

Constraints on Movement*

Mayumi Hosono
Keio University

Abstract

I argue, contra Chomsky (2013, 2015), that internal merge may not be free. It is shown that the Criterial Position (Rizzi 2006, 2010, 2015) is the position in which a raised category completes the valuation of all of its own unvalued features. The Halting Problem, the Extended Projection Principle, and the Empty Category Principle (as well as the disappearance of that effect) are all fully accounted for in terms of feature valuation. This unified account derives from the corollary of the derivational system of Labeling Algorithm (Chomsky 2013, 2015), in which labeling results from feature valuation. In Scandinavian Object Shift and Icelandic Stylistic Fronting, a category that does not have unvalued features can move from/into the Criterial Position (Hosono 2016). Following Chomsky (2013, 2015), who claims that (both external and internal) merge is free, movement from/into the Criterial Position would be allowed to occur with its legitimacy determined by filtering at the interfaces. If such movement is considered to occur exceptionally in narrow syntax, constraints on movement should exist. The argument that far more constraints on movement are imposed by phonology than have been considered so far (Hosono 2016, Richards 2016) indicates not only that internal merge may not be free, but also that narrow syntax will be *crash-proof* (Frampton and Gutmann 2002): the derivational mechanism will produce only well-formed structures that conform to the requirement by phonology, with no filters assumed.

1. Introduction

A sentential element is frozen in some structural positions, the problem called the *Halting Problem* (HP, Rizzi 2006, 2010, 2015). In (1a), the *wh*-object *which dog* moves from its original position to [Spec,(embedded)CP]. When it moves out of that position, the sentence is ungrammatical; see (1b). Such a position as [Spec,(embedded)CP] from where a sentential element cannot move up further is called the *Criterial Position* (CriP).

* Thanks to Johan Brandtler for his helpful comments for my former work which this paper is based on. Special thanks to Anders Holmberg for his insightful comments and helpful suggestions for the former version of this paper. Thanks also to Hisatsugu Kitahara for his insightful suggestions for a series of my works. Part of this work was presented in Keio University Linguistics Colloquium, January 2016. I would like to thank the participants for their many helpful comments.

- (1) a. You wonder [_{CP} [_Q which dog] C_Q John likes [_Q ~~which dog~~]].
 b. *[_Q which dog] do you wonder [_{CP} [_Q ~~which dog~~] C_Q John likes [_Q ~~which dog~~]]?

Another typical CriP claimed in the literature (Rizzi 2006, 2010, 2015) is the subject position, [_{Spec,TP}] traditionally, which is required to be filled by an overt subject in languages such as English (the *Extended Projection Principle* (EPP), Chomsky 1981, 1986, 1995):

- (2) *(John) kisses Mary.

An issue related to the EPP is the *Empty Category Principle* (ECP), which requires that a trace be properly linked with its antecedent (Chomsky 1981, 1986, 1995). In languages such as English, when a *wh*-subject moves to sentence-initial position, the overt complementizer *that* cannot appear, since the overt complementizer intervenes between the raised *who* and its trace (the *that-trace effect*, Chomsky 1981, 1986); see (3a). When the complementizer disappears, the ECP effect also disappears and the entire sentence is grammatical; see (3b).

- (3) a. *Who do you think that [_{wh} ~~read the book~~]?
 b. Who do you think \emptyset [_{wh} ~~read the book~~]?

In this paper, I argue, contra Chomsky (2013, 2015), that *internal merge* (IM) may not be free. It is shown that the CriP is the position in which a raised category completes the valuation of all of its own unvalued features. That is, a category to be raised must have some unvalued feature(s) which is valued by a head in its raised position; after it completes the valuation of all of its unvalued feature(s), it cannot move up further. The HP, the EPP, and the ECP (as well as the disappearance of that effect) are all fully accounted for in terms of feature valuation. This unified account derives from the corollary of the derivational system of *Labeling Algorithm* (LA, Chomsky 2013, 2015), in which labeling results from feature valuation. In Scandinavian *Object Shift* (OS, Holmberg 1986) and Icelandic *Stylistic Fronting* (SF, Holmberg 2000), a category that does not have unvalued features can move from/into the CriP (Hosono 2016). Following Chomsky (2013, 2015), who claims i) that both external merge (EM) and IM are free, and ii) that a *syntactic object* (SO) that is gibberish and not interpreted

appropriately is filtered out at the interfaces, movement from/into the CriP would be allowed to occur with its legitimacy determined by filtering at the interfaces. If such movement is considered to occur exceptionally in *narrow syntax* (NS), constraints on movement should exist. The argument that far more constraints on movement are imposed by *phonology* (PHON) than have been considered so far (Hosono 2016, Richards 2016) indicates not only that IM may not be free, but also that NS will be *crash-proof* (Frampton and Gutmann 2002): the derivational mechanism will produce only well-formed structures that conform to the requirement by PHON, with no filters assumed.

This paper is organized as follows. Section 2 introduces the LA derivational system (Chomsky 2013, 2015). Section 3 examines the properties of the CriP in detail. Sections 4 presents two movement cases, Scandinavian OS and Icelandic SF. Section 5 discusses constraints on movement. Section 6 concludes this paper.

2. Labeling Algorithm and Free Merge

Through the theoretical transition (Chomsky 2004, 2008), Chomsky (2013, 2015) completely eliminates the constraint on movement, contra Chomsky (2001), who claimed that movement occurs when a semantic difference is reflected on the interfaces. Under the long tradition of *X'-bar* theory (Chomsky 1981, 1986, 1995), a head automatically projected itself. Contrary to this tradition, Chomsky (2013, 2015) claims that in configuration [XP, YP], there is no necessity to assume that Y, for instance, always projects. But a SO needs a label so that it can be interpreted at the interfaces. It is labeled in the derivation by LA, which is claimed to be a minimal search of computation.¹

Labeling of SOs proceeds as follows. First, in configuration [v^*/C , XP] where the phase head, either v^* or C, merges to a maximal projection, XP, LA takes the label of that phase head, which results in either [$\langle v^* \rangle v^*$, XP] or [$\langle C \rangle C$, XP].

Secondly, in configuration [R/T, XP] where either a verbal root R or T, which is not a phase head and weak by assumption, merges to XP, either XP itself or a category inside XP, say YP, moves to the Spec of that head to strengthen the head. The raised category and the head share some features, e.g. ϕ -features. Feature sharing between two categories includes the valuation procedure in which

¹ See Collins (2002) and Seely (2006), who claim within the *phase* (Chomsky 2000) framework that labels should be eliminated from the syntactic representation.

one's unvalued features are valued by the other's valued counterpart. After feature valuation, LA takes the shared φ -features and labels the entire projection $\langle\varphi,\varphi\rangle$, which results in either $[\langle\varphi,\varphi\rangle \text{XP}_{[\varphi]} [\text{R}_{[\varphi]}, \text{XP}]]$ or $[\langle\varphi,\varphi\rangle \text{YP}_{[\varphi]} [\text{T}_{[\varphi]} [\text{XP} \dots \text{YP} \dots]]]$.²

Thirdly, in configuration $[\text{XP}, \text{YP}]$ where two maximal projections, XP and YP, merge, one way to label the projection is that either one of them moves out of that configuration, and the remaining maximal projection offers its label. That is, after one of the maximal projections moves out, its copy in the original position becomes part of a discontinuous object. LA is blind to such an element and takes instead the head of the remaining maximal projection as the label of the entire configuration, which results in either $[\text{XP} [\langle\text{Y}\rangle \text{XP}, \text{YP}]]$ or $[\text{YP} [\langle\text{X}\rangle \text{XP}, \text{YP}]]$. The other way is that XP and YP share some features, e.g. φ -features, and feature valuation occurs between them; LA takes the shared features and labels the projection $\langle\varphi,\varphi\rangle$, which results in $[\langle\varphi,\varphi\rangle \text{XP}_{[\varphi]}, \text{YP}_{[\varphi]}]$.

On the basis of the labeling procedure introduced above, the derivation of *John kisses Mary* proceeds as illustrated in (4), which is the final representation of the derivation. Following Chomsky (2013, 2015), let us consider the derivational process until when $\beta\langle\text{v}^*\rangle$ (= v^*P , with the traditional notation) is transferred.

(4) $\text{C} [\langle\alpha\langle\varphi,\varphi\rangle \text{John} [\text{T} [\beta\langle\text{v}^*\rangle \text{John} [\text{kisses}(=\text{R})+\text{v}^* [\gamma\langle\varphi,\varphi\rangle \text{Mary} [\text{kisses}(=\text{R}) [\delta \text{Mary}]]]]]]]]]$ ³

The verbal root R, *kisses*, merges to its internal argument, *Mary*. Since *kisses*(=R) is not a phase head and weak, *Mary* moves to $[\text{Spec},\gamma]$ to strengthen it. The phase head v^* merges to γ . Phasehood is inherited from v^* to R: functional features such as φ -features that are located in v^* are inherited to *kisses*(=R). *Kisses*(=R) and *Mary* in its Spec go on to feature valuation and the latter is assigned an Acc(usative Case). LA labels $\gamma \langle\varphi,\varphi\rangle$. *Kisses*(=R) moves to v^* to become a verbal category.⁴ Phasehood is activated in the original position of R. δ , the complement of R (, which is now vacuous), is transferred.

Then, the external argument of v^* , *John*, and T merge in turn. Since T is

² Here, I tentatively notate all φ -features, both unvalued ones and valued ones, as $[\varphi]$ for explanatory sake. I introduce a detailed derivational process soon below.

³ R has valued φ -features $[\varphi]$, which are inherited from v^* , and T has unvalued φ -features $[\text{u}\varphi]$, which are inherited from C. *John* and *Mary* each have an unvalued Case, $[\text{uCase}]$. I omit them from the notation in (4) for simplicity sake.

⁴ After R moves to v^* , v^* is claimed to be deleted due to its affixal nature (Chomsky 2015). In this paper, I notate $\text{R}+\text{v}^*$ in its final transferred position without a deletion line on v^* .

not a phase head and weak, DP in its complement, i.e. *John* in [Spec, β], moves to [Spec, α] to strengthen it. After *John* moves out, LA finds the phase head v^* and β is labeled $\langle v^* \rangle$. The phase head C merges to α . Phasehood is inherited from C to T, and functional features such as φ -features that are located in C are inherited to T. T and *John* in its Spec go on to feature valuation and the latter is assigned a Nom(inative Case); the unvalued φ -features on T are also valued by the valued counterpart of *John*.⁵ LA labels $\alpha \langle \varphi, \varphi \rangle$. Phasehood is then activated in T. $\beta \langle v^* \rangle$, the complement of T, including $\gamma \langle \varphi, \varphi \rangle$, is transferred.

As can be seen in the demonstration above, movement of a maximal projection does not always produce a new semantic effect such as focus or topic on it in the LA system. A category can of course receive a new interpretation in its raised position. The point here is that in the LA system, there is no constraint on movement that movement occurs when a semantic difference is reflected on the interfaces, contra Chomsky (2001). Eliminating any constraints on movement, Chomsky (2013, 2015) claims that both EM and IM are free. Merge can occur, having no recourse to any triggering features. Among SOs constructed in NS, those which are gibberish and not appropriately interpreted are filtered out at the interfaces (cf. Chomsky et al. 2017).

3. The Properties of the Criterial Position

Chomsky (2015) argues that the HP illustrated in (1a-b) is derived as follows. When *which dog* moves to [Spec,(embedded)C], feature valuation occurs between the unvalued [Q], [uQ], of *which dog* and the valued [Q] of C_Q .⁶ As illustrated in (5a), the projection of C_Q , i.e. β , is labeled $\langle Q, Q \rangle$, with the shared feature Q taken. When *which dog* moves out of [Spec, β], LA takes C_Q as the label of β . This means, according to Chomsky, that the embedded clause is interpreted as a *yes-no* question, a gibberish interpretation, which causes (5b) to be ungrammatical.⁷

⁵ Chomsky (2016) revises his claim, saying that after the phase head C and the raised subject go on to feature valuation, functional features such as φ -features located in C are inherited to T. In this claim, it is not clear for what reason φ -features must be inherited after the valuation procedure occurs between C and the subject. In this paper, I assume the system proposed by Chomsky (2013, 2015).

⁶ See Cable (2010), who claims that C has a valued [Q] in interrogatives in all languages.

⁷ See also Epstein et al. (2015), who claim that since the application of merge is free and ungrammatical cases such as (5b) are derived from their gibberish interpretation at the semantic

- (5) a. You wonder [$\beta_{\langle Q,Q \rangle}$ [which dog] C_Q John likes [~~which dog~~]]. (=1a)
 b. * $[\alpha$ [which dog] do you wonder [$\beta_{\langle C_Q \rangle}$ [~~which dog~~] C_Q John likes [~~which dog~~]]]? (=1b)

Chomsky's account for (5b) is quite odd, however. First, assume that feature valuation does not occur between *which dog* and the embedded C_Q when the former passes through the latter's Spec. Chomsky's account implies that *which dog* can move from its original position through [Spec, β] up to [Spec, α], escaping from LA's minimal search. When *which dog* with [uQ] moves to [Spec, β], LA would expect feature valuation to occur between *which dog* and the embedded C_Q with [Q] and would try to find their shared feature to label the embedded clause as soon as possible. It is quite odd to argue that the raised category LA would definitely target can escape from LA's minimal search and move up further.

Secondly, assume that *which dog* moves out of [Spec, β] after feature valuation occurs between *which dog* and the embedded C_Q. Chomsky's claim that LA takes C_Q as the label of β after *which dog* moves out of [Spec, β] indicates that feature valuation between the embedded C_Q and *which dog* raised to [Spec, β] can be cancelled. After feature valuation between the embedded C_Q and *which dog*, β is already labeled $\langle Q,Q \rangle$, with LA taking their shared features as the label of β . No argument is presented to support the claim that a once labeled SO can be relabeled.

Chomsky (2015) associates the EPP with the ECP, on the other hand. As we saw in section 1, the subject position must be overtly filled in English; see (6a). When the *wh*-subject *who* moves to sentence-initial position, the overt complementizer *that* cannot appear; see (6b). On the contrary, languages such as Italian allow the subject position to be empty; see (7a). When the *wh*-subject *chi* 'who' moves to sentence-initial position, the overt complementizer *che* 'that' can appear; see (7b). Thus, English both has the EPP requirement and obeys the ECP, whereas Italian neither has the EPP requirement nor obeys the ECP. The English case (6c), in which when the overt complementizer disappears, the ECP effect disappears too, is idiosyncratic, according to Chomsky (2015).

- (6) a. *(John) kisses Mary. (=2)
 b. *Who do you think [$\langle C \rangle$ that [α ~~who~~ T [$\beta_{\langle v^* \rangle}$ ~~who~~ read the book]]]? (=3a)
 c. Who do you think [$\langle C \rangle$ \emptyset [α ~~who~~ T [$\beta_{\langle v^* \rangle}$ ~~who~~ read the book]]]? (=3b)

interface, the HP in syntax would be an illusion.

- (7) a. (Gianni) *vincerà*. [Ita.]
 Gianni win-FUT-3sg
 ‘Gianni will win.’
- b. Chi credi [_C che [_{α<φ>} ~~chi~~ *vincerà*+T [_{β<v*>} ~~chi~~ *vincerà*]]]?
 who think-PRES-2sg that win-FUT-3sg
 ‘Who do you think that will win?’

Chomsky (2015) attributes the difference between Italian and English to the strength of T. Based on his claim, the facts above are accounted for as follows. English has a poor inflectional system and has a weak T, which cannot label itself. It needs an overt subject in its Spec to strengthen itself, as illustrated in (6a). When the *wh*-subject *who* moves to the Spec of the matrix C, its copy in [Spec,α] is invisible to LA. α cannot be labeled, which makes (6b) ungrammatical. On the other hand, Italian has quite a rich inflectional system and has a strong T, which can label itself. It does not need an overt subject in its Spec to strengthen itself, as illustrated in (7a). When the *wh*-subject *chi* moves out of [Spec,α], the strong T can label itself (, regardless of whether the complementizer *che* is overt, actually); see (7b).

The account of the EPP and the ECP above is dependent on many stipulations and assumptions. It is stipulated that Italian has a strong T but English has a weak T. It is assumed that a strong T can label itself, whereas a weak T cannot. Chomsky (2015) claims for (6c) that when the complementizer *that* disappears, T acts as a phase head, though, it is assumed, the embedded clause maintains the label of CP. That is, phasehood is inherited from the embedded C to T and activated in T. After the complementizer disappears, T acts as a phase head, and the complement of T, i.e. β<v*>, is transferred. The *wh*-subject in [Spec,α], which is now at the edge of the embedded ‘phase’, can be accessed by the syntactic operations carried out in the matrix phase and move up to sentence-initial position. This account of (6c) is ad hoc, obviously.⁸

Rizzi (2006, 2010, 2015) has argued that the properties of the CriP are accounted for in terms of *criterial Freezing*. A functional head and a sentential

⁸ From another perspective, it would appear that Chomsky’s account for (6c), where T acts as a phase head after the complementizer disappears, is simple and elegant (Anders Holmberg, p.c.). But the entire LA system works on the assumption that only C and v* are phase heads, whereas T and R are not. Chomsky’s account for (6c), in which T can exceptionally be a phasal head, is ad hoc.

element located in its Spec enter a structural, criterial relation; the interpretation of the latter is determined by the features of the functional head, Foc(us), Top(ic), etc. In (5a), the embedded C has Q. The *wh*-object *which dog*, which moves to its Spec, enters the Q-criterial relation with the embedded C and receives the interpretation as a *wh*-operator. In this system, a category raised into the Spec of a head, including the CriP, cannot move up further by definition.

Rizzi (2015) tries to give a unified account for the CriP and the issues related to that position in terms of LA, claiming that when XP and YP have a different label in configuration [$_{\alpha}$ XP YP], one of them can move up. His argument amounts to claiming that in configuration [$_{\alpha}$ XP YP] with XP being in the Spec of the head Y, XP can move up when it does not share any features with Y. In the HP (5a-b), *which dog* and the embedded C share a Q-feature; the former stops in [Spec, β] as in (5a) and cannot move up further as in (5b). In the EPP (6a), the subject shares φ -features with T (*Person* in his term); the subject is frozen in the Spec of that head. In the ECP (6b), the raised *wh*-subject *who* shares φ -features with the embedded T (/Person), which prevents the *wh*-subject from moving up further.⁹

Let us consider the properties of the CriP in detail. First, consider the HP and the properties of [Spec,(embedded)CP]. (8) (=1b) is the final representation of the derivation.¹⁰

(8) * $[_{\alpha} [_Q \text{ which dog}] \text{ do } [_{\beta < \varphi, \varphi >} \text{ you wonder } [_{\gamma < Q, Q >} \{ \text{ }_{\varphi} \text{ which dog} \}] \text{ C}_Q \text{ [John likes } \{ \text{ }_{\varphi} \text{ which dog} \}]]]]?$

After feature valuation occurs between the embedded verbal head *likes* and the *wh*-object *which dog*, the latter is assigned an Acc. It still has [uQ] and moves to [Spec, γ]. Since the verb *wonder* subcategorizes a *wh*-clause, the embedded C has

⁹ In Rizzi's account, it is actually not necessary to refer to LA, since what he refers to as a label corresponds to the feature shared by a head and the category in its Spec. He also makes several assumptions, e.g. the closeness between heads, the maximality condition on projections, etc, which can all be eliminated. To account for the Italian cases (7a-b), *pro*, an argument *pro* (7a) and an expletive *pro* (7b), is assumed. See his work for the details, and also Holmberg (2005) for a convincing argument against assuming *pro*. (6c) is accounted for by assuming that when C disappears, the entire CP system including CP and TP is omitted (Rizzi 2015:335,ft.16).

¹⁰ I leave aside the internal structure of *wh*-phrases. See Cable (2010) for a detailed discussion of that issue. Recall that the *wh*-object has moved to the Spec of *likes*(=R), which process is eliminated from the notation hereafter.

[Q]. The *wh*-object in [Spec, γ] and C_Q go on to feature valuation, and γ is labeled $\langle Q, Q \rangle$. In [Spec, γ], the *wh*-object completes the valuation of its own unvalued features. It is frozen there and cannot move up further to [Spec, α].

Next, let us consider the EPP and the properties of the subject position. (9) (=2) is the final representation of the derivation.

(9) $C_{[\alpha \langle \varphi, \varphi \rangle]}$ John [T [$\beta \langle v^* \rangle$] ~~John~~ [kisses(=R)+ v^*] [$\gamma \langle \varphi, \varphi \rangle$] Mary [kisses(=R)] [δ] ~~Mary~~]]]]]]]

John in [Spec, β] moves to [Spec, α] to strengthen T, which is not a phase head and weak. Feature valuation occurs between T and the raised *John*, and the latter is assigned a Nom; α is labeled $\langle \varphi, \varphi \rangle$. In [Spec, α], the subject *John* completes the valuation of its own [uCase] and stops there.

Finally, let us consider the ECP, which effect does not appear in Italian but appears in English in the unmarked case; see (6-7b). The ECP effect does not appear in English when the overt complementizer disappears; see (6c). Chomsky (2008) proposes the *parallel movement* analysis of *wh*-subjects: a *wh*-subject simultaneously moves from [Spec, v^*P] to [Spec,TP] on one hand and from [Spec, v^*P] to [Spec, CP] on the other in a parallel manner.¹¹ With the parallel movement analysis, the derivation of the ECP proceeds as illustrated in (10a-c), which are the final representations.

(10) a. [$\alpha \langle Q, Q \rangle$ *chi* C_Q [$\beta \langle \varphi, \varphi \rangle$ *credi* [$\gamma \langle C \rangle$ ~~*chi*~~ *che* [$\delta \langle \varphi, \varphi \rangle$ ~~*chi*~~ [*vincerà*+T [$\epsilon \langle v^* \rangle$ ~~*chi*~~ *vincerà*]]]]]]]]]?¹²

b. * [$\alpha \langle Q, Q \rangle$ *who* do+ C_Q [$\beta \langle \varphi, \varphi \rangle$ *you* think [$\gamma \langle C \rangle$ ~~*who*~~ *that* [$\delta \langle \varphi, \varphi \rangle$ ~~*who*~~ [T [$\epsilon \langle v^* \rangle$ ~~*who*~~ *read* the book]]]]]]]]]?

¹¹ Parallel movement as well as other kinds of Merge operations are denied by Chomsky et al. (2017), who claim that merge (or rather, the resulting structure built by merge) should be strictly binary. But the parallel movement analysis must be maintained to label the matrix clause of *wh*-subject interrogatives. Consider the following simple case:

i) * [$\alpha \langle C \rangle$ $C_{[Q, \varphi]}$ [$\beta \langle Q, Q \rangle$ *who* T_[Q, \varphi] [$\gamma \langle v^* \rangle$ ~~*who*~~ *left*]]]]]?

If Q were inherited from C to T in addition to φ -features as illustrated in (i), [uQ] of *who* would be valued by T (and its Case is also assigned a Nom). But since C does not have Q any longer, the matrix clause is labeled $\langle C \rangle$, i.e. as a declarative clause, which leads to a gibberish expression. Thus, Q must not be inherited from C to T. I thank Hisatsugu Kitahara (p.c.) for letting me notice this point. This discussion further concerns the issue on how to tighten feature inheritance, which I leave for future research.

¹² It has been traditionally claimed that a main verb moves to T in the Romance languages (e.g. Emonds 1978, Pollock 1989, Chomsky 1995), which is illustrated in (10a) but irrelevant here.

- c. [α <Q,Q> who do+C_Q [β < φ , φ > you think [γ <C> ~~who~~ \emptyset [δ < φ , φ > ~~who~~ [T [ϵ <v*> ~~who~~ read the book]]]]]]?]

The *wh*-subject *chi/who* moves from [Spec, ϵ] to [Spec, δ] on one hand, and from [Spec, ϵ] to [Spec, γ] on the other. Its [uCase] is assigned a Nom in [Spec, δ] by feature valuation with T, but it still has [uQ] in [Spec, γ]. Since the verb *credi/think* subcategorizes a *che/that*-clause, the embedded C does not have [Q] that can be shared by the *wh*-subject. Feature valuation does not occur between the embedded C and *chi/who* in [Spec, γ]. The *wh*-subject with [uQ] continues to move up to the matrix Spec. In [Spec, α], the *wh*-subject goes on to feature valuation with the matrix C_Q, and its [uQ] is valued. Completing the valuation of all of its own unvalued features, the *wh*-subject stops there.¹³

Note that a *wh*-subject should in principle be able to move across a declarative complementizer cross-linguistically, regardless of whether it is overt (10a-b) or not (10c): since feature valuation does not occur between the embedded C and the *wh*-subject in [Spec, γ], the latter, still having [uQ], continues to move up to [Spec, α], where its [uQ] is valued by the matrix C_Q.¹⁴ There is no difference between Italian (10a), in which an overt complementizer appears, and English (10b), in which an overt complementizer cannot appear. The disappearance of the ECP effect in English as illustrated in (10c) is thus not derived from any

¹³ δ is labeled < φ , φ >, after *chi/who* in [Spec, δ] and the embedded T go on to feature valuation. γ is labeled <C>, after the *wh*-subject moves out of [Spec, γ]. α is labeled <Q,Q>, after *who* in [Spec, α] goes on to feature valuation with the matrix C_Q.

¹⁴ The same argument applies to the v*P phase (Chomsky 2015:10,(3')):

- i) [α who do you [β v* [γ ~~who~~ expect [δ to win]]]]?]

Since the verbal root *expect* does not have Q, [uQ] of *who* is not valued in [Spec, γ]. The *wh*-phrase continues to move up. In the highest Spec, [Spec, α], its [uQ] is valued by the matrix C_Q and, *who* stops there.

Recall also the traditional claim (Huang 1982) that when a *wh*-phrase is extracted from the object position, the complementizer can appear overtly; see (ia). (ib) is the final representation of the derivation. In the same way as *wh*-subjects, the *wh*-object in [Spec, γ] does not go on to feature valuation with the embedded C that does not have Q. It still has [uQ] and continues to move up to [Spec, α]. Its [uQ] is valued in [Spec, α], and it stops there, completing the valuation of all of its own unvalued features. The derivation is licit, whether the complementizer appears overtly or not.

- i) a. Who do you think (that) John loves?
 b. [α <Q,Q> who do+C_Q [β < φ , φ > you think [γ <C> ~~who~~ (that) [δ < φ , φ > John [T [ϵ <v*> ~~John~~ loves ~~who~~]]]]]]?]

constraints in syntax.

Kandybowicz (2006) convincingly argues that the *that*-trace effect is derived from the phonological properties specific to English. According to Kandybowicz, the entire sentence is pronounced with one intonational phrase (iP), when the overt complementizer *that* does not appear; see (11a). When it appears as in (11b), an intermediate phrase (intP) occurs between the matrix verb and the complementizer.

- (11) a. [_{iP} Who do you think __ read the book]?
 b. * [_{iP} Who do you think [_{intP} that __ read the book]]?

Kandybowicz claims that the *that*-trace effect occurs when the overt complementizer *that* is adjacent to a trace within a prosodic phrase (i.e. an intP above) and at the boundary of that prosodic phrase. The *that*-trace effect in English is thus attributed to the phonological properties specific to English, which are outside the NS computation.

As we have seen so far, the CriP is the position in which a raised category *completes the valuation of all of its own unvalued features*. In other words, a category stops in the position where all of its unvalued features are valued. Much literature has preceded this claim in the pre-LA frameworks. Epstein (1992) is the first who claims that a *wh*-phrase cannot move out of the Spec of the embedded C that has [+wh]. Chomsky (2000, 2001) and Bošković (2007) argue that a category can move thanks to its own uninterpretable features.¹⁵ Bošković (2011) argues that after a category has its uninterpretable feature valued by a head, it cannot move out of the Spec of that head.¹⁶ Bošković (2008) claims that after an uninterpretable *wh*-feature is checked in the intermediate Spec, it cannot move up further.

¹⁵ Bošković (2007) claims that in the embedded clause of the C head that does not have [+wh], i.e. the C head that selects a *that*-clause, feature checking does not occur between the C head and a category raised to its Spec. Specifically in (i), feature checking does not occur between *that* and the copy of *what* raised to its Spec:

i) What do you think [_{CP} ~~what~~ that [_{TP} John bought ~~what~~]]?

¹⁶ In his argument, Bošković (2011) assumes both the distinction between interpretable and uninterpretable features (Chomsky 1995) and the distinction between valued and unvalued features (Chomsky 2000, 2001). The system that assumes both interpretability and valuation contained redundancy, as shown in the theoretical development into the current LA system, which assumes only valuation.

A category will continue to move, as long as it keeps some unvalued features. As illustrated in the ECP (10), the *wh*-subject goes on to feature valuation with the embedded T and is assigned a Nom. But it still has [uQ], which cannot be valued by the embedded C. Thus, it continues to move up to the highest Spec, where its [uQ] is valued by the matrix C_Q. This argument applies to all intermediate Spec positions. A *wh*-object, for instance, is assigned an Acc in the valuation procedure with a verbal head but still has [uQ]. It moves to [Spec,v*P] (Chomsky 2000), but its [uQ] cannot be valued by v*. It continues to move up to the highest Spec, where its [uQ] is valued by the matrix C_Q.¹⁷

Chomsky (2013:36,ft.36) poses the question why it is always a subject, not v*P, that moves out.¹⁸ The reason is that v*P does not have any unvalued features, contrary to the subject. The subject must move out of [Spec,v*P], since it moves to [Spec,TP] and its [uCase] is assigned a Nom there.¹⁹ On the other hand, v*P without any unvalued features does not move out in the unmarked case. Thus, it is not the case that any category can move out in an equally free manner: a moved category must have some unvalued feature(s) to enter feature valuation with a head in its raised position.²⁰

This unified account derives from the corollary of the LA derivational system: labeling results from feature valuation in all the cases except when LA takes a phase head as the label. In the HP (8), the *wh*-object in [Spec,γ] goes on to feature valuation with C_Q, and the embedded clause γ is labeled <Q,Q>. In the EPP (9), the subject moves to [Spec,α] and goes on to feature valuation with T. α is labeled <φ,φ>. In the ECP (10), the *wh*-subject in [Spec,γ] continues to move up, since its [uQ] cannot be valued by the embedded C of the verb *credi/think*, which subcategorizes a *che/that*-clause. After the *wh*-subject is raised to [Spec,α],

¹⁷ The literature has claimed, with various arguments, that there is no Agree in intermediate positions in successive cyclic movement. See, e.g. Bošković (2007, 2008) and Cecchetto and Donati (2015).

¹⁸ Chomsky would argue that both a subject and v*P could move out. The structure resulting from subject movement is interpreted at the interface, but the structure resulting from v*P movement would be filtered out at the interface.

¹⁹ It could be argued that after C merges to T and T inherits φ-features from C, feature valuation occurs between T and the subject in [Spec,v*P]; both unvalued φ-features in T and [uCase] of the subject are valued; the subject then moves to strengthen T. In this account, it is not clear why a subject must move to strengthen T after feature inheritance from C to T. I thank Anders Holmberg (p.c.) for letting me notice this possibility. See also footnote 5.

²⁰ The same argument should apply to adverbials, which do not seem to have unvalued features in the unmarked case and do not move out, which issue I turn to later.

it goes on to feature valuation with the matrix C_Q , and α is labeled $\langle Q, Q \rangle$.

In sum, the CriP is the position in which a raised category completes the valuation of all of its own unvalued features. A raised category must have some unvalued feature(s) which is valued by the head in its raised position; after it completes the valuation of all of its unvalued feature(s), it cannot move up further. The HP problem, the EPP, and the ECP (as well as the disappearance of that effect) are all fully accounted for in terms of feature valuation. This unified account derives from the corollary of the LA derivational system, in which labeling results from feature valuation.²¹

4. Movement from/into the Criterial Position

In Scandinavian OS (Holmberg 1986), weak pronominal objects can move across a sentential adverb like a negation (12a), contrary to full NP objects that do not move in the unmarked case (12b).

- (12) a. Jag målade den inte [_{VP} målade den]. [Swe.]
 I painted it not
 ‘I didn’t paint it.’
- b. Jag kysste inte [_{VP} kysste Marit].
 I kissed not Marit
 ‘I didn’t kiss Marit.’

Hosono (2016) claims that Scandinavian OS is the case in which a category that does not have unvalued features can move from the CriP. The derivational process of (12a-b) until when $\gamma_{\langle v^* \rangle}$ ($=v^*P$) is transferred based on the LA system is illustrated in (13a-b).²²

- (13) a. ... C [_{$\alpha_{\langle \varphi, \varphi \rangle}$} jag [T [_{$\beta$} inte [_{$\gamma_{\langle v^* \rangle}$} jag [målade(=R)+v* [_{$\delta_{\langle \varphi, \varphi \rangle}$} den [målade(=R) [_{$\epsilon$} den]]]]]]]]]]]]

²¹ The argument here does not mean that the derivation should be *Greed*-based (Chomsky 1995), in which system a category moves to check its own uninterpretable features. It is not the case as argued there that a category moves for its own needs; here, the category that has some unvalued features must move out, simply. Thanks to Anders Holmberg (p.c.) for letting me notice this point.

²² See Hosono (2016) for the detailed derivational procedure.

EPP, and the ECP (as well as the disappearance of its effect) are all fully accounted for in terms of feature valuation. This unified account derives from the corollary of the LA derivational system, in which labeling results from feature valuation. In this argument, Scandinavian OS and Icelandic SF are both exceptionally allowed to occur. That is, Scandinavian OS can occur, though object pronouns could not move from the CriP without any unvalued features; Icelandic SF can occur, though adverbs could not move into the CriP without any unvalued features.

The point is whether we should regard movement from/into the CriP as movement that would be allowed to occur with its legitimacy determined by filtering at the interfaces, or as movement that can exceptionally occur with its application constrained in some way. Contrary to the former, which will be taken by Chomsky, if the latter is in the right directions, constraints on movement should exist. Recall that it is argued by Chomsky that a SO needs to be interpreted at the interfaces. In the LA derivational system, the structure of a SO built in NS is directly interpreted at the semantic interface. Thus, no constraints on movement are imposed by the semantic interface.²⁵

According to Hosono (2016), constraints on movement are imposed by PHON. On the basis of Hosono (2013), who shows that *downstep* (cf. Gussenhoven 2004) occurs in simple tense forms in which the object pronoun moves, Hosono (2016) argues that movement of the object pronoun occurs when it is required by PHON: it is only when downstep needs to occur that the object pronoun can move. On the basis of Holmberg (2000), who claims that Icelandic SF occurs due to the requirement that something phonologically visible must occupy the Spec of T, Hosono (2016) also argues that Icelandic SF occurs due to the requirement by PHON. Hosono proposes that movement from/into the CriP in which a raised category does not have any unvalued features (which should be valued by a head in its raised position) can exceptionally occur in NS only when it is required by PHON. This constraint on movement from/into the CriP is formulated as follows:²⁶

²⁵ The association between the position that a category occupies in NS and the interpretation that it receives in the semantic component is not new: the *phase* framework since Chomsky (2000) was tied up with the *cartographic* system (Rizzi 1997, Cinque 1999), the latter of which exactly claimed that association.

²⁶ Anders Holmberg (p.c.) suggests that there could be an analogous constraint on movement required by the semantic component. The answer is no, at least within the current theoretical framework. As stated above, the structure of a SO built in NS is directly interpreted at the semantic interface. There is no room for constraints on movement to be imposed by the

- (16) The Constraint on movement from/into the Criterial Position:
 [XP_[uφ] ... [[XP_[±φ]]_{CriP} H ...]] and [[XP_[uφ]]_{CriP} H ... XP_[±φ] ...]
 are allowed in narrow syntax, iff movement is required by phonology.

On the basis of much empirical data that strongly connect syntactic structures with the intonational properties imposed on them, Richards (2016) suggests that the structure that conforms to the requirement by PHON may have already been formed in its syntactic derivation: syntactic derivation can go on so that resulting structures are fit for the requirement by PHON. For instance, in languages such as English in which *wh*-movement is obligatory, a *wh*-phrase moves to [Spec,CP] so that it can compose a phonological phrase with the C head; in languages that have a rich agreement system, agreement morphemes can be part of the prosody of a verbal head, which causes a verb to move to a higher head; and so forth.

Richards' (2016) claim is not incompatible with Chomsky (2013, 2015). Following Chomsky, both EM and IM are free; among SOs constructed in NS, those which do not have a well-formed prosodic structure would be filtered out in PHON. In the same way, on the basis of Richards, NS operations would try to construct SOs that conform to the appropriate phonological properties that they would have in PHON; those which fail in having a well-formed prosodic structure would be filtered out in PHON.

As long as there is evidence that far more constraints on movement are imposed by PHON than have been considered so far, however, IM may not be free, contra Chomsky (2013, 2015). As has been seen so far, the category that does not have unvalued feature(s) cannot move from/into the CriP in the unmarked case. Movement from/into the CriP is exceptionally allowed to occur in NS only when it is required by PHON. Movement in NS is thus constrained by the requirement by PHON.

This argument further indicates that the derivational mechanism will be *crash-proof* (Frampton and Gutmann 2002). Frampton and Gutmann (2002) claim that the derivational mechanism should be constrained within its own system so that only well-formed structures are produced. As has been argued here, the category that does not have unvalued feature(s) cannot move from/into the CriP in the unmarked case. Such movement is allowed to occur in NS only when it is required by PHON, i.e, only when it constructs a SO with a well-formed prosodic structure. The derivational mechanism will then produce only well-formed

from the corollary of the LA derivational system that labeling results from feature valuation. In Scandinavian OS and Icelandic SF, a category that does not have unvalued features can move from/into the CriP (Hosono 2016). Following Chomsky (2013, 2015), who claims that both EM and IM are free, movement from/into the CriP would be allowed to occur with its legitimacy determined by filtering at the interfaces. If such movement is considered to exceptionally occur in NS, constraints on movement should exist. The argument that far more constraints on movement are imposed by PHON than have been considered so far (Hosono 2016, Richards 2016) indicates not only that IM may not be free, but also that NS will be *crash-proof* (Frampton and Gutmann 2002): the derivational mechanism should produce only well-formed structures that conform to the requirement by PHON, with no filters assumed.

References

- Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589-644.
- Bošković, Željko. 2008. On the operator freezing effect. *Natural Language & Linguistic Theory* 26:249-287.
- Bošković, Željko. 2011. On Unvalued Uninterpretable Features. *The Proceedings of NELS 39*, ed. by Suzi Lima, Kevin Mullin, and Brian Smith, 109-120. GLSA.
- Broekhuis, Hans. 2008. *Derivations and Evaluations: Object Shift in the Germanic Languages*. Berlin: Mouton de Gruyter.
- Bruce, Gösta. 1977. *Swedish Word Accents in Sentence Perspective*. Travaux de L'Institut de Linguistique de Lund XII. CWK GLEERUP. PhD dissertation, Lund University.
- Cable, Seth. 2010. *The Grammar of Q: Q-Particles, Wh-Movement, and Pied-Piping*. Oxford: Oxford University Press.
- Cecchetto, Carlos, and Caterina Donati. (Re)labeling. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. New York: Mouton de Gruyter.
- Chomsky, Noam. 1986. *Knowledge of Language*. Westport, CT: Praeger.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by Step*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-156. Cambridge, MA: MIT Press.

- Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.
- Chomsky, Noam. 2008. On Phases. In *Foundational Issues in Linguistic Inquiry: Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133-166. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33-49.
- Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, Strategies and Beyond: Studies in Honour of Adraian Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3-16. Amsterdam: John Benjamins.
- Chomsky, Noam. 2016. Ouzzles about Phases. ms., MIT.
- Chomsky, Noam, Ángel J. Gallego, and Dennis Ott. 2017. Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. Ms., MIT/Universitat Autònoma de Barcelona/University of Ottawa.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Collins, Chris. 2002. Eliminating Labels. In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel D. Epstein and T. Daniel Seely, 42-64. Oxford: Blackwell.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Diesing, Molly. 1997. Yiddish VP Order and the Typology of Object Movement in Germanic. *Natural Language and Linguistic Theory* 15:369-427.
- Emonds, Joseph. 1978. The verbal vcomplex V'-V in French. *Linguistic Inquiry* 9:151-175.
- Epstein, Samuel David. 1992. Derivational Constraints on \bar{A} -Chain Formation. *Linguistic Inquiry* 23:235-259.
- Epstein, Samuel D., Hisatsugu Kitahara and T. Daniel Seely. 2015. *Explorations in Maximizing Syntactic Minimization*. New York: Routledge.
- Erteschik-Shir, Nomi. 2005. Sound Patterns of Syntax: Object Shift. *Theoretical Linguistics* 31:47-93.
- Fox, Danny and David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1-45.

- Frampton, John, and Sam Gutmann. 2002. Crash-Proof Syntax. In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel D. Epstein and T. Daniel Seely, 90-105. Oxford: Blackwell.
- Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Holmberg, Anders. 1986. *Word Order and Syntactic Features in the Scandinavian Languages and English*. PhD dissertation, University of Stockholm.
- Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53:1-39.
- Holmberg, Anders. 2000. Scandinavian Stylistic Fronting: How any category can become an expletive. *Linguistic Inquiry* 31:445-483.
- Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36:533-564.
- Holmberg, Anders, and Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford: Oxford University Press.
- Hosono, Mayumi. 2013. *Object Shift in the Scandinavian Languages: Syntax, Information Structure, and Intonation*. PhD dissertation, Leiden University.
- Hosono, Mayumi. 2016. Exceptional Movement from/into the Criterial Position. *Working Papers in Scandinavian Syntax* 97:23-39.
- Huang, Jim C.-T. 1982. Logical relations in Chinese and the theory of grammar. PhD dissertation, MIT.
- Jónsson, Jóhannes Gíslí. 1991. Stylistic fronting in Icelandic. *Working Papers in Scandinavian Syntax* 48:1-43. Department of Scandinavian Languages, Lund University.
- Josefsson, Gunlög. 2003. Four Myths about Object Shift in Swedish – and the Truth.... In Lars-Olof Delsing, Cecilia Falk, Gunlög Josefsson, and Halldór Á. Sigurðsson, eds., *Grammar in Focus: Festschrift for Christer Platzack 18 November 2003 Volume 1*, pp.199-207. Department of Scandinavian Languages, Lund University.
- Josefsson, Gunlög. 2010. Object shift and optionality. An intricate interplay between syntax, prosody and information structure. *Working Papers in Scandinavian Syntax* 86:1-24.
- Kandybowicz, Jason. 2006. Comp-trace Effects Explained Away. *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. by Donald Baumer, David Montero and Michael Scanlon, 220-228. Somerville, MA: Cascadilla Press.

- Mikkelsen, Line. 2011. On Prosody and Focus in Object Shift. *Syntax* 14:230-264.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Richards, Norvin. 2001. *Movement in Language: Interactions and Architectures*. Oxford: Oxford University Press.
- Richards, Norvin. 2016. *Contiguity Theory*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar*, ed. by Liliane Haegeman, 281-338. Dordrecht: Kluwer Academic Publishers.
- Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-Movement: Moving On*, ed. by Lisa Lai-Shen Cheng and Norbert Corver, 97-133. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 2010. On some properties of criterial freezing. In *The Complementizer Phase: Subject and Operators*, ed. by E. Phoevos Panagiotidis, 17-32. Oxford: Oxford University Press.
- Rizzi, Luigi. 2015. Cartography, criteria, and labeling. In *Beyond Functional Sequence: The Cartography of Syntactic Structures, Volume 10*, ed. by Uri Shlonsky, 314-338. Oxford: Oxford University Press.
- Seely, T. Daniel. 2006. Merge, derivational C-command and subcategorization in a label-free Syntax. In *Minimalist Essays*, ed. by Cedric Boeckx, 182-218. Amsterdam: John Benjamins.
- Sells, Peter. 2001. *Structure, Alignment and Optimality in Swedish*. Stanford, CA: CSLI Publications.
- Vikner, Sten. 2001. The Interpretation of Object Shift and Optimality Theory. In Gereon Müller and Wolfgang Sternefeld, eds., *Competition in Syntax*, pp.321-340. Berlin: Mouton de Gruyter.

mayumi.hosono@keio.jp