For instance, uninterpretable \( \phi \)-features are properties of C
and they are inherited by T from C when C is Merged with TP, as shown in (1):

(1) **Feature-Inheritance**

\[
[\text{CP } C_{\{\phi\}} [\text{TP } T [vP \ldots]]] \rightarrow [\text{CP } C [\text{TP } T_{\{\phi\}} [vP \ldots]]]
\]

Under this phase model, the syntactic properties of non-phase heads like T are functionally determined by phase heads (say, C); the locus of Case/agreement properties is not T but C (Chomsky (2004), Hiraiwa (2005)).

The mechanism of \(\phi\)-feature inheritance, though much entertained in the current Minimalist literature on phase syntax, is yet to be clarified, and we must consider its exact mechanism for the purpose of our discussions in this paper. To be more specific, the question is whether \(\phi\)-features are inherited on their way to probing or they probe after they are inherited by T. We suggest that feature-inheritance to T takes place when \(\phi\)-features probe for a goal. Provided that probes start searching upon their Merge, \(\phi\)-features, which are inherent in C, should start probing upon the Merge of C. This argument naturally leads us to the assumption that \(\phi\)-features in C are inherited by T on their way to probing for their goal: \(\phi\)-features probe from C to be inherited on their way. We thus propose that \(\phi\)-features are inherited via probing:

(2)

\[
\text{Merge of } C_{\{\phi\}} \quad \text{Inheritance via probing} \quad \text{Agree}(T_{\{\phi\}}, \text{DP})
\]

Phase syntax with feature-inheritance gives us a new picture on the derivation of subject raising. Provided that \(\phi\)-features are properties of C and probe after C is Merged with TP, unlike in previous frameworks, where subject raising is followed by the Merge of C, the subject raises to Spec-T

---

1 The question here is why \(\phi\)-features are inherited from phase heads. Some proposals have been made in the literature: for A/A’ distinction at the C-I interface (Chomsky (2008)), for Value-Transfer simultaneity (Richards (2007)), for Case-valuation (Epstein, Kitahara and Seely (2012)). We refer the reader to the works cited above for details. Since the discussion is not central to our argument and any of these proposals suffices for our purpose, we remain neutral on the reason behind feature-inheritance.
inside CP phase, after the phase head is Merged in. Consider the subject raising of John in (3) under feature-inheritance:\(^2\)

(3) John read Barriers.

(4) a. \([CP \ C \ [TP_{\{\phi\}} \ [v^*P \ John \ v^* \ [VP \ read \ Barriers]]]]\)
b. \([CP \ C \ [TP \ John \ T_{\{\phi\}} \ [v^*P \ t_j \ v^* \ [VP \ read \ Barriers]]]]\)

In (4a), \(\phi\)-features in \(C\) are inherited by \(T\) at the \(C\) phase level, and after the inheritance, John Agrees with \(T\) in \(\phi\)-features and raises to Spec-T internally within this phase (= (4b)).\(^3\)

2.2. Operations at the Phase Level

Phase syntax with feature-inheritance implies that \(\phi\)-features in \(C\) probe together with its EF, which raises an operator element such as a wh-phrase to Spec-C. Recall that in phase syntax, \(\phi\)-features are properties of \(C\) and that they probe from \(C\) upon its Merge. This means that when both \(\phi\)-features and EF are present in \(C\), these probes start probing (searching) together when \(C\) is Merged with TP; in other words, A-syntax and A’-syntax both apply at the phase level. To make our discussion more specific, consider the derivation of object wh-movement in (5), which is sketched in (6):\(^4\)

(5) What did John read?

(6) a. \([TP \ T \ [v^*P \ what_i \ [John \ v^* \ [VP \ read \ t_i]]]]\)
b. \([CP \ C \{EF\} \ [TP \ T_{\{\phi\}} \ [v^*P \ what_i \ [John \ v^* \ [VP \ read \ t_i]]]]\)
c. \([CP \ what_i \ C \{EF\} \ [TP \ John_j \ T_{\{\phi\}} \ [v^*P \ t_j \ v^* \ [VP \ read \ t_i]]]]\]

\(^2\) In this paper, since we focus on CP phase, unless otherwise mentioned, the discussion of \(v^*P\) phase, including feature-inheritance from \(v^*\) to \(V\), is omitted for simplicity. Irrelevant details of derivations are also omitted in (4) and elsewhere.

\(^3\) Chomsky (2008: 149) says that \(\phi\)-features raise the subject to Spec-T. Given that Merge (Internal as well as External) is triggered by the edge-feature (EF), this suggests that \(\phi\)-features are equipped with EF. We assume, following Epstein, Kitahara and Seely (2012), that EF is a higher-order property of features and that \(\phi\)-features can have EF. Thanks to the inheritance of \(\phi\)-features, \(T\) bears EF for second-Merge. The so-called EPP (or EPP property) of \(T\) is thus reformulated as EF of \(\phi\)-features inherited by \(T\).

A-movement in the feature-inheritance model is counter-cyclic in that it takes place inside CP phase, and it apparently violates the No-Tampering (Extension) Condition. One solution to this problem is proposed by Epstein, Kitahara and Seely (2012), and Narita (2011). In this paper, we just note that there is a way of solving counter-cyclic A-movement under feature-inheritance syntax, and that it does not pose a problem to Minimalist syntax.

\(^4\) In (6), what moves to the edge of \(v^*P\) in order to move out of the phase (the Phase Impenetrability Condition (henceforth, PIC)).
In (6), Agree in $\varphi$-features does not take place when $T$ is Merged with $v^*P$ (= (6a)). As soon as $C$ is Merged in, its EF and $\varphi$-features start probing together at this phase level, with the latter being inherited by $T$ on their way to probing (= (6b)). As a result of probing, the EF raises what to Spec-C and $\varphi$-features, thanks to their own EF (see fn. 3), move John to Spec-T in the CP phase (= (6c)). This simultaneous probing has the effect that unlike in previous approaches, where wh-movement and subject raising take place at different stages of the derivation (CP and TP, respectively), the C phase level is the only point of the derivation where the two operations are applied: A-syntax does not come before A'-syntax; in other words, there is no A-movement to Spec-T before A'-movement to Spec-C. As a natural consequence of feature-inheritance, A-syntax and A'-syntax are applied together at the same derivational stage (i.e. at the phase level) under phase syntax, not at different stages of the derivation.

In this section, we have discussed two key components of phase syntax and laid out the theoretical background for our discussions in the following sections.

3. Phase-Internal Derivations

In the last section, we have reviewed phase syntax and considered two hallmarks which make phase syntax distinct from previous models of generative syntax. With this theoretical background in mind, in this section we scrutinize phase-internal derivations and claim that they proceed in a step-by-step manner.

3.1. All at Once or Step by Step?

As we have discussed, phases are loci of computation and all operations are driven by probes (uninterpretable/unvalued features) inherent in phase heads at the phase level. Chomsky (2007, 2008) develops this approach to the extent that all operations take place simultaneously at the phase level: if phase heads have all the probes that trigger operations inside phases and probing starts only at the phase level, then all the operations within phases should be applied simultaneously, in a parallel manner, upon the Merge of phase heads. Consider once again (5) under simultaneous derivations. When CP phase is constructed, both a wh-phrase and the subject, which are probed by EF and $\varphi$-features in C respectively, undergo parallel movement at this phase level; as sketched in (7), A-movement and A'-movement apply in simultaneity with each other at the phase level upon the
Merge of C:

(7) \[\text{CP what}_i \text{C}_{\{\text{EF}\}} \{\text{TP John}_j \text{T}_{\{\phi\}_j} \{\text{v}^* \text{P} \text{t}_i \text{t}_j \{\text{v}^* [\text{VP read t}_i]\}]\}]\]

Simultaneous derivations inside phases certainly constitute one reasonable hypothesis against the background of phase syntax discussed in the last section. We argue, however, that such derivations must be considered with caution. For one thing, simultaneous derivations will reintroduce a GB-type representational analysis in the form of phases and diminish the role of derivation, which is undesirable in light of the Minimalist Program. Under simultaneous derivations, all the operations are packaged in one at the phase level, and so-called D-structure, S-structure or Logical Form (LF) in the Government and Binding framework are replaced with phases (CP, v^*P) as the level of operation applications. Such wholesale applications of operations are representational in nature, and return to a GB model of syntactic derivations, where the principles of Universal Grammar apply to linguistic levels at the same time (Epstein and Seely (2002), Grewendorf and Kremers (2009)). Provided that the Minimalist Program is an attempt to derive linguistic properties by seeking an explicitly derivational theory and deducing representational properties from derivation (e.g. Epstein (1999)), it is doubtful whether simultaneous derivations at the phase level are on the right track.\(^5\)

In addition to the representational nature mentioned above, simultaneous applications of operations inside phases are problematic from third-factor considerations. One of the distinguishing characteristics of Minimalist syntax is its assumption that computational complexity matters for language. The notion of computational complexity still being unclear at the present

\(^5\) A reviewer has correctly pointed out that the Minimalist Program may not necessarily be incompatible with the (GB-type) representational approach, citing Rizzi (2004, 2006) as one such example. It might be true that the representational approach can be compatible with the Minimalism. However, it is redundant with the derivational approach in that it must presuppose derivations in order to construct syntactic structures for representations, and representational properties are deducible from derivations; indeed, Rizzi’s (2004, 2006) approach can be derivationally formulated (e.g. Blümel (2012)). Moreover, the representational approach shows a type of delay in that it waits until all the syntactic structures are constructed for the application and evaluation of syntactic operations. We thus assume that the Minimalist Program, where Merge is the only generative engine, is an explicitly derivational approach. I thank the reviewer for bringing representational Minimalism to my attention.
stage of Minimalist research, we can reasonably think that simultaneous applications of operations increase computational complexity in that syntax has to process all the computational operations in a phase at once. Suppose that two operations O1, O2 need to take place: in one derivation, O1 and O2 are processed and completed as soon as O1 and O2 can apply; in the other, O1 and O2 take place and are completed simultaneously. In the former, the number of computations to be processed at a time is just one (O1 or O2), and computational workload can be kept to a minimum. On the other hand, in the latter, the number is two (O1 and O2), and syntax has to process two operations in one go, which will increase computational workload. Since more workload endows syntax with more burdens of computation, computational complexity will increase. Besides, simultaneous derivations imply a kind of delay: instead of applying an operation (O1) when it is applicable, syntax has to wait until the other operation (O2) can apply, which is computationally inefficient. It can be said that the smaller the number of operations subjected to computation at a time, the more efficient it will be, with computational complexity kept to a minimum. This is also suggested by the binary application of Merge, which is considered to be the simplest (hence, the most efficient). Efficient (or simple) computation requires the number of x (x = syntactic objects, syntactic operations) subjected to computation at a time to be as small as possible. Thus, to the extent that computational complexity matters, simultaneous derivations at the phase level should be reconsidered.

We agree with Chomsky (2000 et seq.) that phases play important roles in syntactic derivations; by localizing computation to phases (or phase heads), syntax can be made more restrictive and more efficient. We basically follow this derivation-by-phase model. We argue, however, that syntax is derivational from beginning to end. Provided that sentences are constructed derivationally step by step through phase heads, it naturally follows from this assumption that phase-internal operations should also be applied step by step. Phases are not representational levels where wholesale applications of operations take place; rather, they are simply points of derivations where probes in phase heads trigger syntactic computation one by one, which is completed as soon as the probes find their goals. In the rest of this section, we explore this theoretical hypothesis on phase-internal derivations and claim, with the discussion of subject condition effects, that phase-internal syntactic processes are also step-by-step.
3.2. Phase-Internal Derivations and Subject Condition Effects

It has been noted since early days of generative research that the subject (more precisely, an element moved to Spec-T) is an island for extraction (Chomsky (1973), Huang (1982), Lasnik and Saito (1992) among others). To see some examples, consider the following:

(8)  
   a. *Who_1 did [a picture of _i] cause a scandal? 
   b. *Who_1 was [a picture of _i] taken _j by Bill? 

Wh-movement out of the subject [a picture of who] causes ungrammaticality, and the extraction is ungrammatical regardless of whether the subject is derived (as in a passive sentence (8b)) or not (as in a non-passive sentence (8a)). Various proposals (both Minimalist and pre-Minimalist) have been proposed to explain subject condition effects. To the extent that the Minimalist Program is on the right track, a principled explanation will be achieved for subject condition effects if they follow from the process of syntactic derivations (that is, Merge and Agree). It is demonstrated here that the ungrammaticality of extraction out of subjects is deducible from step-by-step derivations within phases.

Take (8a) as our example and consider its derivation in detail, which is sketched in (9):\(^6\)

\[
\begin{align*}
(9) & \quad \text{a. } [\text{CP } C_{\{EF, \varphi\}} \text{ [TP } \{\varphi\} \text{ [v*P [DP a picture of who] v* [VP cause a scandal]]}]] \\
& \quad \text{b. } [\text{CP } C_{\{EF\}} \text{ [TP } T_{\{\varphi\}} \{\text{v*P [DP a picture of who] v* [VP cause a scandal]}\}]] \\
& \quad \text{c. } [\text{CP } C_{\{EF\}} \text{ [TP } T_{\{\varphi\}} \{\text{v*P [DP a picture of who] v* [VP cause a scandal]}\}]] \\
& \quad \text{d. } [\text{CP } C_{\{EF\}} \text{ [TP } \{\text{v*P [DP a picture of who]} \text{ T}_{\{\varphi\}} \{\text{v*P } t_j \text{ v* [VP cause a scandal]}\}]] \\
\end{align*}
\]

The derivational stage relevant to our discussion is the C phase level, where C, a phase head with EF and \(\varphi\)-features, is Merged with TP (= (9a)). Upon the Merge of C, its EF and \(\varphi\)-features start probing together (= (9b)). In searching for their goals, the \(\varphi\)-probe finds the subject DP earlier than

\(^6\) In this paper, solid lines are used to represent probe-goal (Agree) relations, and dotted lines are employed to represent search (probing).
RECONSIDERING PHASE-INTERNAL DERIVATIONS

the EF-probe finds the *wh*-phrase in the subject because of its embedding structure: the subject $[\text{DP a picture of who}]$ embeds the *wh*-phrase *who* ($=\text{(9c)}$). As a result, under step-by-step derivations, Agree in $\varphi$-features derivationally comes first and the $\varphi$-probe raises the subject to Spec-T upon the Agree, since the $\varphi$-features are inherited by T on their way to probing ($=\text{(9d)}$). This has the effect, however, that the EF-probe cannot find the *wh*-phrase embedded in the subject; subject raising, which takes place earlier in the derivation due to Agree in $\varphi$-features, dislocates *who* to the position outside of the search domain of the EF-probe before the EF can reach down into the subject DP to find its goal. Under step-by-step derivations, the subject Agrees in $\varphi$-features and is moved to Spec-T in the very course of the EF probing for *who*. The search by the EF fails, and *who* cannot move out of the subject ($=\text{(9d)}$). The derivation crashes, which explains subject condition effects.

It should be noted that after the subject has moved to Spec-T, the EF-probe cannot search and find the embedded *wh*-phrase. Recall that under phase syntax, because of simultaneous probing by EF and $\varphi$-features at the phase level upon the Merge of C, EF has already begun its probing (search) along with the $\varphi$-probe by the time the subject moves to Spec-T. Spec-T, which is yet to be created at the time of the Merge of C, does not fall in the search domain of EF. When Spec-T is counter-cyclically created, on the other hand, C has already been Merged with TP and it cannot have any derivational (i.e. search/probe-goal) relation with either Spec-T or its terms (Epstein (1999), Epstein, Kitahara and Seely (2012)). Thus, its EF cannot probe into this position and extraction out of Spec-T is derivationally impossible.\(^7\) Furthermore, under derivational syntax, EF cannot wait to probe until Spec-T is created by subject raising. Unlike in previous frameworks, where C and T are engaged in computation at different stages of the derivation and an independent condition is required to explain the ungrammaticality of (8), the invisibility of Spec-T is derived as one consequence of derivations under our proposal, and phase syntax can give a principled explanation

\(^7\) We should note that under the representational approach, where syntactic relations are arbitrarily defined based on output representations, Spec-T and its terms can fall in the search domain of C and serve as potential goals for its EF even after the subject moves counter-cyclically to Spec-T, which is because C can c-command Spec-T. To the extent that our proposal is correct, the derivational approach is supported.
to subject condition effects.  

Now let us consider (8a) under simultaneous derivations. We should note that simultaneous derivations wrongly predict that extraction from subjects is grammatical. Recall that all the syntactic operations inside a phase are processed at once. This means that one operation is not applied earlier or later than another in a phase; put differently, phases have the effect that all the probe-goal relations within a phase are equidistant. Thus, in the derivation of (8a), illustrated in (10) below, EF and φ-features simultaneously find their goals in the CP phase, which makes subject raising and wh-movement take place in tandem with each other:

(10)  

a. \[CP \ C_{\{EF, \phi\}} \ [TP \ T \ [v^*_P \ [DP \ a \ picture \ of \ who] \ v^* \ [v_P \ cause \ a \ scandal]]]]\]  
b. \[CP \ [C_{\{EF\}} \ [TP \ T_{\{\phi\}} \ [v^*_P \ [DP \ a \ picture \ of \ who] \ v^* \ [v_P \ cause \ a \ scandal]]]]\]  
c. \[CP \ who_i \ C_{\{EF\}} \ [TP \ [DP \ a \ picture \ of \ t_j] \ T_{\{\phi\}} \ [v^*_P \ t_j \ v^* \ [v_P \ cause \ a \ scandal]]]\]  

But this derivation, unless some condition or constraint is assumed, would allow who to be extracted from the subject. We can thus conclude that “all-at-once” derivations pose empirical as well as theoretical problems to syntactic theory. On the other hand, as we have discussed above, under the step-by-step derivations we have proposed, subject condition effects follow straightforwardly from the embedding structure of the subject.

3.3. Predicate Fronting

We have argued that extraction from subjects can be derivationally explained under the phase-based derivations we have proposed. The proposed

---

8 A reviewer has noted the possibility of explaining subject condition effects with what Rizzi (2006) calls Criterial Freezing. Criterial Freezing, however, cannot account for subject condition effects for two reasons. First, Criterial Freezing prohibits the movement of an element which meets a criterion (in our case, the subject which satisfies the Subject Criterion) but it does not prohibit extraction from an element satisfying a criterion, which is evidenced by examples like (26)–(28) below. Thus, extraction from the subject is possible even if it satisfies the Subject Criterion. Second, under feature-inheritance, extraction from the subject will take place before the subject moves to satisfy the Subject Criterion (see (10)), and Criterial Freezing is simply orthogonal to subject condition effects. The proposed derivational approach, on the other hand, can correctly explain the ban on extraction from subjects.
analysis, however, might face a problem if predicate fronting, illustrated in (11), is taken into account:

(11) (Joe though Mary would criticize Bill, and) Criticize Bill, Mary certainly did.

The predicate fronting of this kind (so-called VP-fronting) is one instance of topicalization and can be considered a movement triggered by EF in C. The problem lies in the fact that the subject Mary is embedded in v*P. With this in mind, consider the derivation of (11):

(12) a. \[ CP \{EF\} \{TP \{φ\} \{v*P Mary v*-criticizek [vP t_k Bill]\}\} \]

b. \[ CP \{EF\} \{TP \{φ\} \{v*P Mary v*-criticizek [vP t_k Bill]\}\} \]

c. *=CP \{v*P Mary v*-criticizek [vP t_k Bill]\}, C\{EF\} \{TP \{φ\} t_i\} \]

In the CP phase, the φ-features of C, along with its EF, start probing upon their Merge, and are inherited by T on their way to probing (= (12a)). Since Mary is embedded in v*P, EF would first target v*P, and raise it to Spec-C under the proposed analysis (= (12b)). This has the effect, however, that Mary would be out of the search domain of φ-features (= (12c)); the derivation would crash due to unvalued φ-features in T and (11) would be impossible.

We argue that predicate fronting does not pose a problem to our analysis. The crash of the derivation comes from the fact that v* projects up to “v*P.” Projections or labels are being reconsidered in the current Minimalist syntax because they do not follow from Merge, and require a stipulated mechanism such as feature-percolation, which raises a number of questions on its implementation in light of third-factor principles. With this in mind, suppose that projection labels are only a convenient notational device, eliminated from Minimalist syntax (see Chomsky (2007), Collins (2002), Narita (2011), Seely (2006) for much related discussion). Given this label-free framework, the structure of (11) is represented as (13):

(13) \[ \{C\{EF\} \{TP \{φ\} [Mary v*-criticizek [t_k Bill]]\}\} \]

In (13), Mary c-commands, and is hence structurally higher than, v*-criticizek. In searching for its goal, the φ-probe finds its goal Mary first and raises it to Spec-T, because there is no v*P for EF to target. On the other hand, the EF-probe continues searching. Since v*-criticizek is closer to EF than V, which is also rendered invisible due to cyclic Transfer (the PIC), it is targeted by the probe and v*-criticizek (or v*-headed set {v*,{V, …}}) is raised (topicalized) to Spec-C for the EF. Thus, the derived structure is not
(12c) but (14), which comes out as (11):  

\[
(14) \quad [[v^*-\text{criticize}_k [t_k \text{ Bill}],]_i C_{[\text{EF}]_j [\text{Mary}_j T_{[\phi]} [f_j t_i]]}]
\]

With an independently motivated, label-free assumption taken into account, the proposed analysis does not pose a problem to predicate or VP-fronting.  

It should be noted that the analysis of subject condition effects we have proposed in section 3.2 still goes through under the label-free assumption. Recall the derivation (9). The subject of (8a), [a picture of who], is analyzed as (15) under this assumption. In (15), the determiner a, which is the goal of φ-features, c-commands who, the goal of EF. This means that in the search by the two relevant probes at the C phase level, the φ-probe finds and raises a (or a set headed by the determiner) before the EF-probe.

9 A reviewer has asked why \([v^*-\text{criticize}_k [t_k \text{ Bill}]],\) not \(v^*-\text{criticize}_k\) alone, is fronted. We suggest that \([f_k \text{ Bill}]\) moves with \(v^*-\text{criticize}_k\) because verb-topicalization, as found in some Scandinavian and Germanic languages, is disallowed in English. Consider the following Swedish example from Holmberg (1999: 7):  

(i) Kysst har jag henne inte (bara hållit henne i handen).  

In (i), the verb kysst comes to the left of har, which is in C. Since it moves across har, it is not head-moved to C but topicalized to Spec-C.  

The next question is why verb-topicalization is possible in some languages (Scandinavian and Germanic languages) but not in others (English). We do not have enough space to explore this question and have to leave for future research the discussion of the cross-linguistic differences of verb-topicalization.  

10 It should be noted that the Proper Binding Condition, which is representational in nature, derivationally follows from the proposed analysis.  

11 Some may wonder how the derivation (14) can deal with binding examples of the kind discussed in Huang (1993):  

(i) Criticize himself, (Mary thinks) John will not.  

In (i), the anaphor himself in the fronted v*P is bound by the subject John in Spec-T. On the assumption that anaphors must be locally bound, (i) may suggest that there is a copy of John in the fronted v*P; that is, \([f_k v^*-\text{criticize}_k [t_k \text{ himself}]]\) is preposed (see Huang (1993) for discussion). In the label-free framework, v*P, which includes a copy of the subject and forms a unit of movement, is non-existent and \([t_j v^*-\text{criticize}_k [t_k \text{ himself}]]\) is immovable. Chomsky (2007, 2008) argues, referring to Reuland’s (2001) discovery, that binding relations are reducible to probe-goal relations. Given this analysis, the binding relation in (i) can be established derivationally through a probe-goal relation: V and himself Agree, and T and John also Agree; V, which raises to v*, can be in the search domain of this T at the v*-edge. Provided that probe-goal relations are derivational, John indirectly binds himself in the course of the derivation thanks to the probe-goal relation between T, which Agrees with John, and V (or v*-V), which Agrees with himself, before predicate fronting by EF. Thus, an explanation is available in Minimalist syntax which does not require a copy of the subject in a fronted predicate. I thank a reviewer for clarifying this footnote.
finds who:

(15)

Thus, who gets out of the search domain of EF before the probe can reach the wh-phrase. Subject condition effects still receive a derivational account under the label-free framework.

3.4. Extraction from Objects

Before we leave this section, we should consider how extraction from objects is analyzed under the proposed system of derivations. It has been known that extraction from objects, unlike extraction from subjects, is grammatical. As shown in (16), a wh-phrase embedded in the object can be moved out of it without any problem:

(16) a. Who did John see [pictures of ti]?
    b. Who did Kathy write [an article about ti]?

Take (16a) as our example and consider its derivation, which is sketched in (17):

(17) v*P Phase

a. \[v_p \text{John } v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of who}]]\]
b. \[v_p \text{ who}_i [\text{John } v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of } t_i]]]\]
CP Phase
c. \[c_p C_{\{EF\}} [\text{TP } T_{\{\phi\}} [v_p \text{ who}_i [\text{John } v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of } t_i]]]]]\]
d. \[c_p C_{\{EF\}} [\text{TP } T_{\{\phi\}} [v_p \text{ who}_i [\text{John } v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of } t_i]]]]\]
e. \[c_p \text{ who}_i C_{\{EF\}} [\text{TP } T_{\{\phi\}} [v_p t_i [\text{John } v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of } t_i]]]]]\]f. \[c_p \text{ who}_i C_{\{EF\}} [\text{TP John}_j T_{\{\phi\}} [v_p t_i [t_j v^*-\text{see}_k [\text{VP } t_k [\text{DP pictures of } t_i]]]]]\n
The derivation (17) includes two phases. In the v*P phase, the object Agrees with V in φ-features, which are inherited from v*. Suppose that unlike the subject, the object does not move to Spec-V upon its Agree
in φ-features. The wh-phrase embedded in it is Internally Merged to Spec-v* due to the PIC, which we assume is triggered by EF in v*. This produces (17b). When the next phase head C is Merged in, its EF and φ-features probe together, with the latter being inherited by T on their way (= (17c)). In the derivation of this CP phase, the EF-probe finds who first and moves it to Spec-C (= (17d)); then the φ-probe finds John, and moves it to Spec-T upon the Agree (= (17e)). Unlike in the case of extraction from subjects, both who and John are in the search domains of EF and φ-features in this CP phase when they probe. The absence of object condition effects follows.13

It should be noted that the derivation shown in (17) can also give a derivational solution to the blocking problem caused by a partially moved wh-phrase, which is a problem in Chomsky (2000, 2001). Suppose that T has φ-features of its own and that they probe when T is Merged in. Provided that who moves to Spec-v* on its way to Spec-C for the PIC, it would prevent T from searching for the subject because of its φ-features, and the Agree relation between T and the subject would be blocked (Defective Intervention Effects):

\[
(18) \quad \left[ TP T_{[\phi]} \right] \left[ v^*P \, \text{who}_{i} \, [John \, v^* \, [vP \, \text{see [DP pictures of } t_i]]] \right] \times \text{Agree}(T_{[\phi]}, John)
\]

12 It is assumed that EF (the EPP property) of V from inherited φ-features (see fn. 3) can be satisfied on a free ride via (obligatory) V-to-v* movement/affixation in syntax (Chomsky (2004)). Hence, the object, unlike the subject, can stay in its first-Merged position (but see fn. 13 for the movement of certain objects to Spec-V for the EF). Related discussion will be provided in section 5.2.

A reviewer has asked how this assumption explains the generalization that verb-movement is a prerequisite for Object Shift if it can satisfy EF in V. We suggest that Object Shift is movement to Spec-v* for v*'s own EF, independent of EF in V from inherited φ-features. Thus even if V-to-v* movement satisfies V's EF, Object Shift can be implemented. The verb-movement requirement on Object Shift can be explained if we take Fox and Pesetsky's (2005) cyclic linearization into consideration.

13 Following Mahajan (1992), we assume that illicit extraction from the object (say, a specific object and an object of certain predicates) comes from its movement to Spec-V for specificity:

(i)  a. *Who did you see the picture of?
   b. ??Who did you destroy a picture of?

Given this assumption and v*-to-V feature-inheritance, the illicit extraction in (i) can be explained on par with subject condition effects in the domain of CP phase: in v*P phase, object movement to Spec-V derivationally comes first upon Agree and dislocates an embedded wh-phrase out of the search domain of EF in v*. See Chomsky (1977) and Endo (2007) for related discussion. I thank a reviewer for directing me to the discussion of (i).
One solution to this problem, as Chomsky (2007, 2008) suggests, is simultaneous, parallel derivations at the phase level. As we have argued, however, the step-by-step derivations we have proposed can also solve this problem: a wh-phrase at the edge of v*P vacates the edge and moves to Spec-C derivationally before T{φ} Agrees with the subject. Thus, the blocking problem does not arise and it cannot be evidence for simultaneous derivations within a phase.

To summarize the discussion, in this section, we have explored phase-internal derivations. After discussing theoretical problems of simultaneous, parallel derivations, we have claimed that phase-internal derivations are also applied in a step-by-step manner, giving the analysis of extraction from subjects as empirical evidence in favor of our claim. As we have shown, provided that a wh-phrase is embedded in the subject, the ban on extraction from subjects follows as one natural consequence of phase syntax if syntactic operations are processed one by one inside a phase.

4. Step-by-Step Procedures: Predictions

In the last section, we have claimed that phase-internal derivations are processed step by step, arguing that subject condition effects are a straightforward consequence of this claim. In this section, we discuss two empirical predictions from the proposed analysis of subject condition effects, and demonstrate that our claim is empirically endorsed.

4.1. Extraction from XP

Recall that subject condition effects follow from step-by-step phase-internal derivations because the subject, which embeds a wh-phrase, moves to Spec-T upon its Agree in φ-features and subject raising dislocates the embedded wh-phrase out of the search domain of EF derivationally before the EF-probe can find the wh-phrase. Given that this step-by-step analysis is correct, a wh-phrase will be able to move out of the subject if the subject does not move upon its Agree in φ-features; as sketched in (19), a wh-phrase embedded in the subject stays in the search domain of EF thanks to the absence of subject raising:

\[(19) \quad a. \quad [CP \ C_{EF} \ [TP \ T_{φ} \ [v*P \ [\text{Subject} \ldots \ wh \ldots] \ v^* \ [\text{VP} \ldots]]]]
\]

The derivations illustrated in (19) are in fact borne out. It has been point-
ed out that extraction from subjects varies in grammaticality from language to language, and that in those languages which allow the extraction, subject raising is not obligatory. To see this, consider, for instance, the following examples from German (20) and Spanish (21):

(20)  \[ CP \] Was i haben \[ TP \] denn \[ v^*P \] \[ DP \] t i für Ameisen \[ a \\
postman \[ bitten \]
\‘What kind of ants have bitten a postman?’ (Diesing (1992: 32))

(21) (?) \[ CP \] [De qué conferenciantes] \[ te \] parece \[ CP \] que me van a \\
of what speakers \[ to-you \] seems \[ that me \] will to \\
impresionar las propuestas \[ t i ]\]
\‘Of which speakers does it seem to you that will impress me the proposals?’ (Uriagereka (1988: 118))

In (20), the absence of subject raising is suggested by the fact that the subject \[ DP \] t i \[ für Ameisen] comes to the right of the adverb denn ‘indeed,’ which demarcates the boundary between TP and \( v^*P \); in (21), it has been reported that post-verbal subjects in Spanish stay in Spec-\( v^* \) and do not move at all (Gallego and Uriagereka (2006)). All these examples show that extraction from the subject is possible when it stays in its first-Merged position, and our step-by-step analysis can explain the grammatical extraction.

On the other hand, once the subjects in (20) and (21) move to Spec-T upon their Agree in \( \phi \)-features, the ungrammaticality results just as in English (8):

(22)  *\[ CP \] Was i haben \[ TP \] [DP \] t i für Ameisen \[ a \\
postman \[ bitten \]
\( \)

(23)  *\[ CP \] [De qué conferenciantes] \[ te \] parece \[ CP \] que las propuestas \[ t i \\
me van a impresionar\]

The ungrammaticality of (22) and (23) is explained on par with that of (8).

The above discussion is also confirmed by the following contrast from English:

(24)  a. Which candidate \[ were \] [posters of \[ t i ]\] all over the town?
     b. *Which candidate \[ were \] [posters of \[ t i ]\] all over the town?
     (Lasnik and Park (2003: 651))

In (24a), the Merge of the expletive there blocks the raising of the subject, in which case, unlike in (24b), the \( wh \)-phrase can be extracted from the subject. The examples we have considered so far corroborate the schematic derivations in (19) and hence, our proposed analysis.
Step-by-step derivations further predict that a moved constituent is not always an island for extraction: an element $\beta$ can be extracted from a moved element $\alpha$ if the probe for $\alpha$ and the probe for $\beta$ are in different phase heads. Consider (25), which illustrates this prediction:

(25) a. $[\text{YP} \alpha \ldots \beta \ldots] \text{Y}_{\{\alpha\}}[\ldots t_j \ldots]$  
b. $[\text{XP} X_{\{\beta\}} \ldots [\text{YP} \alpha \ldots \beta \ldots] \text{Y}_{\{\alpha\}}[\ldots t_j \ldots]]$  
c. $[\text{XP} \text{X}_{\{\beta\}} \ldots [\text{YP} \alpha \ldots t_i \ldots] \text{Y}_{\{\alpha\}}[\ldots t_j \ldots]]$

(25) contains two phases: XP and YP. In the lower YP phase, $\alpha$ moves to its edge because of an $\alpha$-probe in its head, Y (= (25a)). When X, which is a higher phase head bearing a $\beta$-probe, is Merged in and probes for $\beta$, $\beta$ will be in the search domain of X and can be a goal for this probe (= (25b)), which raises $\beta$ to the edge of XP (= (25c)). Suppose that $\text{Y}_{\{\alpha\}}$ is the embedded $C_{\{\text{EF}\}}$ and $X_{\{\beta\}}$ is the matrix $v^*_{\{\text{EF}\}}$, which triggers intermediate movement; $\alpha$ and $\beta$ are $wh$-phrases. Then the $wh$-phrase $\beta$ can move out of the other $wh$-phrase $\alpha$, and hence can move into a higher phase from the edge of the embedded CP thanks to the movement of $\alpha$.

This prediction is, in fact, borne out by examples from various languages. Consider the following data:\(^{14}\)

---

\(^{14}\) Some comments should be in order with respect to (26). It has been noted that the English data is less well-formed than its Italian and Spanish counterparts; in addition, the judgment of the data is shaky among speakers, with variations in its well-formedness. For instance, Saito (1992), and Lasnik and Saito (1992) give “?” to (26) while Fiengo et al. (1988) and Fukui (1999) assign “?” to the data. In this paper, we assume that albeit somewhat marginal, the English sub-extraction is basically grammatical and that a $wh$-phrase can be extracted from the $wh$-phrase moved to Spec-C.

A reviewer has asked how the marginality of English sub-extraction, in contrast with the well-formedness in Italian and Spanish, can be captured under the proposed analysis. I have no illuminating explanation of this but suggest that the variation may come from the difference in the left periphery: in Italian and Spanish, with a fine structure of the left periphery, a $wh$-phrase can move either to Spec-Focus or to Spec-Force in the left periphery while in English, with a fused left periphery, Spec-Force (or Spec-C) is the only target of $wh$-movement. Given this assumption, the embedding $wh$-phrase and the embedded $wh$-phrase in (27) and (28) can be moved by EF in different heads (Focus and Force) and one movement does not interfere with the other, which explains the well-formedness of $wh$-extraction. In English, on the other hand, the two instances of $wh$-movement in (26) are driven by EF in Force/C. Consequently, when $who$ is extracted out of $[\text{which pictures of } who]$ in embedded Spec-C, the movement is interfered with by $which$, which is responsible for the marginality. See Maeda (2010) for the discussion of sub-extraction from a different perspective.
(26) Who do you wonder \([\text{CP} \ [\text{which pictures of who}] \text{ John likes } t_j]?)?
(27) ?[Di quale autore], ti domandi \([\text{CP} \ [\text{quanti libri } t_i]siano by which author you wonder how many books have stati censurati } t_j]?

‘By which author do you wonder how many books have been censored?’ (Rizzi (2006: 114))

(28) [De que autora], no sabes \([\text{CP} \ [\text{qué traducciones } t_i] han by what author not you-know what translations have ganado premios internacionales]?

‘By what author don’t you know what translations have won international awards?’ (Chomsky (1986: 26))

The above examples show that the predicted derivations in (25) are correct. Take, for example, the English data. In the embedded clause, the wh-phrase \([\text{which pictures of who}]\) undergoes wh-movement to Spec-C from its first-Merged position by way of Spec-\(v^*\) (= (29a)). When \(v^*\), which bears EF, is Merged in the matrix clause, the embedded wh-phrase who is in the search domain of this \(v^*\) and a probe-goal relation can be established between \(v^*\) and who (= (29b)). Thus, EF in the matrix \(v^*\) raises who to its edge (= (29c)). From the edge, who moves on to Spec-C in the matrix clause for EF in C when C is Merged in, as shown in (17):

\[
\begin{align*}
\text{(29)} & \quad \text{a. } [\text{CP} \ [\text{which pictures of who}] \text{ John T } [v^* t_j v^* [\text{likes } t_i]]])] \\
& \quad \text{b. } [v^* \text{you } v^*_{\text{EF}} \ [\text{wonder } [\text{CP} \ [\text{which pictures of who}] \text{ John T } [v^* t_j v^* [\text{likes } t_i]]])]]
\end{align*}
\]

In this section, we have discussed extraction from XP and shown that the step-by-step analysis we have proposed can explain grammatical as well as ungrammatical extraction from XP. The data we have considered suggest that specifiers are not always islands for extraction. The proposed analysis can explain this empirical fact straightforwardly, and in this respect is superior to such proposals as Uriagereka (1999), Nunes and Uriagereka (2000), and Stepanov (2007). Uriagereka (1999), and Nunes and Uriagereka (2000) argue that constituents are Spelled-Out/Transferred when Merged as specifiers, ruling out extraction from specifiers across the board (see also Narita (2011), who further pursues this idea under a multiple Transfer framework). Stepanov (2007), criticizing this analysis, argues that only
extraction from moved constituents (i.e. specifiers created by movement) causes ungrammaticality due to Chain Uniformity. Given that extraction from moved constituents is not always impossible, however, these analyses cannot account for the data we have considered in this section. Under our explanation, on the other hand, the freezing as well as non-freezing effects on extraction from specifiers follow from the proposed system of phase-based derivations.

4.2. Multiple Wh-Questions

If, as we have argued, step-by-step derivations make a wh-phrase embedded in the subject get out of the search domain of EF and the probe cannot find its goal because of the movement of the subject, then the ungrammaticality should not arise if there is another wh-phrase in the derivation and this wh-phrase is in the search domain of EF to establish a probe-goal relation with C. This “rescue by another wh-phrase” is confirmed by (30) (see Fiengo et al. (1988: 91) among others):

(30)  a. Who, did [pictures of who] t_j please t_i?
     b. Who, were [stories about whom] being told t_i to t_j?

In (30), the wh-phrases other than those embedded in the subjects move to Spec-C and the examples are well-formed. Let us take (30a) first and consider its derivation in detail. In the lower v*P phase, EF in v* raises the verbal object who to its edge for the PIC (= (31a)). In the next CP phase, who at the edge of v*P is in the search domain of C and can be a goal for its EF, which thus raises the wh-phrase to Spec-C (= (31b, c)), while the φ-features of C, which are inherited by T, Agree with the φ-features of the subject, which is raised to Spec-T upon its Agree (= (31c, d)):

(31)  v*P Phase
      a. [v*P who, [DP pictures of who] v* [VP please t_i]]

      CP Phase
      b. [CP C{EF} [TP T{φ} [v*P who, [DP pictures of who] v* [VP please t_i]]]]
      c. [CP who, C{EF} [TP T{φ} [v*P t_i [DP pictures of who] v* [VP please t_i]]]]
      d. [CP who, C{EF} [TP [DP pictures of who] T{φ} [v*P t_i [t_j v* [VP please t_i]]]]]

As shown, the EF-probe in C can successfully find its goal. As a result, the derivation converges in (31).
Essentially the same argument applies to the derivation without a strong v*P phase, which does not necessitate intermediate movement for cyclic Transfer (the PIC). (30b), which is a passive sentence, is an example of this case. Consider (32):

\[(32)\]

\[\begin{array}{l}
a. [vP v [vP were being told [DP stories about whom] to who]] \\
b. [CP C_{\{EF\}} [TP T_{\{\phi\}} [vP v [vP were being told [DP stories about whom] to who]]] \\
c. [CP C_{\{EF\}} [TP [DP stories about whom]j T_{\{\phi\}} [vP v [vP were being told t_j to whom]]] \\
d. [CP who_i C_{\{EF\}} [TP [DP stories about whom]j T_{\{\phi\}} [vP v [vP were being told t_j to t_i]]]
\end{array}\]

The vP phase of (30b) is a weak phase and its v does not bear EF. Hence, neither whom nor who undergoes intermediate movement to the v-edge (= (32a)). In the CP phase, the two probes (EF and \(\phi\)) start searching for their goals upon the Merge of C. Since there is no wh-phrase at the edge of vP, the \(\phi\)-probe finds the subject DP first, and upon the Agree in \(\phi\)-features, the subject raises to Spec-T, just as in (9c, d) (= (32b, c)). The EF-probe, on the other hand, continues searching and can successfully find who in VP, as a result of which the wh-phrase moves to Spec-C for the EF (= (32c, d)). Unlike in (8), the EF in C is satisfied by the Merge of who to Spec-C, and the derivation is saved thanks to this wh-phrase.

We have shown that the grammaticality of (30) is also derivationally explained. As is expected, the extraction of the embedded wh-phrases induces ungrammaticality in (30), and essentially the same explanation that we have claimed for (8) can be given to (33): that is, the subjects raise earlier in the derivation for \(\phi\)-feature Agreement and the raising moves the embedded wh-phrases out of the search domain of EF before the probe can find them:  

\[(33)\]

\[\begin{array}{l}
a. *Who_i did [pictures of t_i] t_j please who? \\
b. *Whom_i were [stories about t_i]j being told t_j to whom?
\end{array}\]

The derivational explanations we have provided for (30) and (33), which we have argued follow from step-by-step derivations inside phases, suggest that C does not enter into multiple probe-goal relations with several wh-

---

15 In (33a), minimal search is also violated: at the derivational stage shown in (31b), who_i at the edge of v*P is closer to EF than who embedded in the subject.
phrases in such language as English, where only a single \textit{wh}-phrase moves to Spec-C. Recall that the search into the subject is derivationally impossible, unless it stays in situ. If so, there is only a single instance of \textit{wh}-movement in multiple \textit{wh}-questions such as (30). Given this conclusion, it follows that multiple \textit{wh}-questions in general do not involve covert movement of in-situ \textit{wh}-phrases: those \textit{wh}-phrases which do not move overtly do not move covertly for their interpretations, either, staying in their first-Merged positions throughout derivations (see Simpson (2000) for persuasive arguments against covert \textit{wh}-movement). To the extent that our discussion is correct, in-situ \textit{wh}-phrases in multiple \textit{wh}-questions are interpreted by interpretive mechanisms in the semantic component, as proposed, say, in Baker (1970) and Stroik (1992).

In summary, in section 4, we have discussed extraction from XP and multiple \textit{wh}-questions as empirical predictions from our analysis of subject condition effects proposed in the last section. We have demonstrated that these predictions are endorsed and receive straightforward accounts under step-by-step derivations. Through the discussion in this section, we can conclude that derivations inside phases are processed in a step-by-step manner.

5. Some Consequences of Step-by-Step Derivations

We have claimed that operations inside phases are not simultaneous, applied in a parallel manner but are processed step by step, showing the validity of this claim with subject condition effects and other empirical examples. To the extent that our claim is on track, it leads us to take a fresh look at some aspects of syntax that have been discussed in the literature. In this section, we focus on this topic and discuss two theoretical consequences of our claim.

5.1. Subject Condition Effects Redux

We have argued that extraction from the subject is impossible in English because unless the expletive \textit{there} is Merged in, the subject raises to Spec-T for inherited φ-features in T upon its Agree, as a result of which a \textit{wh}-phrase embedded in the subject is forced to move out of the search domain of EF. It has been reported, however, that extraction from the raised subject in Spec-T is not always ungrammatical even in English. To see this, consider the following examples, which are grammatical unlike those in (8):
(34) a. It was the CAR (not the TRUCK) of which, the (driver, picture) of which, was found.
   b. Of which car was the hood of which, damaged?

Chomsky (2008) discusses examples such as (34) and argues that operations are simultaneously applied inside phases. As an example, consider (34a). Under simultaneous derivations, the wh-phrase and the subject enter into probe-goal relations with EF and φ-features, respectively, at the same time at the C phase level, undergoing parallel movement:

\[
\begin{array}{c}
\text{a. } \left[ \text{CP } C_{[\text{EF}] \left[ \text{TP } T_{[\phi]} \left[ \text{was found } \text{DP the (driver, picture) of which} \right] \right]} \right] \\
\text{b. } \left[ \text{CP of which } C_{[\text{EF}] \left[ \text{TP } \text{DP the (driver, picture) } T_{[\phi]} \left[ \text{was found } t_1 \right] \right]} \right]
\end{array}
\]

It should be recalled, however, that simultaneous derivations also wrongly rule in ungrammatical examples such as (8), which we have discussed in (10). On the other hand, the analysis of subject condition effects we have proposed will rule out the data in (34); as we have argued, subject raising moves a wh-phrase embedded in the subject out of the search domain of EF. Then, to the extent that our proposal is on track, it follows as one of its consequences that wh-phrases should not be extracted from subjects in grammatical cases like (34). In the rest of this section, we argue that this conclusion is correct and propose one possible analysis of (34) under step-by-step derivations, which we see is empirically endorsed.

A closer look at the data will reveal that grammatical extraction from subjects requires the pied-piping of prepositions (Chomsky (1986), Kuno and Takami (1993), Ross (1967)). As illustrated in (36), once the extracted elements in (34) are wh-phrases without prepositions, the data will turn ungrammatical. Consider (36):

(36) a. *It was the CAR (not the TRUCK) which, the (driver, picture) of which, was found.
   b. *Which car was the hood of which, damaged?

Given this observation that the pied-piping of prepositions has to do with the grammatical status of (34), we argue that wh-phrases with prepositions are not extracted from subjects but are moved from adjoined positions. To be concrete, extraposition underlies (34) and holds the key to grammatical extraction from subjects. Various approaches have been proposed for extraposition in the framework of generative grammar. For example, Mizuguchi (2009) argues that extraposition of a prepositional phrase (PP) does not involve movement but adjunction (External Merge) to XP, proposing that...
an extraposed PP is associated with its host phrase not through movement but through an interpretive mechanism based on mutual c-command relations (see also Culicover and Rochemont (1990) for a base-generation analysis). In this paper, following this analysis of extraposition, we argue that extraposition underlies grammatical extraction from subjects, proposing an extraposition analysis for (34): prepositional wh-phrases are Externally Merged (or adjoined) to XP as extraposed elements, and move from there to the edge of CP. As shown in (37), extraposition of PPs is possible in such examples as (34) (see also (43b)).

(37) a. A driver/picture was found of that car.
   b. The hood was damaged of Bill’s car.

With this proposal in mind, consider the derivation of (34a):

(38) a. \([CP \ C_{[EF]} \ [TP \ T_{[\phi]} \ [v \ P \ [VP \ [VP \ [VP \ was \ found \ the \ (driver, \ picture) \ [PP \ of \ which]]]]]]\]
   b. \([CP \ [PP \ of \ which] \ C_{[EF]} \ [TP \ T_{[\phi]} \ [v \ P \ [VP \ [VP \ was \ found \ the \ (driver, \ picture) \ t_i]]]]\]
   c. \([CP \ [PP \ of \ which] \ C_{[EF]} \ [TP \ the \ (driver, \ picture) \ T_{[\phi]} \ [v \ P \ [VP \ [VP \ was \ found \ t_j \ t_i]]]]\]

In (38a), when C with EF and \(\phi\)-features is Merged in, the probes start searching, and in this process the \(\phi\)-probe is inherited by T. Under the extraposition analysis, the prepositional wh-phrase \([PP \ of \ which]\) will be in the search domain of EF when the probe searches for it, because it is adjoined to VP and is not embedded in the subject. Thus, C can enter into a probe-goal relation with the prepositional wh-phrase and its EF raises it to Spec-C.

---

16 It has been reported that extraposition is degraded when host DPs are definite. Then, the question is why (34) is not degraded. We do not have enough space to address this question and have to leave the discussion for future, but it can reasonably be considered that extraposition from definite DPs is possible in syntax but is ruled out for discourse/functional reasons (see Johnson (1985) and Nakajima (1995)). This is endorsed by examples like (i), where extraposition from a DP with a definite article is not ungrammatical ((ia) is from Drummond (2009) and (ib) from Nakajima (1995: 24)):

(i) a. The picture was accidentally taken of John at his birthday party on Tuesday.
   b. The destruction was ordered of the new building in LA.

These examples suggest that definiteness effects of extraposition depend on non-syntactic factors. In the case of (34), a possible explanation will be that wh-movement of extraposed PPs can somehow remove discourse/functional anomalies brought about by extraposition.

17 Empirical evidence suggests that extraposed elements can be adjoined to TP or VP (Culicover and Rochemont (1990), Hirata (1996), Mizuguchi (2009) and Tanaka (2011), among others). For our purpose here, we assume that extraposed PPs are adjoined to VP.
On the other hand, the φ-probe continues searching, to find [the (driver, picture)] as its goal, and raises it to Spec-T (= (38b, c)). Under the proposed analysis, the grammaticality of (34) is explained on par with that of (39), where adjunct prepositional wh-phrases are moved from VP-adjoined positions:

(39) a. In which city_i did he meet Mary_t?
b. With whom_i did she go to the museum_t?

One straightforward piece of evidence for the extraposition analysis of (34) is that extraposition from subjects, which appears to involve extraction, does not induce ungrammaticality. Consider (40):

(40) [DP A review ____] will appear [PP of this article].

In (40), just as in (8) and (36), part of the subject, which is interpretively associated with the subject as its complement, is in a position away from it. However, (40) is grammatical while (8) and (36) are not. This suggests that extraposition from subjects is essentially different from extraction from subjects in its derivation. The grammaticality of (34) receives a straightforward account if extraposition is assumed.19

Another piece of evidence for the extraposition analysis comes from what is called the unaccusative hypothesis on extraposition. Extraposition has been known to be sensitive to verb types, and is possible only from D-structure Xº-sister positions (Hirata (1996)). This generalization says that extraposition is possible only with unaccusative (including passive) as well as ergative verbs. Consider the following examples (the examples here are cited from Johnson (1985) and Nakajima (1995)):20

(41) a. *[A man ____] hit Mary [PP with hostility toward her].b. *[A new book ____] has attracted many people [PP about the origin of human language].

(42) a. *[A man ____] whispered [PP from Nuie].b. *[A man ____] ran [PP from the EPA].

18 In (40) and elsewhere, underlines in phrases indicate the positions from which PPs are extraposed and with which they are interpretively associated. In (40), [PP of this article] is thus interpreted as the complement of the noun review.


20 In this paper, we do not discuss why extraposition is impossible with (41) and (42), since it is not directly relevant to our discussion in this paper. What is important for our purpose here is that extraposition underlies the (un)grammaticality of such data as (34) and (45). See Mizuguchi (2009) and Tanaka (2011) for recent discussion of the relevant generalization in the framework of the Minimalist Program.
The examples in (41) are examples of transitive verbs and those in (42) are examples of unergative (or “true” intransitive) verbs. In these examples, extraposition is not allowed. On the other hand, examples such as (40), which allow extraposition, include unaccusative and ergative verbs, the subjects of which are Externally Merged as the objects of the predicates (that is, Merged to “D-structure Xº-sister positions”). More examples are provided below:

(43) a. [Men ___] appeared [pp from Tanzania]
    b. [A man ___] was called [pp with blond hair].

(44) a. [A storm ___] followed [pp from the north].
    b. [A picture ___] stands in the hallway [pp by Picasso].

Given this generalization, the extraposition analysis of (34) predicts that prepositional which-phrases cannot be “extracted” from the subjects of transitive and unergative verbs. This prediction is confirmed by the following data:

(45) a. *Of which car did the (driver, picture) cause a scandal?
    b. *Of which car did the driver run in the park?

Under the extraposition analysis, the ungrammaticality of (45) is thus analyzed as that of (41) and (42). Furthermore, the analysis can not only give a straightforward explanation to (45); it can also explain, without any other stipulations, why the examples in (45) are ungrammatical while those in (34) are not. Chomsky (2008: 147–148), discussing (45) in the framework of simultaneous derivations, argues that the examples are ungrammatical because which in the subject is embedded in the lower phase (vp), which has already been passed in the derivation, and suggests that the ungrammaticality can be reduced to a locality condition on which-probing. This explanation, however, is not empirically correct. As we have observed in (20) and (21), which are repeated below, a which-phrase can be extracted from the transitive subjects in the edge of vp:

(20) [CP Was_i haben [TP denn [vP DP t_i für Ameisen] einen
             what have indeed for ants a
    Postbeamten gebissen]]?
    postman bitten
    ‘What kind of ants have bitten a postman?’ (Diesing (1992: 32))

---

21 In English, subject raising is obligatory and a transitive subject always vacates the edge of vp; moreover, transitive expletive constructions as found in Icelandic are also disallowed. The English counterparts to (20) and (21) are impossible and English cannot be used to see if a which-phrase can be moved out of the transitive subject in the edge of vp.
(21) (?)[\textsubscript{CP} [De qué conferenciantes], te parece [\textsubscript{CP} que me van of what speakers to-you seems that me will a impresionar las propuestas tij]? to impress the proposals ‘Of which speakers does it seem to you that will impress me the proposals?’ (Uriagereka (1988: 118))

These examples argue that the terms in the edge of the lower phase are indeed accessible to a probe in the next higher phase. The contrast between (45) and (20)/(21) follows from our proposal without any stipulations on locality.

Finally, extraposition is known to be impossible if a PP is extraposed from a DP which is further embedded in another DP (Akmajian (1975), Ross (1967)). To see this, consider (46):

(46) a. *[\textsubscript{DP} A review of [\textsubscript{DP} a new book ____]] came out yesterday [\textsubscript{PP} about French cooking].

b. [\textsubscript{DP} A review of [\textsubscript{DP} a book ____]] appeared last year [\textsubscript{PP} by three authors].

(47) a. [\textsubscript{DP} A new book ____] came out yesterday [\textsubscript{PP} about French cooking].

b. [\textsubscript{DP} A book ____] appeared last year [\textsubscript{PP} by three authors].

In (46a), the extraposed PP [about French cooking] cannot be interpreted as the modifier of [a new book]; likewise, (46b) does not allow the interpretation in which [by three authors] modifies [a book]. On the other hand, these interpretations are possible with (47), where the DPs are not embedded in another DP.

Given this restriction on extraposition, the proposed extraposition analysis predicts the same boundedness effect on prepositional wh-phrases: due to the boundedness effect on extraposition, prepositional wh-phrases cannot be “extracted” out of a DP which is further embedded in another DP. This prediction is borne out. Consider the following examples:

(48) a. *About which subject did a review of a new book come out yesterday?

b. By how many authors did a review of a book appear last year? (Mizuguchi (2009: 324))

(49) a. About which subject did a new book come out yesterday?

b. By how many authors did a book appear last year?

The wh-counterpart of (46a) is ill-interpreted while that of (46b) only allows a single interpretation. Thus, (48a) does not have the interpretation in which [About which subject] is the modifier of [a new book];
likewise, (48b) is well-formed only under the interpretation that asks the number of the authors who wrote a review, not a book: [By how many authors] is interpretively associated with [a review of a book], not with [a book]. Meanwhile, if [a new book] and [a book] are not embedded as in (49), the interpretations ruled out in (48) are available, just as in (47). The interpretations of (48) and (49) parallel those of (46) and (47), and the extraposition analysis of extraction from subjects can give a straightforward explanation to the interpretive differences in question.

It should be noted that the examples in (8) can still be ruled out under the extraposition analysis proposed here. As we can see from (50), DPs, unlike PPs, cannot be extraposed, hence adjoined, and wh-phrases alone cannot be Merged to VP as extraposed elements:

\[(50) *\text{{[DP A review of ____] came out yesterday [DP this article].}}\]

In this section, as one consequence of our claim in this paper, we have reconsidered “grammatical” extraction from subjects. Noting that prepositional wh-phrases can be extracted from subjects and that PPs can be extraposed, we have argued that extrapoision, which Externally Merges an extraposed phrase with XP, explains the grammaticality of (34). We have validated the extraposition analysis by demonstrating that prepositional wh-phrases show parallelisms with extraposed prepositional phrases.

5.2. Subject Wh-Movement

The other consequence of our claim is concerned with subject wh-movement. Under the assumption that phase-internal derivations are simultaneously operated in a parallel manner at the phase level, Chomsky (2008) argues that a subject wh-phrase moves to Spec-C and Spec-T simultaneously from its first-Merged position, creating two distinct chains. To see this, consider (51) and its derivation illustrated in (52):

\[(51) \text{Who arrived?}\]

\[(52) \text{a. } [\text{CP } C \text{ [TP } T \text{ [vP } v \text{ [VP arrived who]]}]\]

\[\text{b. } [\text{CP who}_i \text{ C [TP who}_j \text{ T [vP } v \text{ [VP arrived } t_{i,j}]]}]\]

At the C phase level, the simultaneous movement of who creates two chains \((\text{who}_{i}, t_{i})\), \((\text{who}_{j}, t_{j})\), with there being no direct relation between \(\text{who}_{i}\) and \(\text{who}_{j}\).

Given our claim, this derivation must be reconsidered. If phase-internal derivations proceed step by step, operations (A-movement and A'-movement) are applied one by one, which means that a subject wh-phrase
should not move simultaneously as shown in (52). In subject \textit{wh}-movement, the goals of the \textit{EF}-probe and the \textit{φ}-probe are the subject \textit{who}, and the two probes find it simultaneously through their minimal search. As we have argued, since step-by-step derivations rule out simultaneous movement, which will increase computational complexity, the Internal Merge of \textit{who} should take place one step at a time. However, there will then arise computational ambiguity in the application of the Internal Merge: which Internal Merge (A-movement or A′-movement) applies first? This ambiguity in syntax is unwanted based on third-factor considerations (efficient computation) and should be removed from syntax.

In addition to this theoretical argument, there is an empirical problem with (52) as well. Provided that \textit{who}_{i} and \textit{who}_{j} form distinct chains, with there being no direct relation between the two, the head of each chain should be pronounced, just as in (5). As shown in (53), however, this is not the case in subject \textit{wh}-movement:

\begin{equation}
\text{(53) } ^* \text{Who}_{i} \text{ who}_{j} \text{ arrived } t_{i,j} ?
\end{equation}

\begin{equation}
\text{(5) } \text{What}_{i} \text{ did John}_{i} t_{j} \text{ read } t_{i}?
\end{equation}

Unlike in (5), only the head of A′-chain (\textit{who}_{i}, t_{i}) is pronounced in (51). The above arguments suggest that (52) is not the derivation of subject \textit{wh}-movement.

We propose that subject \textit{wh}-movement creates only one chain (\textit{who}_{i}, t_{i}). Recall our discussion in section 2. In the phase syntax we have in mind, EF and φ-features are properties of C and probe from the phase head upon their Merge, in the process of which φ-features are inherited by T (“feature-inheritance via probing”). This means that EF in C and inherited φ-features in T start out as a single probe. We argue that if EF and φ-features target different goals as in (5), the probing bifurcates and they probe for their goals, each entering into its own probe-goal relation; on the other hand, if the goals of these probes are one and the same element, the probing does not bifurcate and they enter into a single probe-goal relation with their shared goals. We propose that this is what happens in subject \textit{wh}-movement in (51), where the goals of EF and φ-probes are \textit{who}:

\begin{equation}
\text{(54) } [ \text{CP } C_{[EF]} \text{ TP } T_{[φ]} \text{ vP ... who ... } ] ] \text{ Probe-Goal (C/T-who)}
\end{equation}

Given the above argument, \textit{who} enters into a single probe-goal relation with the two probes. On the assumption that movement (Internal Merge) presupposes a probe-goal relation, there is only a single application of Internal Merge as a result of the probe-goal relation in (54). EF borne by C can only be satisfied by the Merge of an (operator) element to its edge;
if *who* were Merged to Spec-T, the EF would never be satisfied in any fashion. Thus, the Internal Merge targets C, not T, and the *wh*-phrase is Merged to its edge. The proposed derivation is (55), where the theoretical and empirical problems we have noted do not arise:

(55) \[[CP \{who\}, C_{EF}] [TP T_{\{\phi\}} [v_P v [v_P \{arrived t]\}]]\]

On the other hand, there is an alternative mechanism to satisfy the so-called EPP (property) of T. Recall that the EPP of T is a side effect of \(\phi\)-features inherited by T: the EPP is EF from \(\phi\)-features (see fn. 3). In (54), since \(\phi\)-features are inherited on their way to probing, T bears EF for the inheritance, having the EPP property. We propose that in (55), this EF (or EPP) can be satisfied via T-to-C movement. Mizuguchi (2008) proposes, building on Alexiadou and Anagnostopoulou (1998) (see also Chomsky (2000)), that T-to-C movement, just as V-to-T (more precisely, V-to-\(v^*\)-to-T) movement, can function to satisfy the EPP property of T, and argues that the failure of this movement, which leaves the EPP property unsatisfied, explains *that*-trace effects. See Mizuguchi (2008) for details. Following this proposal on the EPP satisfaction, we argue that in (55), where a single probe-goal relation Internally Merges *who* to Spec-C to meet its EF, the EPP property of T is satisfied by the mechanism of T-to-C movement.22

Data from various languages empirically support the derivation proposed in (55). First, consider agreement phenomena in Kinande. This language shows canonical subject-verb agreement, which is morphologically realized as *a*- on verbal predicates. An interesting observation is that when the subject undergoes *wh*-movement, the canonical subject-verb agreement marker disappears, and an anti-agreement marker *u*- has to show up instead. To see this contrast, consider (56) and (57), where the subject undergoes move-

22 See also Miyoshi and Tozawa (2011), who independently suggest that the EPP can be satisfied by head-adjunction.

If T-to-C movement can satisfy EF in T, then the prediction is that in (i), the subject can stay in Spec-\(v^*\):

(i) Did John kiss Mary?

Evidence suggests, however, that the subject does move to Spec-T in (i). We argue that T-to-C movement for the EF satisfaction is a last resort operation and is triggered only when a DP cannot move to Spec-T. This argument is not unreasonable if T-to-C movement, unlike V-to-\(v^*\) movement, is a phonological process (cf. Chomsky (2001)) and is syntactically motivated only when the subject is not Internally Merged to Spec-T as in (55). Thus, T-to-C movement in (i) plays no role in the EF satisfaction. I thank a reviewer for bringing the discussion of (i) to my attention.
Subject raising in (56) shows canonical agreement while subject wh-movement in (57) shows anti-agreement. In the derivations of (56) and (57), φ-features probe from C and the subjects Agree with T in its inherited φ-features. Hence, it can be supposed that φ-feature Agreement is irrelevant to the distinction between canonical agreement and anti-agreement. Given the irrelevance of φ-feature Agreement, the agreement contrast above most naturally follows from the assumption that the morphological realization of the canonical agreement marker is contingent upon movement into Spec-T: the marker a- appears (or is realized in the phonological component) iff an item Agrees in φ-features and is moved to Spec-T. Then, the absence of the canonical subject-verb agreement marker a- in (57a) suggests that the wh-subject moves at one fell swoop to Spec-C without moving to Spec-T. Anti-agreement from Kinande provides evidence for the derivation proposed in (55).23

Another piece of evidence for (55) is found in West Ulster English discussed in McCloskey (2000). This variety of English shows an interesting contrast in subject movement:

(58) a. Who was arrested all t₁ in Duke Street?
   b. *They were arrested all t₁ last night.

As shown in (58), in West Ulster English, subject wh-movement can strand a quantifier in the object position while subject raising from the object position cannot. McCloskey (2000) argues that (58b) is ungrammatical because...
the DP associated with the quantifier moves to the specifier of the DP headed by the quantifier ([DP all-D [DP they]]) on its way to Spec-T in order to strand the quantifier:

\[(59) \quad [CP [TP they, [were [VP arrested [DP t_i [D' all t_i]] last night]]]]\]

DP-movement to Spec-D is considered an instance of A'-movement (provided that the edge of DP is an A'-position). Then further movement from this position to Spec-T, an A-position, will count as "improper" movement (A-to-A'-to-A), which results in ungrammaticality.25

Given this explanation of (58b), the grammaticality of (58a) implies that no improper movement of who takes place; if the wh-phrase moves to Spec-T from Spec-D as shown in (59), the ungrammaticality would result because of improper movement:

\[(60) \quad [CP who, [TP who, was arrested [DP t_i [D' all t_i]] in Duke Street]]\]

*Improper Movement

It can thus be concluded that in subject wh-movement, the subject does not move to Spec-T, which supports the one-fell-swoop derivation proposed in (55).

We now turn to the evidence for the other side of (55): that is, the EPP property is still present in T for (the inheritance of) φ-features and is satisfied by grammatical means other than the Internal Merge (i.e. movement) of a subject wh-phrase. If, as we have claimed, the subject does not move to Spec-T in subject wh-movement, it is then expected that the derivation (55) allows another XP to be Merged to Spec-T for its EF/EPP. This prediction is confirmed by languages such as Yiddish and Swedish. Consider the following examples:

---

24 Provided that not all D-heads allow movement to their edges, we can reasonably assume that the movement of they to Spec-D is driven by EF inherent in quantifier heads like all. Thus, who cannot move to Spec-D on its way in (9) and cannot move out before subject movement. I thank a reviewer for clarifying this point.

25 We assume in this paper that the A/A'-distinction is deducible from phases: the specifier/edge of a phase head (e.g. C, v*/v, D) counts as an A'-position while that of a non-phase head (e.g. T, V, N) as an A-position (cf. Chomsky (2008)). Provided that the A/A'-distinction is crucial to the interfaces and that phases constitute self-contained syntactic objects feeding the interface systems through cyclic Transfer (Epstein and Seely (2006), Hiraiwa (2005)), it is not unreasonable to suppose that the A/A'-distinction is expressed in phases and follows from the phase/non-phase distinction.
(61) Ikh veys nit [CP ver [TP es iz gekumen]].
I know not who Expl has come
‘I don’t know who has come.’ (Diesing (1990: 68))

(62) Vilken elev trodde ingen [CP att [TP han skulle fuska]]?
which pupil thought nobody that he would cheat
‘Which pupil didn’t anyone think would cheat?’ (Engdahl (1982: 166))

These data show that expletives and resumptive pronouns can be Merged to Spec-T in subject wh-movement, suggesting that the EPP property is present in T for inherited φ-features.\(^{26}\) In addition, the data support our argument that a subject wh-phrase moves to Spec-C from its first-Merged position without moving to Spec-T at the same time: if a wh-subject moves to Spec-T for φ-features as shown in (52), it will prevent expletives and resumptive pronouns from being Merged there. Evidence from English (63) demonstrates that subject raising disallows the Merge of expletives:

(63) (*There) [posters of the President]i (*there) were t_i all over the town.

To summarize this section, we have considered, as consequences of our claim, two aspects of syntax that have been discussed in the literature. To the extent that phase-internal derivations proceed step by step, grammatical extraction from subjects and subject wh-movement must be given fresh analyses. We have argued that grammatical extraction from subjects does not involve extraction and that a subject wh-phrase does not undergo simultaneous, parallel movement to Spec-C and Spec-T. These proposals, we have shown, are theoretically valid and are empirically endorsed by examples from various languages.

6. Conclusion

In this paper, we have explored phase-internal derivations and considered their consequences for other aspects of syntax. The central claim in this paper has been that derivations proceed step by step from beginning to end. With the discussion of subject condition effects, we have argued that

\(^{26}\) Miyoshi (2009) suggests, based on (56) and (57), that features (including EPP) are not inherited in subject wh-movement. This suggestion, however, is questionable if feature-inheritance is independently required (see fn. 1). Furthermore, as we have seen, there is evidence for an EPP property in T, which implies that features are indeed inherited by T in subject wh-movement.
phase-internal derivations are not applied all at once in a parallel manner at the phase level, providing theoretical and empirical arguments against this proposal. We have shown that our claim can explain both grammatical and ungrammatical extraction from constituents; in addition, grammatical behaviors of multiple wh-questions also follow from our proposal. The proposed step-by-step derivations have brought two theoretical consequences, and we have proposed novel analyses for grammatical extraction from subjects and subject wh-movement, which are, as we have seen, theoretically and empirically valid. To the extent that our claim in this paper is correct, representational properties of syntax are eliminated from phase-internal computation: syntax is step-by-step throughout its derivations. The discussions in this paper thus strengthen the view that step-by-step derivations explain language.

REFERENCES


[received September 20, 2012, revised and accepted January 31, 2013]

---

Division of Language Education, Premedical Sciences
Dokkyo Medical University
880 Kita-Kobayashi, Mibu, Shimotsuga
Tochigi 321-0293
e-mail: mizuguch@dokkyomed.ac.jp