# Be Careful Where You Transfer: Deriving Consequences from the Nature of Transfer and its Interaction with Labeling toward Establishing a Parametric Syntax on Scrambling in the Free-Merge Model of Grammatical Architecture*

Uchishiba, Shin'ya
Independent Researcher

**Abstract**

In this paper, I investigate the nature of Transfer and its interaction with Labeling, a syntactic operation concerning the endocentricity of phrase structure, and thereby I aim to untangle the following problems which the recent minimalist theorizing has been facing: (i) a query concerning the tension between the theoretical expectation inspired by Free-Merge, under which movement as a variant of Merge comes free, and the empirical fact, which suggests that movement is not so free; (ii) the question of why languages vary with respect to phrasal permutation or scrambling; and (iii) the issue called Problem of Projection, which has remained notoriously burdensome to Labeling. I argue that the solution of (iii) adumbrates how to solve the other two. As for (iii), it is demonstrated that the effect of Transfer paves a way to solve this issue. It is further shown that Transfer interacts with Labeling to restrict the outcome of movement (but not movement per se). It is then argued that a considerable attention has to be paid to where Transfer applies, so as to restrict the outcome of movement. It is also argued that this restriction is subject to the functional vs. lexical distinction of syntactic categories. To verify the approach to be presented in this paper, English and Japanese are examined as a case study. The ultimate enterprise I aim at in this paper, then, is to lay out a parametric approach to cross-linguistic variation regarding scrambling based on the Free-Merge model of grammatical architecture.

*Keywords: Transfer, Labeling, Free-Merge, Problem of Projection, phrasal permutation or scrambling, dichotomy of syntactic categories, duality of semantic interpretation, English vs. Japanese*

## 1. Introduction

In the recent minimalist theorizing (Chomsky 2000, et seq.), Transfer (or Spell-Out as its precursor) is modelled as a mapping operation. Standardly, this operation is said to take on the following properties: (i) it sends syntactic objects (henceforth, abbreviated as SO(s)) to the interpretive components interfaced with the narrow syntactic computation, thereby making them "forgotten" or inaccessible to further syntactic operations; (ii) as a consequence from (i), it reduces the burden of computational memory; (iii) for attaining the purpose of (ii), it chunks the derivation down into as small as possible units (often referred to as *phases* [Chomsky 2000, et seq.]); (iv) so, it applies several times in a cyclic fashion. These properties have been revealed, developed, and further sophisticated by a series of prestigious work by Uriagereka (1999, et seq.) and his followers, which has substantiated the so-called Multiple Spell-Out model of grammatical architecture. It is by mounting this model into the current minimalist framework that a plenty of fruitful successes has been achieved both on the theoretical

* Portions of this paper were presented at the 21st Annual International Conference of the Japanese Society for Language Sciences and the 37th Conference of the English Linguistic Society of Japan. I thank the audiences at these venues for their comments and questions. Comments for the present paper are welcome at my email address: j.linguistar@gmail.com.

and the empirical grounds.

In this paper, I will throw light on the nature of Transfer from different perspectives. More specifically, I would like to propose that we have to be more attentive as to (i) where Transfer applies and (ii) how this operation should interact with other syntactic operations. With respect (ii), as an operation with which Transfer interacts, I will concentrate on Labeling, which plays a crucial role on the alignment of endocentricity of syntactic structure. I will then argue that these two facets of Transfer (i.e., the site of Transfer and its interplay with Labeling) provide a window into how to cope with the following three problems which I suspect have remained to be solved in the recent minimalist theorizing.

I. Why doesn't movement look free, although it comes free in principle? (Tension problem)
II. Why can movement permute phrases in some languages but not in others? (Variation problem)
III. How should Labeling be able to detect the head of an SO composed of two phrases? (Problem of Projection)

In what follows, I describe the theoretical background behind these three problems and thereby illuminate why they are issues to be solved. Prima facie, the third problem, a.k.a. Problem of Projection, appears to be independent from the other two. However, as will be clear later, the former problem bears a close connection to the latter two. Specifically, the modus operandi to be put forth for solving Problem of Projection provides a clue to puzzle out both the tension problem and the variation problem. Ultimately, we will untangle these three problems by investigating the nature of Transfer and its interplay with Labeling, and thereby end up with laying out a parametric approach to cross-linguistic variation regarding phrasal permutation or scrambling within the current minimalist framework of generative grammar.

**1.1. Tension problem**

In his recent paper (Chomsky 2020), Chomsky has established that the concatenating operation Merge is unrestrained and hence comes free. Namely, the application of this operation does not hinge on any triggering force like feature checking and hence is considered non-teleological. Merge has two variants: one is external, and the other is internal. When Merge is external, it simply concatenates two SOs into a bigger one: in this case, both two may be picked out from the lexicon "external" to the derivational workspace; or one of them may be so, while the other may be already present in the derivational workspace; or each of them may be independently present in the derivational workspace. When Merge is internal, it picks out an SO from another one already built in the derivational workspace, then concatenating these two into a bigger one. This variant of Merge, Internal-Merge in short, thus induces the same effect as the traditional manipulation of movement, i.e., displacement. Given that Merge is free, then it follows naturally that its internal variant is likewise free. That is, movement operates freely. However, looking across languages over the world, we immediately find that they teach us the opposite. For example, English, one of the well-studied languages, does not allow for movement of an object phrase over a subject phrase in a normal sentence.

(1) a. John praised Mary.
b. *Mary John praised. (*Mary* is neither a focus nor a topic.)

A problem, then, is how, if Merge is free regardless of its variant, the grammatical model of language should explain the empirical fact that phrases cannot move so freely.

### 1.2. Variation problem

A further complication comes from the fact that some languages are allowed to arrange word-order relatively freely. Representative among them is Japanese, another one of the well-studied languages. In this language, an object phrase may either follow or precede a subject phrase even in a normal sentence, as exemplified below:

(2) a. John-ga Mary-o home-ta.
John-NOM Mary-ACC praise-PAST
'John praised Mary.'
b. Mary-o John-ga home-ta.

The phenomenon attested in (2) has been referred to as Scrambling in the traditional literature (Fukui 1986, et seq.; Hoji 1985; Kuroda 1988; Ross 1967; Saito 1985; among many others). The issue of why scrambling is feasible in some languages but not in others has been a leitmotif since the earliest era of generative grammar (Hale 1980, 1983; Chomsky 1981). The simplest way to approach this issue might be to assume that "free word-order" languages like Japanese are endowed with a certain feature triggering scrambling, while "rigid word-order" languages like English are lack of any such feature. In fact, this or a similar approach has been occasionally adopted in the literature (Grewendorf and Sabel 1999; Kawamura 2004; Miyagawa 2001, 2003; inter alia). However, postulating such feature as a triggering force for movement does not make any sense under the present-day model of generative grammar. As this model presupposes that all languages employ movement freely alike, the issue at stake resurfaces in the form of how to restrict word-order possibilities without restricting movement per se.

### 1.3. Problem of Projection

Labeling is a syntactic operation innovated in the current model of generative grammar (Chomsky 2013, 2015; see also Biskup 2017; Collins 2002; Epstein, et al. 2014; Narita 2012, 2014; Rizzi 2015, 2016; Saito 2016, 2018; Takita, et al. 2016; Vercauteren 2017; among others). Its role is to determine the head of an SO, centered on the endocentric projection of syntactic structure. If the syntactic computation cannot fix the head of an SO in an unambiguous way, this means that it fails to assign any label to that SO. Unlabeled SOs are said to be illegible in the interpretive components, hence preventing the convergence of derivation therein. Namely, such SOs crash the the derivation. In order for the derivation to converge, SOs have to have their labels determined at least until they are sent off to the interfaces. In what follows, however, I will not refer to the timing of Labeling in particular.

Now, with this in mind, let us consider the following structure:

(3) $[_{\alpha}$ XP, $[_{YP}$ ... <XP> ...]]

(3) represents the structure of an SO created via Internal-Merge: XP moves from the inside of YP, and then merges with YP. The second merging operation then delivers α as its product. (N.B.: the usage of the familiar notations

like XP and YP is only for expository purposes. The tenet flowing coherently throughout this paper adheres to the thesis of Bare Phrase Structure [Chomsky 1995, et seq.], which eliminates any such representational level from syntactic entities in the faculty of language.) Here, I speak of it that XP is at the edge of YP and YP is at the edge of XP. This is because "Merge is invariably "to the edge"" (Chomsky 2008: 138). In addition, I use the angle-bracket notation to indicate copies: in (3), the copy of XP left behind inside YP is represented between angle-brackets. XP's movement in (3) comes free under the present model of generative grammar. Therefore, we cannot maintain that this movement should not take place for the reason why there is no triggering force for it. However, if the outcome of the movement, i.e., α in (3), fails to have its head or label determined in an unambiguous way, it cannot stand legible on the interfaces. Namely, the derivation involving this outcome results in crashing on the interfaces. Thus, owing to Labeling, we can restrict the outcome of movement, but not movement per se (cf. Vercauteren 2017). The tension problem is then resolved: movement is unrestrained in and of itself, but the outcome of movement is restrained under Labeling. This manifests an important aspect of the approach to be pursued and developed in this paper.

Factually, however, a syntactic configuration like (3) poses a bothersome problem on Labeling. Given that Labeling obeys minimality or closeness much like other syntactic operations, the syntactic computation cannot determine the head of α in (3) in an unambiguous way. This is because the heads of XP and of YP are both equally close to α in that neither of them c-commands the other. Thus, when an SO consists of two phrases, its label cannot be unambiguously fixed. On top of this, a syntactic configuration like (3) is pervasively attested over the process of sentence-structure building: for example, a subject phrase merges with a predicative phrase, a wh-interrogative phrase merges with a sentential phrase, and so on and forth. This problem is called Problem of Projection (abbreviated as PoP), already acknowledged by several generativists including Chomsky (see Chomsky 2013, 2015 and references cited above). Those researchers have been struggling with this problem, but they have not yet reached any persuasive solution to it.[1]

### 1.4. The theoretical approach in a nutshell

Therefore, before we tackle the tension problem and the variation problem, we have to solve PoP somehow. In this paper, I would like to offer a solution based on the effect of Transfer. Here, please remember the properties of Transfer, mentioned at the outset of this paper. Among them, the first property is crucial to our purpose: Transfer makes an SO to which it applies inaccessible to further syntactic operations. As Labeling is a syntactic operation, it follows that this operation cannot access the inside of any SO that Transfer has already applied to. With this in mind, let us return to (3), and then suppose that Transfer applies to XP at the edge of YP. This operation then renders XP inaccessible to further syntactic operations. As a result, Labeling for α becomes unable to access the inside of XP. As it is only the head of YP that remains accessible, Labeling for α accesses this head and then unambiguously determines it as the head of α. That is, α is unambiguously labeled as Y(P). If Transfer applies to YP at the edge of XP in (3), the head of XP is determined as the label of α (however, we will find this derivational option to result in crashing the derivation, for the reason which I will describe later). Accordingly,

[1] Whether all SOs should have unambiguous labels is a pending issue and waits for empirical investigation (Chomsky et al. 2019). In fact, Fukui (2011), Bauke & Roeper (2017), and Blümel (2017) suggest the possibility that unlabeled SOs may be real or can be justified. In this paper, I do not intend to argue against this possibility. The purpose thereof is not do so, but rather to reveal what empirical and theoretical consequences can be derived by investigating the nature of Transfer and its interaction with Labeling.

an SO, when composed of two phrases, receives an unambiguous label simply by transferring either of them. Labeling thus avoids PoP with the aid of Transfer. This is a consequence from the interplay of Transfer with Labeling.

Certainly, the effect of Transfer is helpful to evaporate PoP. However, this Transfer-based solution of PoP alone is not sufficient to grapple with the tension problem and the variation problem. In order to do so, we have to grant the following (as a first approximation): (i) the site to which Transfer can apply is restricted, and (ii) this restriction is subject to variation. Let us begin with (i). As stated above, movement comes free, but its outcome is restrained under Labeling. That said, the outcome poses PoP on Labeling, since it consists of two phrases and hence cannot receive an unambiguous label as it stands. This problem is circumvented by transferring the moved phrase at the edge of another phase in which it originally resided. Now, if Transfer is unrestrained (much like Merge), it can apply to any SO at any site. It follows then that this operation can apply to moved phrases at any site. If so, Labeling cannot restrict the outcome of movement as we expected before, and hence the tension problem remains to be resolved. Namely, movement can halt anywhere, because it always produces an unambiguously labelable SO through the mediation of the "inaccessibility" effect of Transfer. In order to overcome this state of affairs, we have to restrict the site to which Transfer can apply. In (3) for instance, if Transfer fails to apply to XP at the edge of YP for some reason, Labeling becomes unable to detect the head of α in an unambiguous way. α is then left unlabelable, hence illegible in the interpretive components. In this case, XP's movement in (3) results in breading an illegible SO which yields the ungrammaticality (but this movement is not ungrammatical in and of itself). Thus, in terms of the Transfer-based solution of PoP, restricting the site of Transfer is tantamount to restricting the outcome of movement under Labeling. And then, varying this restriction on the site of Transfer in some way leads to varying the outcome of movement accordingly.

However, I surmise that the idea that some restriction should be imposed on Transfer should be rejected for a conceptual reason. The reason that I have in mind is that such idea is at odds with it that Merge is free. Merge and Transfer are different operations, of course. That said, it is hardly possible to explain why the former is free, on one hand, and why the latter is not, on the other hand. In other words, if Transfer is restrained as stated above, we are faced with a conceptual issue of why only Merge is unrestrained. This conceptual issue, if a genuine one, obliges us to reconsider how to resolve the tension problem and the variation problem in terms of the Transfer-based solution of PoP.

I would like then to suggest that the outcome of Transfer, but not Transfer per se, should be subject to restriction. I think that, in order to restrict the outcome of Transfer, we have to be attentive to another effect that Transfer has. The effect is that an SO, when transferred, is interpreted at the site where it is transferred. Namely, Transfer makes an SO to which it applies interpreted at the site where it operates. I consider this effect a corollary of the mapping function of Transfer. To see why, consider the following toy structure:

(4) $[_{\alpha}$ XP, YP]

In (4), the two phrases XP and YP merges with each other, forming α. α is unlabelable as it is, yielding PoP. If XP for instance is transferred, α turns to be unambiguously labelable in the manner described above. When transferred,

XP is mapped onto the interpretive interfaces. More precisely, this mapping operation sends to the interpretive components the instruction that XP should be interpret at the site where it is transferred: in the perceptual-sensorimotor system, XP is phonologically (or orthographically) realized at that site, and, in the conceptual-intentional system, it is semantically interpreted at the same site.[2,3] Focusing on the semantic side of this mapping effect, we can say that, for a given SO, the site where it is transferred corresponds to the site where it is semantically interpreted. I submit this as an axiom. In what follows, I will be exclusive to the semantic side of the mapping effect of Transfer, putting aside its phonological effect, otherwise necessary.

Let us now suppose that the edge of YP in (4) is so "thorny" a position that XP is hard to interpret in an appropriate way. Then, α will become interpretively defective in the semantic component, albeit being labeled in an unambiguous way. This is simply because XP cannot stand legible as a part of α. Behind this reasoning is Principle of Compositionality (Chomsky 2007; see also Boeckx 2008; Davidson 1967a, 1967b; Pietroski 2002;), which dictates that every complex SO (like α in (4)) is compositionally interpreted: namely, the interpretation of a complex SO depends not only on how each of the constituents of which it consists (i.e., its daughters) is interpreted, but also on how these constituents are interpretively related. In (4), therefore, when XP as one of α's daughters cannot be appropriately interpreted, its mother constituent α is regarded as interpretively defective, even though labelable.

As for how XP in (4) should be interpreted as a part of α after it has been transferred, I suggest that this matter depends on YP but not XP per se. More specifically, I propose that the category-type of YP designates the mode of interpreting XP as a part of α. As is well known, syntactic categories are bifurcated into two: lexical vs. functional (Fukui 1988, 1995; Grimshaw 2000; Riemsdijk 1998; Zushi 2003, 2005). I accord to this bifurcation of syntactic categories in this paper. I suggest, then, that if YP is lexical, XP should be lexically interpreted, and if YP is functional, XP should be functionally interpreted. Accordingly, when Transfer applies to XP, its outcome is restricted so as to be interpreted in the mode which the category-type of YP designates. This is the way I will propose in this paper to restrict the outcome of Transfer. In a nutshell, how to restrict the outcome of Transfer is an interpretive matter, depending on the syntactic context where the outcome in question arises. This interpretive constraint on the outcome of Transfer is applicable to the variation problem, because the outcome in question varies with the category-type of the SO at whose edge it arises.

### 1.5. Organization of this paper

Having set up the problems to be dealt with in this paper and having outlined the theoretical approach to them therein, I now lay out the organization of the rest thereof. In Section 2, I give an overview of Chomsky's (2013, 2015) solution of PoP, and then point out its theoretical drawbacks. Afterward, I show that the Transfer-based

---

[2] Chomsky (2005, 2014) posits Externalization as a variant of Transfer, specializing it in mapping to the interface with the perceptual-sensorimotor system. For the reason described just below in the text, I will not pay attention to this specialized operation mostly in this paper.

[3] This might infer, if taken at face value, that, as for each SO, the site where it is phonologically realized coincides with the site where it is semantically interpreted, and vice versa. This is not the case, of course, when it comes to so-called reconstruction and covert movement. In the former case, the moved phrase is taken to be semantically interpreted at its base position, but not at the site where it is phonologically realized. In the latter case, it is occasionally thought that the moved phrase is phonologically realized at its base position but semantically interpreted at its halting site. How to handle these cases is out of the scope of this paper.

solution of PoP in this paper is immune from those drawbacks. In Section 3, I elaborate the theoretical approach in more details. In particular, I make a more specific statement about what "being lexically/functionally interpreted" means, and why, given that a certain SO is transferred at the edge of another one, the former's mode of semantic interpretation should be designated by the latter's category-type, but not by the former's. Section 4 is devoted to solving the variation problem by implementing the approach elaborated in Section 3. Therein, I take English as representative among "rigid word-order" languages, and Japanese as representative among "free word-order" ones. In Section 5, the predictions of the approach in this paper are examined and eventually shown to be borne out. To this end, it will be demonstrated that the present approach is well-designed on the empirical ground.

**2. How to solve PoP**

As a first approximation, let us follow Rizzi (2016), assuming the labeling algorithm defined in (5) (cf. also Vercauteren 2017):

(5) $H_1$ is the closest head to α iff,
   I. α contains $H_1$, and
   II. there is no $H_2$ such that (i) α contains $H_2$, and (ii) $H_2$ c-commands $H_1$.

(Rizzi 2016: 18)

Suppose here that Labeling applies to an SO consisting of a head, say X, and a phrase, say YP, viz. $[_{\alpha}$ X, YP]. In this case, the labeling algorithm in (5) trivially determines X as the head or the label of α: α contains X and there is no other head such that it c-commands X within α. However, as far as Labeling obeys (5), PoP inevitably arises in (4) above, viz. $[_{\alpha}$ XP, YP]. The labeling algorithm in (5) cannot unambiguously fix the label of α in this toy structure, because the heads of XP and of YP are equally close to α in that neither of them c-commands each other.

**2.1. Chomsky's solution(s): movement and feature sharing**

Chomsky offers two solutions to PoP. One is based on movement, and the other is based on feature sharing. To see how he attempts to solve PoP, consider the derivation shown below:

(6) i. $[_{\alpha}$ XP, YP] (= (4))
   ii. $[_{\beta}$ XP, $[_{ZP}$ Z, $[_{\alpha}$ <XP>, YP]]]

The derivation proceeds from (6i) to (6ii) via the following two steps: α undergoes External-Merge with the head Z, and then the resulting product ZP undergoes Internal-Merge with XP, thereby β created. By the second merging operation, XP moves to the edge of ZP, leaving its copy at the edge of YP. Chomsky suggests that any copy left behind by movement does not partake in the labeling algorithm. This follows if we define the notion of containment as follows: α contains β iff α dominates all the occurrences of β (Rizzi 2016; Vercauteren 2017). In (6ii), there are two occurrences of XP: one is at the edge of ZP, and the other is at the edge of YP. Each of them is the same SO as XP, but they are discrete occurrences of XP. In terms of the notion of containment thus defined, α in (6) ceases to contain XP, because it no longer dominates all the occurrences of XP. As a result, it follows that α contains YP (and all that YP contains) alone. Hence, the head of YP is unambiguously chosen to fix the label of α. Namely, α is labeled as Y(P). Movement thus circumvents PoP.

This movement-based solution of PoP seems to be cogent. However, it is not sufficient by itself to evaporate PoP completely. For it yields another unneglectable problem. The problem is that movement breeds another unlabelable SO. In this respect, consider (6ii) again. Therein, although XP's movement renders α labelable, it breads β, which is not unambiguously labelable as it is due to consisting of XP and ZP. In terms of the movement-based solution of PoP, the unambiguous determination of β's label requires further movement from β: i.e., either XP or ZP should move away. And, this further movement breeds another unlabelable SO, the unambiguous determination of whose label requires further movement out of it. Thereafter, the same cycle is repeated ad infinitum. Thus, the movement-based solution of PoP coerces the infinite application of movement. Chomsky himself acknowledges this problem, calling it "Halting Problem".

Chomsky's second solution directly pertains to this problem. He claims that the labeling algorithm can make reference to features that two Merge-members share. More precisely, given that two SOs merge with each other and that they share features that match, the labeling algorithm then bluffs these matching features into the label of the SO consisting of those two. In (6ii) for instance, suppose that XP and ZP shares certain features, called Fs. Then, the labeling algorithm bluffs those features into the label of β as {F, F}. In this way, Chomsky exploits the mechanics of feature sharing to solve PoP (see also Vercauteren 2017). In addition, this solution based on feature sharing leads to circumvent Halting Problem. In (6ii), as β receives a label in terms of feature sharing, no further movement is required from β. Accordingly, movement can halt only when the moved phrase shares certain features with another phrase with which it merges. However, we should not pretermit that this solution is not coherent with the primary role of Labeling, i.e., it does not provide any way to detect the head of an SO. Taking β in (6ii) for instance, bluffing its label as {F, F} is quite different from determining which head, X or Z, fixes its label.

**2.2. The Transfer-based solution of PoP and its virtue**

In this paper, as a "third" solution of PoP, I propose to take the best advantage of the effect of Transfer. The present model of generative grammar devises Transfer as a mapping operation to the interpretive interfaces. It hypothesizes that, once a given SO has undergone Transfer, it is no longer accessible to further syntactic computation. What I would like to propose is to mount this "inaccessibility" effect of Transfer into the labeling algorithm. Namely, I would like to reformulate (5) above into the following:

(7) $H_1$ is the closest head to α iff,
   I. $H_1$ is accessible (i.e., it is not in any transferred domain),
   II. α contains $H_1$, and
   III. there is no $H_2$ such that (i) $H_2$ is accessible, (ii) α contains $H_2$, and (iii) $H_2$ c-commands $H_1$.

Given this revised version of the labeling algorithm, PoP evaporates straightforwardly. To see how, consider (6ii) once again. Therein, XP's movement yields PoP, since it produces β, which is unlabelable in terms of consisting of XP and ZP. Suppose here that Transfer applies to XP at the edge of ZP. This operation then renders XP inaccessible to further syntactic computation. Therefore, when Labeling applies to β, it cannot access the head of XP. Only the heads that Labeling can access within β are inside ZP, since the latter remains accessible to syntactic

computation. In addition, among those heads, ZP's one counts as the closest to β, since it c-commands all the others. Consequently, ZP's head is unambiguously determined as the label of β. Thus, by virtue of the effect of Transfer, Labeling circumvents PoP.

This Transfer-based solution of PoP is immune from the drawbacks which Chomsky's solution(s) suffers from. At first, Halting Problem does not arise in the present solution, because movement halts just when the moved phrase is transferred. In (6ii), when XP is transferred at the edge of ZP, it halts there. Secondly, the present solution is true to the primary role of Labeling. In (6ii), in terms of this solution, ZP's head and only this head is detected as the label of β, as demonstrated above. I consider this trueness of the present solution to the primary role of Labeling a theoretical motivation to advocate it rather than Chomsky's one. However, I must haste to add the caveat that I do not intend to turn down Chomsky's first solution of PoP. As mentioned above, I advocate this movement-based solution as cogent in this paper, since it maintains the primary role of Labeling as well. In (6ii), when Labeling applies to α after XP has moved away, it can determine YP's head and only this head as the label of α. The sole problem against the movement-based solution of PoP is Halting Problem. That is, this solution alone is not sufficient to solve PoP completely. Eventually, I reject Chomsky's second solution based on feature sharing, for the reason why it is aberrant in that it does not carry out the primary role of Labeling.

In (6ii), one may wonder whether or not Transfer applies to ZP vis-a-vis (the higher occurrence of) XP. There seems to be no reason to rule out this derivational option per se, given that Transfer is free much like Merge. If ZP is transferred, the resulting syntactic configuration is represented as follows:

(7) $[_{XP}$ XP, //$[_{ZP}$ Z, $[_{YP}$ <XP>, YP]]//]

In (7), I use the double-slash notation to represent SOs which have already been transferred: ZP in (7), being transferred, is represented between double-slashes. Therein, as shown, β in (6ii) inherits projection from the higher occurrence of XP at the edge of ZP. This inheritance owes to the accessibility of XP's head and the inaccessibility of ZP's one. The latter is a consequence of Transfer. Namely, in the case under consideration, the head of the higher occurrence of XP counts as the nearest to β among the accessible heads within it. At this point, we should note that the top-most XP node dominate the lower occurrence of XP at the edge of YP. I submit that a syntactic configuration like (7) is defective in that it fails to form a discontinuous object of the moved constituent. The term "discontinuous object" is approximately synonymous to the more familiar term "chain". Suppose that the formation of a discontinuous object requires a lower occurrence of the moved constituent to be c-commanded by its higher occurrence (much as the formation of a chain does). Then, the two occurrences of XP in (7) fail to constitute a discontinuous object of XP, since the higher one dominates and hence does not c-commands the lower one. Given that the failure to form a discontinuous object of the moved constituent means that each of its occurrences cannot be properly interpreted because of losing its own identity (metaphorically speaking, each of them has no knowing of "Who I am"), a syntactic configuration like (7) should be eliminated as an illegible object. A lesson we can elicit from this reasoning is that, when an SO moves to and halts at the edge of another one, the former but not the latter should be transferred. Otherwise, an illegible SO like (7) would inevitably come into place. In what follows, according to this lesson, in a syntactic configuration like (6ii) I will take into consideration only the possibility that Transfer applies to the moved phrase (XP in (6ii)) but not to the phrase which houses the

edge at which it halts (ZP therein).[4]

## 3. The theoretical approach in details

As I outlined in Introduction, the theoretical approach to be advanced in this paper consists of the following ingredients:

(8) i. Transfer-based solution of PoP (see Section 2.2)
ii. Coincidence of the site where an SO is transferred with the site where it is interpreted (a corollary of the mapping function of Transfer)
iii. Interpretive constraint on the outcome of Transfer
iv. Lexical vs. functional interpretation

As for (8i), I explicated it in detail in Section 2.2. As for (8ii), I consider it a corollary of the mapping function of Transfer. When Transfer applies to an SO, it sends to the interpretive components the instruction that this SO should be interpreted at the site where it operates. SOs must be legible in the interpretive components after they have been transferred. The legibility of SOs depends not only on their internal state but also on their syntactic context. Labelability and the presence of unvalued features are instances of the internal factors to affect the legibility of SOs: if an SO is transferred to the interfaces with being unlabelable or with containing any unvalued feature in itself (or with both), it is doomed to crash the derivation. The syntactic context under which SOs are generated is another factor to impinge on their legibility. Namely, in order to be rendered legible, an SO has to be generated under a syntactic context where it can stand legible. Therefore, if an SO is transferred under a syntactic context where it cannot do so, it cannot be appropriately interpreted even though it is labelable or contains no unvalued feature in itself. (8iii) states this. As for (8iv), it is already known in the literature that languages have the dual system of semantic interpretation: one aspect of the semantic system is lexical, and the other is functional (Culicover 1970; Jackendoff 1969; Jenkins 1972; among others). I suggest that, when a certain SO is transferred at the edge of another one, the type of the former's semantic interpretation depends on the latter's category-type. Namely, if the latter is lexical, the former is lexically interpreted in the semantic component, and, if the latter is functional, the former is functionally interpreted therein. All this above is the outline of the theoretical approach

---

[4] An antithesis of this possibility is submitted by several researchers. For example, Donati (2006), Cechetto & Donati (2010, 2015), and Donati & Cecchetto (2011) point out the possibility that moved phrases may project as in (7), and then demonstrate that this possibility opens up a novel approach to free relative clauses (see also Chomsky 2008). In (i) below for instance, *what* is schizophrenic in that it behaves not only as the internal argument of the embedded verb but also as that of the matrix verb.

(i) *John read [what you read].*

(ii) *... read* [$_{what\text{P}}$ *what* [$_{\text{Relative Clause}}$ *you* <*what*>]]

This schizophrenic property of free relative pronouns can be accounted for as in (ii) above, where *what* moves to the edge of the embedded relative clause and determines the label of the node dominating it and that clause. In this movement dependency, *what* acts in its base position as the internal argument of the embedded verb, and it behaves in its derived position as if it were the internal argument of the matrix verb. If this analysis of free relative clauses is correct, it contradicts, and hence undermines, the approach to be advanced in this paper. However, it seems to me that it is not the case that only this analysis can account for the schizophrenic property just mentioned of free relative pronouns. See Riemsdijk (2006) for an alternative, and see Caponigro (2019) for an argument against Donati & Cecchetto (2011). Unfortunately, the limited space does not allow me to recapitulate the analyses presented in the references just cited.

to be pursued and developed in this paper.

### 3.1. Lexical vs. functional interpretation

I have to be more specific about what the lexical/functional interpretation is and why, given that a certain SO is transferred at the edge of another one, the former's mode of semantic interpretation is designated by the latter's category-type, but not the former's. To start with the first question, I follow the semantic division which has already been made in the classical literature (see the references cited above). The division classifies theta-theoretic relations as the lexical type of semantic interpretation, and discourse-oriented relations like focus-background and topic-comment relations as the functional type. To the former type, I would like to add modification and predication relations without any argument (but see below). It is here that we should note that these semantic concepts, irrespective of whether lexical or functional, are all relational: i.e., the semantic interpretation of a given SO can only be understood in terms of its interaction with another SO. This point will become clearer in what follows.

As for the lexical type of semantic interpretation, I suggest that all the semantic concepts that it subsumes, viz. theta-theoretic, modification, and predication relations, are common in that they share "&"-relations. The insight behind this originates from the (neo-)Davidosonian model of eventish semantics (Davidson 1967a, 1967b; see also Boeckx 2008 and Pietroski 2002). To see this, let us consider sentence "Brutus stabbed Caesar with a knife." This sentence can be paraphrased into "there was an event of stabbing, the event involved Brutus as the executor, and Caesar as the victim, and the event involved the use of a knife" (the original sentence and its paraphrase are both borrowed from Boeckx 2008). In this paragraph, we can find that the theta-theoretic relations and the modification relation involved in the sentence at hand are all represented as &-relations. This point can be crystalized by the following formal representation of this paraphrase in terms of the (neo-)Davidsonian model of eventish semantics: $\exists$e [Stab(e) & Agent(Brutus, e) & Patient(Caesar, e) & With(knife, e)]. The first and the second &-relations represent the theta-theoretic relations and the third one the modification relation: the former two stands for "Brutus is the executer of stabbing" and "Caesar is the victim of stabbing", and the latter stands for "the event of stabbing is with a knife." Predication, since Jespersen (1924), has been known to be primitive among the three lexical relations. Namely, this lexical relation is semantically denotated by a simple &-relation. For example, sentence "John is smart" can be roughly paraphrased as "there is a person who is smart, *and* the person is John". This paraphrase, though highly informal, indicates that the predication between *John* and *smart* in the sentence holds without the mediation of any special means. In a nutshell, all the semantic interpretations of the lexical type are translated into &-relations in terms of the (neo-)Davidsonian model of eventish semantics.

The semantic interpretations of the functional type are quite different from those of the lexical type in nature. As an example, let us take *Mary* in sentence "Mary, John praised." This noun phrase is dually interpreted: it is lexically interpreted as the internal argument of the verb *praise*, and functionally interpreted either as a focus or as a topic. The first interpretation is based on the lexical property as an argument-taker of the verb *praise* to which *Mary* relates. Put differently, this lexical property reflects on the lexical interpretation of *Mary* as an internal argument. Namely, the semantic interpretation of the lexical type involves a reflection of the lexically idiosyncratic property of either of the lexical items relating to each other under this interpretation. The second interpretation of *Mary*, on the other hand, is not implied either by its lexical meaning or by any of the lexical

properties of *John* and *praise* in the rest of the sentence. Whether *Mary* is interpreted as a focus or as a topic depends on the construal of the rest of the sentence. When the rest of the sentence is construed as a background, it requires *Mary* to be interpreted as a focus, and when construed as a comment, it requires *Mary* to be interpreted as a topic. The crossing is never possible. Whether the rest of the sentence is construed as a background or a comment is likewise not determined by any of the lexically idiosyncratic properties of *John* and *praise* therein. In the first place, all the semantic contents just mentioned, viz. focus, topic, background, and comment, are not attributable to the lexically idiosyncratic properties of lexical items.

Further noteworthy concerning the functional type of semantic interpretation is that these semantic units are complimentarily distributed. That is, *Mary* in the sentence used is never focalized when the rest of the sentence is construed as a comment, and the former is never topicalized when the latter is construed as a background. Accordingly, getting to know the functional interpretation of *Mary* amounts to getting to know how the rest of the sentence is construed, and vice versa. In addition, foci and topics cannot be interpreted by themselves without backgrounds and comments, respectively. Focalization and topicalization presuppose the presence of a constituent construed as a background and of one construed as a comment, respectively. Likewise, the latter two constituents, in order to be construed as such, demands what is specified as a foreground (i.e., a focus) and what the relevant sentence or discourse is about (i.e., a topic), respectively. In consideration of all this above, we can view focus and background, on one hand, and topic and comment, on the other hand, to be in a pairwise relation, and each of these members to be semantically complementary to its own partner (Krifka 1992; Neeleman and van de Koot 2008). The concept of complementation on this view will become important in the discussion to continue below.

In this paper, I suggest that semantic interpretations of the functional type are phenotypes of feature agreement on the conceptual-intentional interface. Feature agreement, Agree in short, is a syntactic operation to manipulate the valuation of unvalued features (Chomsky 2000, 2001). As mentioned before, unvalued features are uninterpretable in the interpretive components, and hence they have to get valued until they are sent off thereto. In order to get valued, unvalued features have to probe their valued counterparts. This probing process runs through the c-commanding domain of an unvalued feature (more aptly, of the SO whose head bears an unvalued feature). When an unvalued feature does well to find out its valued counterpart within its c-commanding domain, these two are said to match each other. This feature matching assimilates the value of the unvalued feature with that of its valued counterpart. The unvalued feature thus gets valued. This entire process (probing, matching, and value-assimilation/valuation) is referred to as Agree.

What I intend to propose is that syntactic dependencies established under Agree are expressed as functional interpretations in the semantic component. In order to illustrate how the former should be translated into the latter, I continue to use the sentence “Mary, John praised.” Here, I would like to, following Miyagawa (2010, 2017a,b), exploit discourse-oriented features (δ-features in short) as formal features responsible for denoting the discourse-oriented contents mentioned above, viz. topic, focus, background, and comment. In this paper, two values are posited for δ-features: one is “*B*” (the acronym of “Background”), and the other is “*C*” (the acronym of “Comment”). I suppose that unvalued δ-features and their valued counterparts are deployed to focalized or topicalized constituents and to syntactic heads specifying the illocutionary type of sentence, viz.

Complementizers (Comp(s) in short), respectively. When the head Comp has a δ-feature with *B*-value, it specifies the sentence it heads as a background, and when it has a δ-feature with *C*-value, it specifies the sentence it heads as a comment. This sentence-type specification property of δ-features is active only on Comps. Conversely speaking, this property is deactivated on any other syntactic element.

Let us now look at (9):

(9) [*Mary*{δ:__}, [$_{\text{CompP}}$ Comp{δ:*B*/*C*} ...]]

(9) zooms in only the relevant portion of the syntactic structure for the sentence used. There I use the curly-brace notation to indicate features: the δ-features are represented between curly-braces. In addition, I represent the attributes and values of features to the left and right, respectively, of the colon. The underbar is used to indicate that the feature is inherently unvalued. As seen in (9), *Mary* is assigned an unvalued δ-feature (represented as {δ:__}), the head Comp a valued δ-feature (represented as {δ:*B*/*C*}, where the slash means "or"), and the former merges with the SO which the latter heads, viz. CompP. When the δ-feature on Comp is valued as *B*, it stimulates CompP to translate into a background, and when valued as *C*, it stimulates CompP to translate into a comment. The δ-feature on *Mary*, being unvalued, has to prove its valued counterpart. In (9), this feature does well to find out the δ-feature on Comp as its valued counterpart, because the latter resides within the former's c-commanding domain. The δ-feature on *Mary* then results in matching the δ-feature on Comp, whereby the former's value assimilates into the latter's. That is, when the δ-feature on Comp has *B*-value, the δ-feature on *Mary* gets valued as *B*, and when the former has *C*-value, the latter gets valued as *C*. In this sense, we can say that *Mary* is syntactically dependent on CompP in that the former's unvalued δ-feature acquires a value from the latter's valued one (cf. De Vos 2007).

A question which remains to be addressed is how *Mary* should be interpreted either as a focus or as a topic. The δ-feature on this phrase has no semantic efficacy so as to translate it into such discourse-oriented content. As I mentioned above, δ-features functions as stimuli to construe the SO whose heads bear them either as backgrounds or as comments, relative to their own values. That said, it does not make any sense to state that the δ-feature on *Mary* stimulates it to get interpreted either as a background or as a comment. In the first place, this semantic effect of δ-features is active only on the syntactic heads Comps. How, then, should *Mary* be functionally interpreted? It is here that the concept of complementation noted above is helpful. My answer to this question is that the functional interpretation of *Mary* is the result of complementation. Let me be more specific about this. In (9), CompP is construed either as a background or as a comment, relative to the value of the δ-feature on its head. Whichever way it is construed, however, the ensuing semantic unit cannot stand alone: backgrounds and comments cannot be appropriately interpreted without their functional pair-mates foci and topics, respectively. To complete the construal of Comp thus demands complementation. We noted above that CompP is in an Agree-dependency with *Mary* in terms of δ-features. Through this syntactic dependency, the complementation is executed. That is, the construal of CompP as a background or a comment demands its dependent *Mary* to be interpreted so as to be complementary to it: when CompP is construed as a background, *Mary* should be interpreted as a focus, because these two semantic units are complementary to each other; when CompP is construed as a comment, *Mary* should be interpreted as a topic for the same reason. The functional interpretation

of *Mary* thus results from the complementation of CompP's construal. This is how to assign a functional interpretation to *Mary*. Throughout the argumentation so far, we should note that the way in which a functional interpretation comes about is reflected by the Agree-dependency between two SOs under this interpretation. This is what I intended in stating that semantic interpretations of the functional type are phenotypes of Agree. Given that this is correct, it follows that the manifestation of a functional interpretation presupposes the establishment of Agree between two SOs under this interpretation.

**3.2. How to determine the type of semantic interpretation**

Having elucidated what the lexical/functional interpretation is and how these two semantic types differ, let me turn to the second question, viz. why, given that a certain SO is transferred at the edge of another one, the type of the former's semantic interpretation should be determined by the latter's category-type, but not the former's. To consider this why-question, let us reuse the toy structure in (4), repeated as (10) below:

(10) $[_{\alpha}$ XP, YP] (= (4))

In (10), let us suppose as before that Transfer applies to XP at the edge of YP (whereby α is to be labeled as Y(P)). This operation, as its mapping function, sends to the semantic component the instruction that XP should be interpreted at the site where it is transferred. Here we should note that the information sent to the semantic component at this point is only about XP and does not include any information about YP. Therefore, at this point, the semantic component does not know how XP should be interpreted. Recall that the semantic interpretation of the sort under consideration is relational in that it can only be understood in terms of the interaction between SOs. Therefore, just after XP has been transferred, it cannot enter an interpretive relation with any other SO in the semantic component. Next, suppose that, as a subsequent step of derivation, Transfer applies to α, viz. Y(P). This instance of Transfer inevitably has to be implemented for the sake of the evaluation of the full derivation. It is at this very point that the information about YP is sent off to the semantic component. It is worthy to note here that this information includes the following sub-information: α contains XP and YP, XP has already been transferred at the edge of YP, and α is equivalent to YP in terms of their labels. This whole information is then enough to instruct the semantic component to interpret XP at the edge of YP. In order to do so, the semantic component has to make reference to XP's interpretive relation with YP. This owes to Principle of Compositionality. Namely, getting to know how XP is related with YP is a necessary step for the sake of interpreting α (now labeled as Y(P)) in a compositional manner. Under the present approach, the mode in which XP is related with YP is two-fold: one is lexical and the other functional. Which way is opted for depends on the category-type of YP. Why this is so is at issue now. My answer to this why-question is because, just when the semantic component receives the information about α including the category-type of YP, it recognizes that XP is interpretively related with YP. Namely, the information about α triggers the recognition or emergence of the interpretive relation between XP and YP, and it is this very information that the semantic component refers to for the sake of determining how to interpret XP under the relation having thus emerged. It should be noted here that α is equivalent to YP in terms of their labels. This means that α inherits the category-type information from YP. By referring to this information, the semantic component determines which way XP should be interpreted, lexically or functionally. In a nutshell, in order to determine the type of semantic interpretation, the semantic component refers to the category-type of an SO whose information triggers the emergence of an interpretive relation. In the case under consideration, α

amounts to that SO. As noted above, the information about XP, just when sent to the semantic component, does not trigger the emergence of any interpretive relation between XP and YP. Therefore, the semantic component does not refer to this information in order to determine how to interpret XP at the edge of YP: in the first place, it does not yet recognize any information about YP. This is the reason why XP has no right to designate its own mode of semantic interpretation.

### 3.3. Summary

In this section, I have elaborated the theoretical approach in this paper. In particular, I devoted much space to illuminating what the lexical/functional interpretation is and how these two semantic interpretations differ. Moreover, I elucidated how to determine the type of the semantic interpretation which a certain SO receives at the edge of another one. The following lists up the findings in this section:

(11) i. Lexical interpretations = theta-theoretic, modification, predication relations = &-relations (in the sense of the (neo-)Davidsonian eventish semantics)

ii. Functional interpretations = discourse-oriented relations (e.g., focus-background and topic-comment relations) = phenotypes of feature agreement/Agree

iii. Flow towards the determination of the type of semantic interpretation
In [$_{\alpha}$ XP, YP],

a. Transfer applies to XP (for instance).

b. It sends to the semantic component the instruction to interpret XP at the site where it applies (at this point, the semantic component receives only the information about XP).

c. Transfer apples to α (which is now labeled as Y(P)).

d. It sends to the semantic component the instruction to interpret α at the site where it applies.

e. At (iiid), the semantic component is informed that α contains XP and YP, XP has already been transferred at the edge of YP, and α is equivalent to YP in terms of their labels.

f. Based on the information at (iiie), the semantic component interprets XP at the edge of YP according to the instruction at (iiib).

g. The semantic component determines how to interpret XP at the edge of YP, relative to the category-type of α. This is because α's category-type is part of the information at (iiie), which triggers the emergence of the interpretive relation between XP and YP. Note that α's category type is identical with YP's, because α and YP both have the same label.

According to the classical literature, the semantic interpretation of the lexical type subsumes theta-theoretic relations, and that of the functional type encompasses discourse-oriented relations such as focus-background and topic-comment relations. In this paper, modification and predication relations are added to the lexical type of semantic interpretation. The semantic interpretation of the lexical type is taken to be synonymous to "&"-relation in the sense of the (neo-)Davidsonian eventish semantics. The semantic interpretation of the functional type, on the other hand, is considered a phenotype of feature agreement/Agree on the conceptual-intentional interface. Namely, that a given SO is functionally interpreted presupposes that this SO enjoys an Agree-dependency with another SO. Agree is not involved in the expression of the lexical type of semantic interpretation, which, instead, reflects a certain lexically idiosyncratic property of either of the SOs that interpretively relate with each other.

The functional type of semantic interpretation does not have such reflection.

The flow towards the determination of the type of semantic interpretation is summarized in (11iii) above. The point to be noted is (11iii-g), which states how to determine the type of the semantic interpretation which a certain SO receives at the edge of another one. According to (11iii-g), when an SO is transferred at the edge of another one, the former should be lexically interpreted if the latter is lexical, and functionally interpreted if the latter is functional. In the first case, the semantic component recognizes the interpretive relation between XP and YP, classifying it as lexical by virtue of YP being lexical. This lexical relation thus recognized should be mapped onto the conceptual-intentional interface as any of the following lexical interpretations: a theta-theoretic, modification, or predication relation. If XP has no lexical meaning or denotation suitable to any of such lexical interpretations, it cannot be semantically interpreted at the edge of YP in an appropriate way. This results in crashing the derivation. Hereafter, I refers to the requirement that in [$_{\gamma}$ α, β], where α is transferred and β is lexical, α should be lexically interpreted at the edge of β, as Lexical Criterion (henceforth abbreviated as LC). In addition, I say that, if α is lexically interpreted at the edge of β in an appropriate way, LC is saturated. The saturation of LC is referred to as LC-Compliance. Accordingly, when Transfer applies to the edge of a lexical category, its outcome has to saturate LC or attain LC-Compliance.

In the second case, where YP is functional, the interpretive relation recognized between XP and YP is classified as functional. This relation, hence, should be mapped onto the conceptual-intentional interface as a functional interpretation such as a focus-background or topic-comment relation. As this type of semantic interpretation is taken as a phenotype of Agree on the conceptual-intentional interface, the failure to establish an Agree-dependency between XP and YP means that XP cannot be functionally interpreted at the edge of YP. This likewise results in crashing the derivation. Hereafter, I refer to the requirement that in [$_{\gamma}$ α, β], where α is transferred and β is functional, α should be functionally interpreted at the edge of β, as Functional Criterion (henceforth abbreviated as FC). I say, moreover, that if α is functionally interpreted at the edge of β, FC is saturated. The saturation of FC is then dubbed FC-Compliance. Accordingly, when Transfer applies to the edge of a functional category, its outcome has to saturate FC or attain FC-Compliance. As noted above, the saturation of FC or FC-Compliance presupposes that α enjoys an Agree-dependency with β.

Summarizing what I have stated in the last two paragraphs:

(12) i. Lexical Criterion (LC):
In [$_{\gamma}$ α, β], where α is transferred and β is lexical, an interpretive relation should be established between α and β, and classified as lexical.
ii. LC-Compliance:
In [$_{\gamma}$ α, β], where α is transferred and β is lexical, α is lexically interpreted.

(13) i. Functional Criterion (FC):
In [$_{\gamma}$ α, β], where α is transferred and β is functional, an interpretive relation should be established between α and β, and classified as functional.
ii. FC-Compliance:

In [$_{\gamma}$ α, β], where α is transferred and β is functional, α is functionally interpreted.

All in all, (8) (at the outset of this section), (11), (12), and (13) are altogether combined into the theoretical approach in this paper. Armed with this approach, we will tackle the two conundrums, viz. the tension problem and the variation problem.

### 3.4. Tutorial

Before implementing this approach, I would like to prepare a tutorial to help the understanding of what will be accounted for in the subsequent sections. The tutorial proceeds on the following two toy configurations:

(14) a. [$_{\alpha}$ XP, [$_{YP}$ $Y_L$ ... <XP> ...]]
b. [$_{\beta}$ WP{F:*val*}, [$_{ZP}$ $Z_F${F:*val*} ... <WP> ...]]

In (14a) XP moves from the inside of YP to its edge, and in (14b) WP moves from the inside of ZP to its edge (I continue to use the angle-bracket notation for representing copies). YP and ZP are categorially specified as lexical and functional, respectively. I mark this categorial specification by attaching the subscripts "L" and "F" to the heads of YP and of ZP, respectively. This subscript notation is used in what follows to specify heads or phrases as lexical or as functional. In (14b), the features Fs (represented between curly-braces) are allocated to WP and ZP. The feature on the head of ZP is inherently valued, but the feature on WP is not (as indicated by the underbar). As the former counts as the valued counterpart of the latter and is c-commanded by the latter, these two match each other. Hence, the latter's value assimilates into the former's, as shown.

XP's movement in (14a) and WP's movement in (14b) result in breeding α and β, respectively. These two resulting SOs are not unambiguously labelable as they are. In order to render them labelable, XP and WP have to be transferred at the edges of YP and of ZP, respectively. As a result, α and β are labeled as Y(P) and Z(P), respectively, in the manner illuminated in Section 2.2. When transferred, XP and WP halt at those sites. In addition, they are compelled to get interpreted thereat. This is due to the corollary of the mapping function of Transfer, viz. the coincidence of the site where an SO is transferred with the site where it is interpreted. How to interpret XP and WP should meet LC and FC, respectively. Namely, XP should be lexically interpreted at the edge of YP, and WP should be functionally interpreted at the edge of ZP. This is because YP and ZP are lexical and functional, respectively. Particularly, WP should enjoy an Agree-dependency with ZP, since establishing this syntactic dependency is a prerequisite for saturating FC. In (14b), as stated above, Agree is successfully conducted between WP and ZP in term of their shared features. Therefore, it is guaranteed that the outcome of transferring WP at the edge of ZP in (14b) saturates FC, insofar as the Agree-dependency between WP and ZP can be translated into a certain functional interpretation. If LC and FC fail to be met in (14a) and (14b), respectively, this means that XP and WP cannot stand legible in the semantic component. Their illegibility hence prevents the convergence of derivation, even though transferring them ensures the labelability of α and β.

The lesson from this tutorial teaches us how the present approach solves the tension problem and the variation problem. As for the first problem, our approach solves it as follows. First and foremost, movement always produces an SO consisting of two phrases, one of which has been moved. And, once this moved phrase is

transferred, the movement halts. As the effect of Transfer, the moved phrase becomes inaccessible to further syntactic operations, whereby the SO produced by the movement becomes unambiguously labelable. Moreover, as a corollary of the mapping function of Transfer, the moved phrase gets semantically interpreted at the site where it has been transferred. The semantic interpretation which the moved phrase thus receives has to comply with LC or with FC: if the moved phrase halts at the edge of a lexical category, it has to be lexically interpreted at that site, while if it halts at the edge of a functional category, it has to be functionally interpreted at that site. Otherwise, the moved phrase would fail to stand legible in the semantic component, hence the derivation would crash. Accordingly, in order for the derivation to converge, the movement should halt at the site where the moved phrase can be semantically interpreted in compliance with LC or with FC. This is how to resolve the tension problem in terms of the present approach: movement comes free, but it doesn't look so free for the reason why its canonical halting site is thus restricted. In this sense, we have to be careful where Transfer applies. This solution to the tension problem can be extended to the variation problem, because the canonical halting site of movement varies with the category-type of the SO at whose edge the moved phrase halts. As for this problem, we will delve into more details in Section 4.

The lesson from this tutorial also provides us with at least two predictions to be confirmed. One is that, when a moved phrase halts at the edge of a lexical category despite of its lexical meaning being unsuitable to any lexical relation (i.e., theta-theoretic, or modification, or predication relation) with that lexical category, the sentence should be deemed ungrammatical. The other is that, when a moved phrase halts at the edge of a functional category without enjoying any Agree-dependency with that functional category, the sentence should be ruled out as ungrammatical. The confirmation of these two predictions awaits us in Section 5.

## 4. The theoretical approach in action

### 4.1. Prologue

We are now in a position to implement the theoretical approach assembled in the last section. Here, we examine English and Japanese as a case study. These two languages have so far been frequently compared with each other in the sense that the former is representative among "rigid word-order" languages and the latter is representative among "free word-order" languages. For space, we confine ourselves to word-orders in clausal domain, putting aside word-orders in nominal domain from the scope of the present paper.

In general, the clausal architectures of both languages are assumed to be basically the same. The following scheme represents the basic structure up to TP in clausal domain:

(15) $[_{TP}$ T{φ:*val*; Case:*Nom*}, $[_{\alpha}$ DP1{φ:*val*; Case:__}, $[_{vP}$ $v_L$, $[_{VP}$ ... DP2 ... ]]]]

DP1 and DP2 in (15) stands for an external argument and an internal argument, respectively, of the verbal predicative projection, which is represented as vP therein. The internal argument DP2 is base-generated within the lower verbal stratum VP, and the external argument DP1 is merged with vP, the upper verbal stratum. The second merging operation, as shown, produces the complex SO α, which at the present is not unambiguously labelable. As will be clear soon later, the label of α is fixed as v(P). In this sense, the basic clausal configuration in (15) is in accordance with Split-VP Hypothesis (Larson 1988) and Predicate-Internal Subject Hypothesis

(Fukui and Speas 1986; Kitagawa 1986; Koopman and Sportiche 1991; Kuroda 1988). Moreover, (15) displays the conventional featural constellation: the tense head T, which is merged with α, is assigned the set of unvalued φ-features and a Case feature valued as *Nom*(inative), and DP1 the set of valued φ-features and an unvalued Case feature. The term φ-features is used according to the notational convention for representing the sets of agreement features including number, gender, and person features. The unvalued φ-features on T can be valued via Agree at the stage where T is merged with α. This is because at that stage they can do well to find out the φ-features on DP1 as their valued counterparts within the c-commanding domain of T. For avoiding clutter, the features of DP2 and others are omitted.

Here, I assume, following Fukui (1988, 1995) and Zushi (2003, 2005), that only the difference between English and Japanese with respect to (15) lies in the categorial specification of T. These authors suggest that T in English is specified as functional and the corresponding head in Japanese as lexical. Namely, the categorial specification of T is parametrized as below:

(16) Parameter for Categorial Specification of T:
Functional (e.g., English) vs. Lexical (e.g., Japanese)

As a piece of evidence for (16), the authors present the fact that English has an overt expletive *there* and Japanese has no such element. Their reasoning is this: expletives have no lexical content and their structural niche is assumed to be confined to the frontal domain (or, to use a familiar term, Specifier) of TP; it is standardly assumed that lexical categories require any element within their domains to be rich in lexical content, while functional categories do not; consequently, the availability of expletives in English suggests that T in this language is functional, and the absence of such elements in Japanese indicates that the corresponding head in this language is lexical. This reasoning provides the parametric setting in (16).

Putting aside for a moment what empirical consequences the parameter setting in (16) has under the present approach, let us now return to (15), then paying attention to α therein. As mentioned before, α is not unlabelable as it stands. We have two options at hand to render α labelable. One is to transfer DP1, and the other is to move it. When the first option is opted for, α is labeled as v(P). This is because the head of DP1 becomes inaccessible to Labeling for α due to the inaccessibility effect of Transfer and the head of vP remains accessible to that operation. The outcome of transferring DP1 at the edge of vP has to saturate LC, because vP is classified as lexical. And, indeed it does, since DP1 can be theta-theoretically related with vP as its external argument. The option under consideration, however, results in preventing the derivation from converging, since DP1 is transferred with its Case feature unvalued. A caution is in order here. The application of Transfer to DP1 at the stage of (15) is not prohibitable in and of itself. Recall that Transfer is unrestrained just like Merge. When Transfer applies to DP1 at the stage of (15), it simply crashes the derivation. That's all.

When the second option is opted for, it derives the following derivation:

(17) $[_{\beta}$ DP1{φ:*val*; Case:<u>*Nom*</u>}, $[_{TP}$ T{φ:<u>*val*</u>; Case:*Nom*}, $[_{\alpha}$ <DP1>, $[_{vP}$ $v_L$, $[_{VP}$ ... DP2 ... ]]]]]

As seen, DP1's movement delivers two occurrences of it at the edges of vP and of TP each. As a result, α ceases to contain DP1, since the former no longer dominates all the occurrences of the latter. The label of α is then determined by the head of vP and only by this head. This is because the latter is contained within α, remains accessible, and not c-commanded by any other accessible head. Accordingly, the label of α under the second option is the same as under the first one. However, the second option yields a different consequence from the first one. The consequence is that it becomes feasible to establish an Agree-dependency between DP1 and TP in terms of their Case features. As the edge of TP c-commands its head, the unvalued Case feature on (the higher occurrence of) DP1 does well to find out and match the valued Case feature on the head of TP. As a result, the former's value assimilates into the latter's. The unvalued Case feature on DP1 thus gets valued as *Nom*.

In the way thus, DP1's movement in (17) renders α labelable. However, it simultaneously breeds another unlabelable SO, which is represented as β in (17). In order to render β labelable, we have two options as before: one is to move DP1 and the other is to transfer it. Although the Transfer-option at the stage of (15) ends in crashing the derivation due to the presence of the unvalued Case feature on DP1, the same option at the stage of (17) does not incur the same result. This is because, as stated above, DP1's movement feeds the valuation of the unvalued Case feature on DP1. In (17), therefore, the option to transfer DP1 at the edge of TP does not prevent the convergence of the derivation. Moreover, implementing this option renders DP1 inaccessible to further syntactic operations, whereby the label of β is fixed as T(P) because the head of TP remains accessible and c-commands all other accessible heads within β. In passing, the option to move DP1 from the edge of TP yields the same result: by virtue of this movement, β ceases to contain DP1, hence its label is unambiguously fixed as T(P) since the head of TP c-commands all other heads that remain accessible within β, except for the head of DP1. (However, this option demands further extension of the derivation in (17) by merging β with another head.)

It is here that the parameter setting in (16) comes into the picture. Given (16), the outcome of transferring DP1 at the edge of TP in (17) varies between English and Japanese. In English, the outcome in question has to saturate FC. This is because TP in this language is categorially specified as functional. In order to attain FC-Compliance, DP1 has to enjoy an Agree-dependency with TP, since establishing this syntactic dependency between these two is a prerequisite for saturating FC. Indeed, DP1 stands in such dependency, as stated above. Here, a question immediately arises what kind of functional interpretation DP1 obtains. For answering this question, I tentatively assume, following De Vos (2007), that DP1 is functionally interpreted as a subject. For De Vos, given that a certain phrase enjoys an Agree-dependency with TP in terms of their Case features, this phrase is assigned the grammatical function subject. The notion of subject, much like focus, topic, etc., is relational. In this sense, a subject cannot stand semantically by itself, and hence demands complementation. As the functional complement of subject, I tentatively assume the notion of presentation. This notion, which was originally ventilated by Brentano (1973), is a hard one to grasp, so I do not afford to devote much space to it. For the present purpose, it suffices to state that a presentation is equal to a description of an individual or an eventuality (Ladusaw 1994) and that the constituent denoting an individual or an eventuality which satisfies a description is identified as a subject. Moreover, I tentatively assume that the syntactic constituent responsible for presenting a description is TP and that the feature stimulating TP into a presentation is a Case feature with Nominative value. All in all, I would like to suggest that subject and presentation are in a pairwise relation and complementary to each other, and moreover that this subject-presentation relation is the functional interpretation into which an Agree-

dependency between TP and its dependent in terms of their Case features translate. This translation runs as follows: at first, TP is construed as a presentation by virtue of the Nominative-valued Case feature on its head, and then, in order to stand as such, it assigns the grammatical function subject to its dependent (DP1 in the case at hand) through the Agree-dependency between these two. Accordingly, DP1 in (17) is functionally interpreted as a subject as a consequence of its complementation to TP as a presentation.

The grammatical function of DP1 in (17) as subject holds in Japanese as well. However, this interpretation is orthogonal to whether or not Japanese licenses DP1 as legible at the edge of TP. According to (16), TP in this language is categorially specified as lexical, as opposed to the corresponding phrase in English. Therefore, Transfer at the edge of TP in Japanese has to saturate LC as its outcome. For attaining LC-Compliance, DP1 in (17) has to bear any of the following lexical relations with TP: a theta-theoretic, or modification, or predication relation. However, the first and the second lexical relations do not hold between DP1 and TP in (17): neither is TP an argument-taker in the theta-theoretic sense, nor is DP1 a modifier. The only option left to us is a predication relation. It is this relation that I would like to suggest holds between the two. In fact, Fukui (1988, 1995) and Zushi (2003, 2005) treat Japanese TP as a predicate and attribute its predicate-like property to its lexical nature. To the extent that DP1 in (17) denotes a referent suitable to a predication relation with TP, the outcome of transferring it at the edge of TP is deemed LC-Compliant.

If this is correct, it is then predicted that Japanese TP accommodates more than one constituents in its frontal domain insofar as it can predicate them. In order to confirm this prediction, let us consider the following schema:

(18) $[_{\gamma}$ YP, $[_{\beta}$ XP, $[_{TP}$ $T_L$ ... ]]]

In (18), XP merges with TP as a lexical category, thereby β created, and with β, YP merges, then γ created. β and γ are not unambiguously labelable, since either of them consists of two phrases. When XP and YP each are transferred, β and γ are both rendered labelable in the manner explicated so far. After XP has been transferred, the head of TP counts as the nearest to β among the accessible heads within it, and hence this head determines the label of β. β is thus labeled as T(P). The outcome of transferring XP at the edge of TP has to saturate LC, since TP is a lexical category. After YP has been transferred, the head of TP continues to be accessible, and counts as the nearest to γ. Hence, this head also fixes the label of γ, i.e., γ is likewise labeled as T(P). It should be noted here that β and TP are equivalent to each other in terms of their labels and hence categorially classified as lexical alike. Therefore, the outcome of transferring YP at the edge of β likewise has to be LC-Compliant. In order that XP and YP may be lexically interpreted at their respective sites, they each only have to be semantically related with TP and β, respectively, via predication, since Japanese allows TP to behave as a predicate by virtue of its lexical nature. The prediction, then, is that Japanese does not prevent the derivation in (18) from converging to the extent that XP and YP denote referents suitable to predication.

In fact, it has been well documented in the literature on Japanese syntax that Japanese allows for the so-called multiple “subject” construction (Akiyama 2004, 2005; Fukui 1986, 1988, 1995; Hiraiwa 2001; Ura 2001; Vermeulen 2005; Zushi 2003, 2005; among others). Some examples are provided below:

(19) a. John-ga me-ga ao-i.
John-NOM eye-NOM blue-PRES
'As for John, his eyes are blue.'
b. Syutoken-ga zinkoo-ga oo-i.
metropolitan.area-NOM population-NOM large-PRES
'The metropolitan area has a large population.'

In (19a), the portmanteau phrase *ao-i* (consisting of the adjective and the present tense morpheme) is predicated of the nominal phrase *me*, and these two phrases combine to form a larger predicate, which describes the eyes-being-blue property of the referent denoted by another nominal phrase *John*. We can easily construct the syntactic structure for (19a) by substituting *John*, *me*, and *ao-i* for YP, XP, and TP, respectively, in (18). Much the same goes for (19b). The fact that this construction is utterly grammatical in Japanese convincingly demonstrates the well-formedness of the derivation in (18) in this language. In passing, Japanese basically allows an indefinite number of "subjects" to appear in a single sentence insofar as the clause to the right of each subject describes the property of its referent (e.g., *Nihon-ga syutoken-ga zinkoo-ga oo-i*. 'As for Japan, the metropolitan area has a large population.')

The grammaticality of the multiple "subject" construction in Japanese sharply contrasts this language with English, which accommodates only one subject per sentence. (20) below for instance is deadly ungrammatical in English.

(20) *John, his eyes are blue.

The ungrammaticality of (20) falls into place under the present approach. Consider the following structure:

(21) $[_{\gamma}$ YP, $[_{\beta}$ XP, $[_{TP}$ $T_F$ ... ]]]

(21) is no different from (18) except for the categorial specification of T. That is, TP in (21) is functional, in contrast to the corresponding phrase in (18), which is lexical. In (21), for the sake of determining the label of β in an unambiguous way, XP has to be transferred. When this instance of Transfer is implemented, β is labeled as T(P). YP, too, has to be transferred, for the sake of determining the label of γ in an unambiguous way. This instance of Transfer feeds the unambiguous determination of γ's label as T(P). All this is much the same as in (18). However, each instance of Transfer in (21), unlike in (18), has to saturate FC as its outcome. Namely, XP and YP in (21) should be functionally interpreted, because their respective sites each are the edge of a functional category. Note that β receives the same categorial specification as TP. If XP bears an unvalued Case feature, it can and must enter into Agree with TP for the purpose of assigning a value to that Case feature. Hence, if so, XP can be functionally interpreted as a subject vis-a-vis TP as a presentation. The outcome of transferring it at the edge of TP in (21) is then deemed FC-Compliant. How about YP, then? If this phrase bears an unvalued Case feature as well, can it likewise enter into Agree with TP for the same purpose? I reckon that the answer should be negative. For, Agree is exclusive in nature, i.e., a one-to-one relationship (Fukui 1986, 1988, 1995; pace Hiraiwa 2001a, 2001b; Ura 2000). This exclusiveness prohibits YP from entering Agree with TP in terms of their Case

features, because the latter has already "used up" its own Case feature for valuing the Case feature of XP. In other words, the Case feature of TP is no longer active for newly establishing an Agree-dependency with YP. Without this syntactic dependency, YP cannot be functionally interpreted at the edge of β. Hence, FC will not be saturated if YP is transferred at the edge of β. Even if YP does not bear an unvalued Case feature, the result is no different, because, if so, this phrase has no way to enter into Agree with TP. Consequently, among XP and YP in (21), only the former can be interpreted at the edge of TP in compliance with FC insofar as it bears an unvalued Case feature, the latter inevitably causing a violation of FC regardless of whether it bears an unvalued Case feature. Namely, only XP, but not both, can be interpreted as a subject. If the derivation in (21) withholds Transfer from applying to YP at the edge of β, it will not incur a violation of FC, but instead it will leave γ unlabelable throughout, hence failing to converge. In the end, this derivation cannot be saved from the failure of convergence. This explains the ungrammaticality of (20). It is then concluded that English accommodates only one subject per sentence. The uniqueness of subject in English sentence, as argued so far, follows from the functional nature of T in English and the exclusiveness of Agree. Accordingly, the typological difference between English and Japanese with respect to the multiplicity or uniqueness of subject per sentence lends support to postulating a categorial parameter such as in (16) concerning T in these two languages.

**4.2. An excursus: Deriving the EPP effect**

Before stepping to the main issue to be addressed in this section, I would like here to point out a theoretically intriguing consequence of the present approach. In order to understand the point clearly, consider the following structural representation:

(22) $[_{TP}$ //DP1{φ:*val*; Case:*Nom*}//, $[_{TP'}$ T{φ:*val*; Case:*Nom*}, $[_{vP}$ <DP1>, $[_{vP'}$ $v_L$, $[_{VP}$ ... DP2 ... ]]]]]

(22) represents the derivational structure resulting from the labeling of α and β in (17). The double-slash notation continues to be used for representing SOs having been transferred: DP1 therein is already transferred, hence represented between double-slashes. In addition, the prime notation is used to indicate non-terminal phrasal nodes whose labels are identical to those of the nodes that immediately dominate them. vP′ and TP′ correspond to such nodes. The usage of these notations is only for visual purposes. What we should note here is that the derived configuration in (22) exhibits the effect of the Extended Projection Principle (Chomsky 1981), usually abbreviated as EPP. The EPP, roughly speaking, dictates that every sentence should have a subject. Given that the top-most TP in (22) corresponds a sentence, the configuration therein exactly satisfies the requirement of this principle, since, as argued above, DP1 at the edge of TP′ functions as a subject. More important is that the derivational plot to derive (22) has no recourse to any stipulative device like the EPP feature or the like. The point in this plot is that DP1, which has the fate to function as a subject, cannot help but undergoing movement to the edge of TP′ and Transfer thereat, and otherwise the derivation would be doomed to crash due to both the presence of the unvalued Case feature on DP1 and the failure to determine the label of β (which is represented as TP in (22)) in (17) in an unambiguous way. In order for there to be no misunderstanding about this, it should be repeated that the movement of DP1 in question is not triggered either by the need to value the unvalued Case feature on DP1 or by the unambiguous determination of the label of α (which is represented as vP in (17) and (22)) in (15). Please recall here the leitmotif of the current minimalist framework of generative grammar on which the present paper bases itself, viz. that Merge comes free regardless of its variant. If that movement does not happen, the

result is merely that the derivation ends in crashing. That's all. Conversely speaking, the derivation will converge if the movement happens to take place. Simply, the derivation has no option but the invocation of the movement in order to avert crashing. Furthermore, after DP1 has moved to the edge of TP, it has no other option rather than undergoing Transfer thereat: otherwise, the label of β in (17) could never be determined in an unambiguous way. This Transfer-option is not forced, either, since Transfer is free just like Merge. In a nutshell, the convergence of (22) is nothing but a product of chance. It is unquestionably clear that it is feasible to derive (22) without postulating any triggering force such as the EPP feature or the like. Namely, without recourse to such stipulative device, the present approach fares well to explain the EPP effect, viz. why every sentence should have a subject. This not only implies that it is no longer necessary to postulate the EPP as an independent principle, but also that the EPP effect is merely a by-product of the implementation of all the syntactic operations indispensable for the convergence of the derivation. This theoretical consequence is considered desirable in the light of the minimalist tenet which encourages any endeavor to reduce the number of theoretical devices as far as possible.

### 4.3. Resolving the variation problem: a case study of English and Japanese

The main problem to be addressed in this section emerges in the following derivational structure:

(23) $[_{\gamma}$ DP2, $[_{TP}$ //DP1{φ:*val*; Case:*Nom*}//, $[_{TP'}$ T{φ:*val*; Case:*Nom*}, $[_{vP}$ (<DP1>, $[_{vP'}$ $v_L$, $[_{VP}$ ... <DP2> ... ]]]]]]

(23) is derived from (17) by moving DP2 from the inside of VP to the edge of TP. As a result, γ is created, which is not unambiguously labelable as it is. This movement is not prohibitable in and of itself, since movement is considered unrestrained as a variant of Merge in terms of the current model of grammatical architecture. If DP2 lingers at the edge of TP, it should undergo Transfer thereat: otherwise, γ would be left unlabelable throughout, hence preventing the convergence of derivation. After DP2 has been transferred, the label of γ is fixed as T(P), since the head of TP remains accessible within γ and is not c-commanded by any other accessible head within it. Given that asymmetrical c-command governs phonological ordering relations among constituents (viz., Kayne's [1994] Linear Correspondence Axiom), the resulting configuration linearly places DP2 to the left of DP1, since the former asymmetrically c-commands the latter but not vice versa. As DP1 and DP2 in (23) correspond to an external argument (or a subject phrase) and an internal argument (or an object phrase), respectively, the linear order thus exhibited shows the precedence of the object phrase over the subject phrase. This precedence relation is impossible in English except for special grammatical circumstances like sentences involving topicalization or focalization, but not in Japanese without limitation in principle (ignoring discourse-oriented or pragmatic factors). The issue to be addressed, then, is why this is so. I repeat with emphasis that, in terms of the current model of grammatical architecture, the movement yielding the preceding relation in question is in principle feasible in both English and Japanese. The issue at stake, hence, can be restated as why these two languages differ with respect to phrasal permutation or scrambling despite the liberty of movement as a variant of Merge.

#### 4.3.1. The case of English

The reason why English disallows the preceding relation in question is because in this language the derivation in (23) is not eligible to converge. The failure in convergence of (23) in English is attributed to the failure to establish an Agree-dependency between DP2 and TP (more aptly, the head of TP) therein. As TP in English is functional in terms of the parameter setting in (16), DP2 has to be functionally interpreted at the edge of TP. Namely, the

outcome of transferring DP2 at the edge of TP has to attain FC-Compliance. For this purpose, DP2 has to enjoy an Agree-dependency with TP. However, it is no longer possible to establish such a syntactic dependency between DP2 and TP. This is because the head of TP is already inactive for establishing a new Agree-dependency in terms of Case features: it has already entered Agree with DP1 for the valuation of the latter's unvalued Case feature. Once again, Agree is exclusive in nature, and hence, once a given feature enters this syntactic relation with its matching feature, it can never do so with another. The failure to establish an Agree-dependency between DP2 and TP ends in the failure to attain FC-Compliance with respect to the outcome of transferring DP2 at the edge of TP. For this reason, DP2 cannot be semantically interpreted at the edge of TP in compliance with FC. This then renders DP2 illegible at that site, hence preventing the convergence of derivation in (23). If DP2 desists from undergoing Transfer at the edge of TP, the failure to saturate FC will be circumvented. However, if it does so, the label of γ will never be fixed in an unambiguous way. The indeterminacy of γ's label likewise results in the failure in convergence. In the end, (23) for English falls into an insurmountable dilemma: if DP2 is transferred at the edge of TP, FC will be violated, on one hand, while, if not, γ will be left unlabelable throughout, on the other hand. In either case, the derivation is doomed to fail to converge. This explains why English does not tolerate the preceding relation in question.

According to what has been stated above, we can expect hat English allows an object phrase to precede a subject phrase if the former acquires a functional interpretation at the edge of an additional functional projection over TP. In fact, this language can place an object phrase in the initial position of sentence when this phrase is functionally interpreted as a focus or as a topic. This fact was already instantiated by the sentence we used before, repeated below as (24):

(24) Mary, John praised.

In (24), the object phrase *Mary* is preposed to the sentence-initial position, where it is assigned either a focus or topic interpretation. Standardly, focalized or topicalized phrases are assumed to be placed in the frontal domain of the functional projection called CompP. In addition, let us here recall the featural system in the present approach that derives discourse-oriented relations like focus and topic interpretations. To recap it, we are following Miyagawa (2010, 2011a,b), exploiting δ-features as formal features responsible for expressing such relations and assuming that a valued δ-feature and its unvalued counterpart are assigned to the head of CompP and to a phrase to be focalized or topicalized, respectively. Positing "*B*(ackgound)" and "*C*(omment)" as the values for δ-features, we are assuming that *B*-valued and *C*-valued δ-features stimulate CompP into a background and a comment, respectively. According to all this above, the relevant portion of the syntactic structure for (24) and the relevant featural constellation therein can be represented as below:

(25) [$_{\pi}$ *Mary*{δ:__}, [$_{\text{CompP}}$ Comp{δ:*B*/*C*}, [$_{\text{TP}}$ *John praised* <*Mary*>]]]

In (25), the object phrase *Mary* merges with CompP by moving to its edge from the inside of TP. The SO produced by this movement is represented as π. Consisting of the two phrases *Mary* and CompP, π is not unambiguously labelable as it is. To avoid this situation, *Mary* should be transferred at the edge of CompP. After this instance of Transfer has been implemented, the head of CompP counts as the nearest to γ among the accessible heads within

it. As a result, this head fixes the label of γ: i.e., γ is labeled as Comp(P). Moreover, this instance of Transfer phonologically realizes *Mary* at the edge of CompP. Given that asymmetrical c-command governs the precedence relation among constituents, the linear order emerges of *Mary* preceding over the subject phrase *John*. The outcome of transferring *Mary* at the edge of CompP has to be FC-Compliant, since CompP is categorially classified as functional. In order to saturate FC, *Mary* has to enjoy an Agree-dependency with CompP (more aptly, the head of CompP). As seen in (25), Agree can be established between these two phrases in terms of their δ-features, since *Mary* is moved to the position where it enables its unvalued δ-feature to find out the valued counterpart on the head of CompP within its c-commanding domain. This feature matching assimilates the value of the unvalued δ-feature on *Mary* into the value of the valued counterpart on the head of ComP. In terms of the present approach, this Agree-dependency is expressed as a functional interpretation on the conceptual-intentional interface, relative to the value of the δ-features involved. When the δ-feature on the head of CompP is valued as *B*, the Agree-dependency translates into a focus-background relation, where *Mary* is focalized in the discourse-oriented sense as a result of the complementation to CompP as a background. When the same feature is valued as *C*, the Agree-dependency translates into a topic-comment relation, where *Mary* is topicalized in the discourse-oriented sense as a result of the complementation to CompP as a comment. Whichever way *Mary* is interpreted saturates FC. Consequently, it can be concluded that the derivation in (25) is eligible to converge. The present approach is thus compatible with the fact that English sanctions the linear order of an object phrase preceding over a subject phrase when the former is functionally interpreted as focus or as topic.

**4.3.2. The case of Japanese**

The next task to be tackled is to explain why Japanese allows the derivation in (23) (repeated below) to converge.

(23) $[_{\pi}$ DP2, $[_{TP}$ DP1{φ:*val*; Case:*Nom*}, $[_{TP'}$ T{φ:*val*; Case:*Nom*}, $[_{vP}$ (DP1), $[_{vP'}$ $v_L$, $[_{VP}$ ... (DP2) ... ]]]]]]

In (23), once again, in order to fix the label of π in an unambiguous way, Transfer should apply to DP2 at the edge of TP. In terms of the parameter setting in (16), TP in Japanese is categorially classified as lexical. In (23) for Japanese, therefore, DP2 should be semantically interpreted at the edge of TP in compliance with LC. To attain LC-Compliance, DP2 has to bear any of the lexical relations with TP. Under the present approach, theta-theoretic, modification, and predication relations and only these are regarded as lexical relations. Here, we are immediately faced with a worrisome matter: which of these lexical relations links DP2 with TP semantically? In the sample derivation of (23), DP2 is treated as an internal argument, and hence a modification relation is not appropriate to its lexical relation with TP. Moreover, TP is not an argument-taker in the theta-theoretic sense, and hence, though DP2 itself is an internal argument, a theta-theoretic relation is rejected as inappropriate between DP2 and TP. The last candidate is a predication relation. Unfortunately, however, to the best of my knowledge, there is no evidence showing that an object phrase in the sentence-initial position is interpreted so as to be predicated by the rest of the sentence. If any of these lexical relations were not to be available between DP2 and TP in (23), the former could not stand legible at the edge of the latter. This means that, in (23) for Japanese, LC is inevitably violated if DP2 is transferred at the edge of TP. This violation crashes the derivation in (23). Even though this instance of Transfer is not invocated, the result is no different. This is because, if Transfer does not apply to DP2 at the edge of TP, γ will never receive an unambiguous label, thereby likewise crashing the derivation in (23). Accordingly, if no lexical relation holds between DP2 and TP in (23), the derivation therein

for Japanese falls into a dilemma much like for English: if DP2 is transferred at the edge of TP, LC will be violated, on one hand, while, if not, γ will be left unlabelable throughout. This dilemma dooms (23) to fail to converge even for Japanese, putting us far from explaining why this language allows for scrambling.

To disperse this worrisome matter, I would like to put forth the following condition:

(26) Given that an SO α is discontinuous such that it originally occurs as the sister of β and finally as the sister of γ, where β and γ are both lexical, then any lexical relation borne by α with β is available between α and γ, iff γ is an extended lexical projection of β.

This condition builds itself on the notion of extended lexical projection. The insight behind this notion was originally ventilated by Grimshaw (1991) (see also Grimshaw 2000), who suggested that several distinct projections be strung out in terms of their categorial properties in common (cf. also Keine 2016, 2019, 2020; van Riemsdijk 1998, 1999; Williams 2009).[5] Essentially following this insight, I define the notion of extended lexical projection as in (27):

(27) α is an extended lexical projection of β, iff:
   i. α and β are both lexical,
   ii. α dominates β, and
   ii. there is no functional category γ such that γ is dominated by α and dominates β.

In what follows, I dub the condition in (26) *Conservation of Lexical Interpretation within a Discontinuous Object* just for the mnemonic purpose, abbreviating it as CLIDO.

To see how CLIDO works, let us consider the following two toy configurations:

(28) a. $[_{\alpha}$ XP, $[_{ZP}$ $Z_L$, $[_{YP}$ $Y_L$ <XP>]]]
   b. $[_{\alpha}$ XP, $[_{ZP}$ $Z_L$, $[_{WP}$ $W_F$, $[_{YP}$ $Y_L$ <XP>]]]]

In both (28a) and (28b), XP moves to the edge of the lexical category ZP from the inside of another lexical category YP, and thereby its two occurrences are each distributed at these two distinct positions. This movement also breeds α, the SO which is not unambiguously labelable as it stands. In order to fix the label of α, XP should be transferred at the edge of YP: otherwise, α would be left unlabelable, hence crashing the derivation. This instance of Transfer has to attain LC-Compliance as its outcome, since ZP is a lexical category. In all these

[5] Grimshaw originally proposed the following condition:

(i) X is a head of YP, and YP is a projection of X iff:
   a. YP dominates X
   b. YP and X shares all categorial features
   c. All nodes intervening X and Y share all categorial features, (where a node X intervenes between X and YP if YP dominates X and N, N dominates X, N does not dominate YP.),
   d. No node intervening between X and YP is lexical.

(Grimshaw 2000: 117, typos remain original)

respects, (28a) and (28b) are no different. The only difference between these two lies in the presence or absence of the functional category WP between the two lexical projections ZP and YP. According to the definition in (27), it follows from this difference that ZP in (28a), but not in (28b), can be treated as an extended lexical projection of YP. This, in turn, means in terms of CLIDO that, given that the lower occurrence of XP is lexically related with Y, this lexical relation is available between the higher occurrence of XP and ZP in (28a), but not in (28b). Therefore, in (28a), when XP is transferred at the edge of ZP, it can be lexically interpreted at this position by means of accessing its lexical relation with Y. As a result, this instance of Transfer can attain LC-Compliance as its outcome, even though XP is not semantically related with ZP per se. The same does not hold in (28b). Therein, the higher occurrence of XP cannot access any lexical relation available between the lower occurrence of XP and Y, since ZP is not an extended lexical projection of YP. Put differently, the intervention of the functional projection WP between ZP and YP blocks the access by the higher occurrence of XP to any lexical relation borne by the lower occurrence of XP with Y. In (28b), therefore, when XP is transferred at the edge of ZP, it should be lexically related with ZP per se. Otherwise, this instance of Transfer could never attain LC-Compliance as its outcome, and hence the derivation would fail to converge. If the derivation withholds itself from transferring XP at the edge of ZP, it will avert the situation just stated. However, if so, α will remain unlabelable throughout, thereby crashing the derivation. Anyway, in (28b), unless some lexical relation holds appropriately between XP and ZP, the derivation is doomed to fail to converge.

Armed with CLIDO, let us consider (23) for Japanese once again. At first, according to the definition in (27), TP in Japanese is qualified as an extended lexical projection of another lexical category which it dominates, to the extent that no functional category intervenes between these two. This is because Japanese categorially specifies T as lexical in terms of the parameter setting in (16). Therefore, for Japanese, TP in (23) can be treated as an extended lexical projection of V, since TP and V are both lexical, the former dominates the latter, and there is no functional category such that it is dominated by the former and dominates the latter. Note that vP in (23) is another extended lexical projection of V, because treating it as such meets the definition in (27). Then, in (23) for Japanese, the spine of extended lexical projection is stretched out from V to TP. In terms of CLIDO, this spine of extended lexical projection makes it possible for the higher occurrence of DP2 at the edge of TP to access any lexical relation available between the lower occurrence of DP2 and V. As we are now treating DP2 as an internal argument of V, CLIDO allows this theta-theoretic relation to be accessed by the higher occurrence of DP2 at the edge of TP. This guarantees that, in (23) for Japanese, DP2 can be lexically interpreted at the edge of TP. Namely, by virtue of CLIDO, LC-Compliance is ensured with respect to the outcome of transferring DP2 at the edge of TP in (23) for Japanese, despite the lack of any direct semantic relation between DP2 and TP per se. Consequently, it follows that the derivation in (23) for Japanese is eligible to converge. This disperses the worrisome matter at issue. Namely, the success in convergence of the derivation in (23) for Japanese ensures the licitness of the linear order of an object phrase preceding a subject phrase in this language. The present approach thus fares well to deduce the principled reason why Japanese tolerates such preceding relation.

In sum, the present approach attributes the feasibility of scrambling in Japanese to the conspiracy of the categorial specification of T as lexical in this language and CLIDO in (26) (based on the definition of extended lexical projection in (27)). As TP in Japanese is lexical, Transfer at the edge of this phrase, if it happens, has to saturate LC as its outcome. Therefore, if an object phrase moves to that site from the inside of VP and undergoes Transfer

thereat, it has to be lexically related with TP via theta-theoretic, modification, or predication relation. As noted above, none of these lexical relations is appropriate between those two phrases. Here, CLIDO figures into the picture. By virtue of this condition, the object phrase, though not semantically related with TP per se, can be lexically interpreted at the edge of this *lexical* category by accessing the lexical relation borne by its occurrence in its original position with V (viz. its theta-theoretic relation with V). This is because the definition in (27) qualifies Japanese TP as an extended lexical projection of a lexical category which it dominates, unless no functional category intervenes between these two. Namely, to the extent that this caveat is met, Japanese is allowed to stretch the spine of extended lexical projection to TP from the lexical head V that it dominates, and to make use of this spine of extended lexical projection to allow the object phrase having moved to the edge of TP to access the lexical relation borne by its occurrence left behind in its original position with V. Therefore, when transferred at the edge of TP, the object phrase can be semantically interpreted at that site in compliance with LC, even though it is not semantically related with TP per se. In addition, as the effect of Transfer, the object phrase is rendered inaccessible to further syntactic operations, and hence does not participate in Labeling for the complex SO consisting of it and TP (viz., π in (23)). This SO, eventually, is labeled as T(P), since the head of TP counts as the nearest to it among the accessible ones within it. In this way, the legibility of an object phrase having been moved to and transferred at the edge of TP and the labelabilty of the SO composed of the object phrase and TP are both ensured in Japanese thanks to the categorial specification of T in this language as lexical. This boils down into an explanation for why Japanese allows for scrambling. When an object phrase moves to the edge of TP and undergoes Transfer thereat, it phonologically realizes at that site. This feeds the linear order of the object phrase preceding the subject phrase. The feasibility of scrambling in Japanese thus follows.

It is not the case that Japanese can scramble only arguments like DP2 in (23). In this language, modifiers including adverbial phrases and postpositional phrases can be scrambled as well. In addition, Japanese can scramble more than one such phrases. (29) and (30) below exemplify several attestable patterns of scrambling in Japanese, where scrambled phrases are bolded for emphasis:

(29) a. John-ga hon-o Mary-ni age-ta. (baseline sentence)
John-NOM book-ACC Mary-DAT give-past
'John gave a book to Mary.'
b. **Hon-o Mary-ni** John-ga age-ta. / **Mary-ni hon-o** John-ga age-ta.
(two internal arguments scrambled)

(30) a. Mary-ga kooen-de tanosiku ason-da. (baseline sentence)
Mary-NOM park-in enjoyably play-PAST
'Mary enjoyed playing in the park.'
b. **Tanosiku** Mary-ga kooen-de ason-da. (only adverbial phrase scrambled)
c. **Kooen-de** Mary-ga tanosiku ason-da. (only postpositional phrase scrambled)
d. **Tanosiku kooen-de** Mary-ga ason-da. / **Kooen-de tanosiku** Mary-ga ason-da.
(both adverbial phrase and postpositional phrase scrambled)

(29) shows that more than one internal arguments can be scrambled: in (29b), the two internal arguments (*hon*

and *Mary*) of the verb *age*(*-ta*) precede the subject phrase *John*. (30) shows that adverbial phrases and postpositional phrases can be scrambled as well: in (30b) the adverbial phrase *tanosiku*, which denotes the manner in which Mary played, is placed at the initial position of the sentence, in (30b) the postpositional phrase *kooen-de*, which denotes the location in which Mary played, is placed at the same position, and in (30c) both precede the subject phrase. Note in passing that Japanese does not necessarily focalize or topicalize the left-most scrambled phrase (*hon* or *Mary* in (29b), *tanosiku* in (30b)/(30d), and *kooen* in (30c)/(30d)) in a sentence. All these patterns, but not limited to, are utterly grammatical. This liberty of scrambling is the most salient characteristic of Japanese.

The grammaticality of the scrambled patterns attested in (29) and (30) falls into place under the present approach. Let us start with (30b) and (30c). For the reader who keeps up with what has so far been explicated in this section, why these sentences are grammatical is easy to understand. Consider:

(31) $[_{\pi}$ XP, $[_{TP}$ //DP1{φ:*val*; Case:*Nom*}//, $[_{TP'}$ T{φ:*val*; Case:*Nom*}, $[_{vP}$ <DP1>, $[_{vP'}$ $v_L$, $[_{VP}$ ... <XP> ... ]]]]]]

(31) is much the same as (23) except for XP, which replaces DP2 in (23). XP in (31) is meant to behave either as an adverbial phrase or as a postpositional phrase, in whichever case XP modifies VP. As shown, XP moves from the inside of VP to the edge of TP. This movement thus distributes two occurrences of XP each at those sites, and also breeds π, the complex SO which is not unambiguously labelable as it is. The unambiguous determination of π's label requires XP to undergo Transfer at the edge of TP. This instance of Transfer phonologically realizes XP thereat, hence feeding the liner order of XP preceding DP1: in the case that XP is an adverbial phrase, the linear order is as attested in (30b), and in the case that it is a postpositional phrase, the linear order is as attested in (30c). The outcome of transferring XP at the edge of TP in (31) has to be LC-Compliant. And, it is so, indeed, for the same reason as the outcome of transferring DP2 at the edge of TP in (23) for Japanese. To recap the reason, owing to the parameter setting in (16) and the definition in (27), TP in Japanese is qualified as an extended lexical projection of another lexical category which it dominates, to the extent that there is no intervening functional category such that it is dominated by the former and dominates the latter. In (31), then, the spine of extended lexical projection can be stretched out from V up to TP. Through this spine of extended lexical projection, the higher occurrence of XP at the edge of TP is allowed to access the lexical relation borne by the lower occurrence of XP with VP. This owes to CLIDO. Namely, XP in (31) can be lexically interpreted at the edge of TP under its modification relation with VP, even though such relation does not hold between XP and TP per se. LC-Compliance is thus attained with respect to the outcome of transferring XP at the edge of TP in (31). This instance of Transfer then feeds the unambiguous determination of π's label by rendering the head of XP inaccessible to further syntactic operations. The head of TP, on the other hand, remains accessible. Since this head c-commands all other accessible heads within π, it determines the label of π in an unambiguous way. That is, π is labeled as T(P). In (31), accordingly, the legibility of XP at the edge of TP and the labelability of π are both ensured. The grammaticality of (30b) and (30c) is thus explained under the present approach.

Let us turn to (29b) and (30d), which exemplify multiple scrambling. To understand why these sentences are grammatical, let us assign them the following syntactic structure as a first step:

(32) $[_{\mu}$ XP1/2, $[_{\pi}$ XP2/1, $[_{TP}$ //DP1{φ:*val*; Case:*Nom*}//, $[_{TP'}$ T{φ:*val*; Case:*Nom*}, $[_{vP}$ <DP1>, $[_{vP'}$ $v_L$, $[_{VP}$ ... <XP1> ... <XP2> ...]]]]]]]

(32) assumes: (i) XP stands for either an internal argument or a modifier, which means that both XP1 and XP2 may be internal arguments or modifiers; and (ii) XP1 and XP2 undergo movement in any order, but the first XP's movement breeds π and the second one μ, where π and μ are both unlabelable as they are.[6] To rephrase (ii) in other words, movement may at first merge XP1 with TP, breeding π, and next XP2 with π, breeding μ, where XP1 and XP2 are interchangeable (i.e., XP2 may at first move to the edge of TP, and next XP1 may move to the edge of γ). After each movement has taken place, the two occurrences of XP1 and the two occurrences of XP2 are distributed as depicted. If we fare well to prove the derivation in (32) eligible to converge, we can assure that XP1 and XP2 therein can precede the subject phrase DP1 in any order regardless of whether they are internal arguments or modifies. This amounts to explaining the grammaticality of multiple scrambling in (29b) and (30d).

In order to probe the derivation in (32) eligible to converge, we need to ensure the following two: the labelability of π and μ and the legibility of XP1 and XP2 at their respective halting sites. At first, the labelability of π and μ requires Transfer to apply to the edges of TP and of π each: i.e., the higher occurrences of XP1 and XP2 should be transferred. After each instance of Transfer has been implemented, the labels of π and μ are both fixed as T(P). This is because implementing each instance of Transfer renders the head of TP the nearest to both π and μ among the accessible ones within each of them. That π is labeled as T(P) means that π receives the same categorial specification as TP, i.e., π is categorially specified as lexical. This, in turn, means two things, which are interrelated with one another. One is that each instance of Transfer has to saturate LC-Compliance as its outcome. The other is that, in terms of the definition in (27), it is possible to stretch the spine of extended lexical projection from V to π, because there is no functional category such that it is dominated by π and dominates V. CLIDO then allows any lexical relation borne by the lower occurrences of XP1 and XP2 with V(P) to be transmitted to their higher occurrences through the spine of extended lexical projection from V(P) to π. Namely, XP1 and XP2 can be semantically interpreted in compliance with LC at the positions where their higher occurrences are present. Consequently, the labelability of π and μ and the legibility of XP1 and XP2 at their respective halting sites in (32) are both ensured. This proves the derivation in (32) eligible to converge. The grammaticality of (29b) and (30d) thus follows. This account implies that Japanese allows any argument and/or any modifier of V(P) to appear to the left of a subject phrase in any order to the extent that TP is qualified as an extended lexical projection of V. The liberty of scrambling in Japanese thus falls into place under the present approach.

### 4.4. Epilogue

We have so far devoted ourselves to untangling the three problems repeated below:

I. Why doesn't movement look free, although it comes free in principle? (Tension problem)
II. Why can movement permute phrases in some languages but not in others? (Variation problem)
III. How should Labeling be able to detect the head of an SO composed of two phrases? (Problem of Projection)

Problem I or the tension problem stems from the concept of Free-Merge, which enacts that all variants of Merge

[6] Minimality or closeness for movement is not taken into account here. In addition, I do not consider the tucking-in mode of movement (cf. Richards 1997, 2001).

including movement are freely applicable. This concept raises the theoretical expectation that movement takes place freely, which obviously contradicts the empirical fact that movement is not so unrestrained. This problem is further complicated by the presence of so-called "free word-order" languages like Japanese. Therefore, to the extent that Free-Merge is viable, we are responsible for illuminating not only why movement appears to be severely restricted, but also why it can alter word-order freely in some languages but not in others (i.e., Problem II or the variation problem). In this paper, we have approached these problems by taking the best advantage of the solution to Problem III or PoP.

PoP arises whenever two phrases composes an SO: when Labeling applies to this SO, it cannot determine in an unambiguous way which of these two phrases projects up to the SO in question. Such an SO remains unlabeled throughout the derivation, then deemed as illegible on the interpretive interfaces. The presence of unlabeled SOs hence causes ungrammaticality. Movement cannot avert PoP, since it always produces a complex SO by merging the moved phrase with the root phrase. This inescapability of movement from PoP, however, merits attention to solve the tension problem, because it means that movement cannot produce a grammatical SO unless the PoP caused by it is gotten rid of. Curiously, PoP alludes us a hint to approach the tension problem.

As a way out of PoP, we proposed to exploit the "inaccessibility" effect of Transfer. This solution runs as follows: when one of the two phrases composing a complex SO is transferred, this phrase is rendered inaccessible to further syntactic operations including Labeling, hence ceasing to participate in the labeling algorithm for determining the label of that SO; the other phrase, on the other hand, remains accessible, and hence it (more aptly, its head) fixes the label of the SO in question. PoP is thus resolved through the mediation of Transfer. PoP of the sort raised by movement is likewise dissolved by asserting that the moved phrase should be transferred at its halting site (a caution is in order that the root phrase within which the moved phrase originally resides should not be transferred, because otherwise the resulting derivation could not converge due to the failure to form a licit discontinuous object of the moved phrase). It is worthy to note here that this solution concomitantly overcomes Halting Problem, a drawback plaguing Chomsky's solution of PoP based on movement.

However, if Transfer is allowed to apply to any phrase at any site, the tension problem remains intact, because, if so, movement always ends up with producing an unambiguously labelable SO with the aid of Transfer. In order to avoid this undesirable state of affair, another effect of Transfer needed to come into the picture. The effect was that an SO having been transferred is interpreted at the site where it has been transferred, which owes to the mapping function of Transfer. Moreover, we suggested that the resulting interpretation of an SO having been transferred varies with where this SO has been transferred. (It should be recalled here that semantic interpretations of the sort on which we have focused on are all relational.) According to the two robust systems, viz. the dichotomy of syntactic categories and the duality of semantic interpretation, we proposed that, when a certain phrase is transferred at the edge of a lexical category, it should be lexically interpreted, on one hand, and when transferred at the edge of a functional category, it should be functionally interpreted, on the other hand. Referring to the first interpretive requirement as Lexical Criterion (LC) and the second one as Functional Criterion (FC), then our proposal boils down to it that a certain SO, when transferred at the edge of another SO, should be semantically interpreted at that site in compliance with LC or FC, relative to the category-type of the latter SO. If this interpretive condition is not met, the SO thus transferred is regarded as illegible in the semantic component.

Therefore, even though Transfer applies to one of the two phrases composing a complex SO, it will results in crashing the derivation in spite of rendering that complex SO unambiguously labelable, if its outcome cannot comply with either LC or FC. Withdrawing this instance of Transfer from applying ends up with the same result, since the complex SO in question is to be left unlabelable. Accordingly, the full convergence of derivation demands a considerable attention to be paid to where Transfer should apply (this is exactly what the main title of this paper denotes). Notice, however, that the application of Transfer per se is unrestrained in principle like Merge. What is severely restricted is its outcome. Namely, the syntactic computation can operate Transfer freely at any site, but its outcome should be ensured legibility so as to allow the derivation to fully converge.

This further implies that a special attention also has to be paid to where movement halts, because, as just stated above, its halting site corresponds to where the moved phrase undergoes Transfer. At this point too, a notice is in order that the syntactic computation does not impose any restriction (except for minimality) on movement per se, since movement can operate freely in principle as a variant of Merge. What is restricted is the halting site of movement, to say more aptly, the outcome of transferring the moved phrase at its halting site. This is the answer which we reached as for the tension problem in this paper. In a nutshell: once a moved phrase undergoes Transfer, it terminates its movement; at its halting site, then it has to stand legible; otherwise, the derivation would crash in the semantic component.

This approach to the tension problem can be extended to the variation problem. As stated above, the outcome of Transfer varies with where it arises: when Transfer apples to the edge of a lexical category, its outcome has to comply with LC, on one hand, and when applying to the edge of a functional category, its outcome has to comply with FC, on the other hand. In order to be LC-compliant, the transferred phrase only has to bear a lexical meaning suitable under a lexical relation, such as theta-theoretic, modification, or predication relation, with the lexical category at whose edge it has been transferred. The LC-compliance can also be attained even when a phrase is transferred at the edge of a lexical category with which it cannot be semantically linked but which dominates another lexical category with which it can be. A caveat is that there should be no functional category such that it is dominated by the former lexical category and dominates the latter one. What guarantees the LC-compliance in this special case is Conservation of Lexical Interpretation within a Discontinuous Object (CLIDO), repeated below:

(26) Given that an SO α is discontinuous such that it originally occurs as the sister of β and finally as the sister of γ, where β and γ are both lexical, then any lexical relation borne by α with β is available between α and γ, iff γ is an extended lexical projection of β.

Thus, the edge of a lexical category can accommodate not only a phrase which bears a lexical relation with that lexical category but also one which bears a lexical relation with another lexical category dominated by it without any intervening functional category between the two.

In order to be FC-compliant, on the other hand, the transferred phrase has to enjoy an Agree-dependency relation with the functional category at whose edge it has been transferred, because such a syntactic relation, as we argued, is presupposed by assigning a functional interpretation to that transferred phrase. By virtue of its own nature, viz.

exclusiveness, Agree cannot relate a single phrase with more than one phrases by matching the former's unvalued feature with its valued counterparts on the latter. Therefore, in case that more than one phrases are transferred at the edge of a functional category, only one of them can saturate FC if it enjoys an Agree-dependency relation with that functional category and all the others fail to do so even though they are endowed with any feature matching the one on the functional category. Consequently, it follows that each functional category accommodates only one transferred phrase at its edge. In other words, the edge of a functional category is "closed off" once there takes place an instance of Transfer in compliance with FC. In this sense, we can term functional categories as *closed* categories, as opposed to lexical categories, which, as noted above, "open" their edge so as to accommodate more than one phrases. The latter categories, hence, can be termed as *open* categories (Fukui 1988, 1995).

This contrast in openness or closedness between lexical and functional categories constitutes a ground for yielding variation of a certain kind between languages. Supposing that a certain category in one language is lexical, it can accommodate a phrase at its edge if the latter is lexically linked with the former via a theta-theoretic, modification, or predication relation, or, alternatively, if the latter is so with another lexical category which the former dominates in the absence of any intervening functional category. The number of phrases at the edge of a lexical category is in principle unlimited to the extent that such phrases meet these if-conditions. It is not a matter whether they are externally merged with that lexical category or moved to its edge. Given, on the other hand, that the corresponding category in another language is functional, the number of phrases at its edge is restricted only to one. This unique phrase has to enjoy an Agree-dependency relation with that functional category, and such a syntactic relation cannot be established with any other phrase by virtue of the exclusiveness of Agree. Therefore, when a phrase enters an Agree-dependency relation with a functional category and then undergoes Transfer at its edge (regardless of whether this phrase is externally merged with that functional category or moved to its edge), other phrases can no longer enter such a syntactic relation with the same functional category and hence violate FC if they undergo Transfer at its edge. In sum: in one language, where a certain category is lexical, more than one phrases are permitted to stay at the edge of that category, while in another language, where the corresponding category is functional, only one phrase is allowed to do so. Namely, the former language permits one phrase already having halted at the edge of the lexical category in question to be stacked upon by another phrase moving to the edge of the same lexical category, in contrast to the latter language, which cannot do the same thing in the corresponding functional category. Then, a sort of variation regarding word-order arrangement (aka scrambling) is predicted to arise between these two languages in that the former language tolerates such arrangement in the domain of a certain lexical category but the latter one cannot in the domain of the corresponding functional category. We have verified this prediction by examining English and Japanese as a case study. Following Fukui and Zushi, we assumed the tense head T in English as functional and the corresponding head in Japanese as lexical. If what has been stated just above is on track, it should be the case that TP in English can accommodate at most one phrase at its edge and the corresponding category in Japanese can accommodate more than one phrases at its edge. Namely, English cannot allow for word-order arrangement in the domain of TP, but Japanese can do so. This meets the fact. We can thus conclude that our approach to the variation problem is well-designed at least for English and Japanese. What remains to be pursued is whether or not every language that can tolerate scrambling specifies its tense head T as lexical. It has been reported in the literature that Hindi, Russian, and German, among others, behave like Japanese in that they can arrange word-order in the clausal domain relatively

freely (for Hindi, see Mahajan 1990; for Russian, see Abels 2007, 2009, and Müller & Sternefeld 1993; for German, see Fanselow 2004 and Grewendorf & Sabel 1999, inter alia). Does the availability of scrambling in these languages mean that their tense head T is specified as lexical much like Japanese T? If so, are there any empirical evidence in favor of the lexical status of the tense head T in these languages? As for Japanese, we referred to Fukui and Zushi, who reasoned that this language is not equipped with any expletive element like English "there" by virtue of its tense head being lexical. Alternatively, if the tense head T is not lexical in Hindi, Russian, and German, what other factors makes scrambling available in these languages? I cannot afford to delve into these issues any further, hence leaving them to future research.

As we have reviewed just above, English TP is functional, hence closed off once FC is saturated by transferring a certain phrase at its edge, while Japanese TP is lexical, hence opened so as to accommodate more than one phrases at its edge. We argued that this categorial distinction between TP in English and the corresponding category in Japanese paves a way to explain why the latter language can, but the former one cannot, tolerate scrambling. What I must confess here is that this is not the first to approach the typological difference between English and Japanese in terms of the categorial specification of the tense head T in these two languages. Fukui already proposed such an approach in the pre-minimalist era of generative grammar. In a series of his monumental work (Fukui 1986, 1988, 1995), he termed functional and lexical categories as *closed* and *open* categories, respectively, then suggesting that the tense head T (which he referred to as INFL) in English and the corresponding head in Japanese belong to the former and the latter, respectively. He thus reached the same consequence as we did in this paper. Although almost all the theoretical tools and theorems innovated in that era and adopted in his approach have been abolished in the course of the transition from the pre-minimalist to the present-day minimalist model of generative grammar, his core insight that the dichotomy of syntactic categories plays an important role in approaching the parametric variation between languages is carried over to our approach in this paper without fading.

This line of approaching cross-linguistic variation is rooted in what is called Lexical Parametrization Hypothesis. The germ of this hypothesis was sprouted by Borer (1984) and matured by subsequent proponents including Fukui (1988, 1995), Manzini and Wexler (1987), Zushi (2003, 2005), and so on. The name of this hypothesis is after Manzini and Wexler, who enacted that "[v]alues of a parameter are associated not with particular grammars but with particular lexical items" (Manzini and Wexler 1987: 424). This hypothesis was very appealing to many linguists in that it threw a new light on language acquisition. Namely, it parallels the burden of fixing the values of parameters of a given language with that of acquiring lexical items of that language. Fukui further advanced this parametric approach to language variation, by asserting that only functional categories are subject to parametric variation. Under this more restrictive hypothesis, which he dubbed Functional Parametrization Hypothesis, the syntactic characteristics of a given language are determined by the properties of functional categories of the language. Our approach is faithful to this restrictive version of the Lexical Parametrization Hypothesis: we did not associate values of a parameter with grammatical operations such as Merge and Transfer, keeping them invariable across languages; we postulated (16) as a categorial parameter for the tense head T. In this sense, the enterprise in which we have so far engaged in this paper can be viewed as an attempt to revive the parametric approach just overviewed within the present-day minimalist framework of generative grammar.

Our enterprise does not end here. In the next section, we will examine the empirical predictions of the present approach in this paper. Doing so, I believe, will provide a further corroboration for the empirical adequacy of our approach. In particular, it will be demonstrated that our approach is also well-designed to explain why so-called improper movement is illicit.

## 5. Proof of Predictions

We noted in Section 3.4 that our approach has at least the following two predictions: (i) when a moved phrase halts at the edge of a functional category with which it does not enter Agree, the sentence should be ungrammatical, and (ii) when a moved phrase halts at the edge of a lexical category with which it is not lexically related, the sentence should be so, too. If both predictions are empirically borne out, the adequacy of our approach will be apagogically endorsed on the empirical ground. In this section, I will demonstrate that both are in fact empirically borne out. We will start with the first prediction, and then turn to the second one, which in due course we will find pertains to the illicitness of improper movement.

### 5.1. The first prediction

To verify the first prediction, I present (34) and (35) below:

(34) a. Who thinks [$_{CP}$ that Mary bought what]?
b. *Who thinks [$_{\alpha}$ what [$_{CP}$ that Mary bought]]?

(Bošković 2007: 592, 619)

(35) a. There seems [$_{TP}$ to be a man in the garden]
b. *There seems [$_{\beta}$ a man [$_{TP}$ to be in the garden]]

(Bošković 2007: 600)

The (a)-sentences in (34) and (35) are baseline. The (b)-sentences therein are then derived from these sentences: in (34b) the embedded wh-phrase *what* appears to have moved to the edge of the embedded CP, and in (35b) the embedded subject phrase *a man* appears to have moved to the edge of the embedded TP. These sentences, however, are both ungrammatical. Let us here suppose for convenience that the two instances of movement in the (b)-sentences deliver the complex SOs α and β, respectively, as depicted: α in (34b) is composed of *what* and the embedded CP, and β in (35b) is composed of *a man* and the embedded TP. These two SOs are both unlabelable as they are. To render them labelable, Transfer has to apply to *what* and *a man* at their respective halting sites. It should be noted here that the syntactic contexts surrounding these halting sites share two points in common. For one thing, the embedded CP and the embedded TP are both functional. Therefore, our approach dictates that Transfer at the edges of these functional categories should attain FC-Compliance as its outcome. For the other thing, the heads of these functional categories are assumed to bear no formal feature for inducing Agree. This assumption is surely uncontroversial. The complementizer *that* in (34) is not equipped with any such feature concerning the formation of interrogative question. In (34b), therefore, Agree cannot be established between *what* and *that*. In (35), the raising predicate *seem* takes as its complement the embedded TP headed by the infinitival *to*, which is conventionally assumed to bear no φ-feature as opposed to the finite tense head. Therefore, in (35b) too, there is no way to establish Agree between *a man* and *to*. Of importance for our concern here is that, within

the current generative model of grammatical architecture, the absence of any Agree-inducing feature on the heads of the functional categories in question is irrelevant to whether the instances of movement under consideration is licit or not. Namely, the absence of such features does not mean that *what* and *a man* cannot move to the edges of the functional categories. Recall here that movement comes free as a variant of Merge in terms of the current minimalist framework of generative grammar. What the absence of such features has to do with under our approach is the outcomes induced by Transfer at the edges of the functional categories. This absence means, as just stated, that *what* in (34b) and *a man* in (35b) do not enjoy any Agree-dependency with the embedded CP and the embedded TP, respectively, which in turn means under our approach that the former two cannot be functionally interpreted at the edges of the latter two. Therefore, if these two moved phrases are transferred at their respective halting sites, FC will be violated. This violation results in ungrammaticality. If the moved phrases do not undergo Transfer at the sites in question, the violation of FC will be obviated. However, if so, α in (34b) and β in (35b) would be left unlabelable throughout, thereby yielding ungrammaticality alike. At any rate, the (b)-sentences in (34) and (35) should be predicted to be ungrammatical, which meets the fact. This ungrammaticality thus convincingly demonstrates that, if movement halts at the edge of a functional category with which the moved phrase does not enjoy any Agree-dependency, it will inevitably crash the derivation. This proves our first prediction empirically correct.

**5.2. The second prediction and improper movement**

The story for explaining the ungrammaticality of (34b) and (35b) does not end yet, however. For, there is another scenario for the placement of the embedded interrogative *what* in (34b) and the embedded subject *a man* in (35b) relative to their neighborhoods. The scenario is: *what* and *a man* move from their original positions to the edges of the matrix VPs each, as sketched in (36) below with irrelevant details omitted.

(36) a. ... [$_{vP}$ [$_{v\text{-}V}$ think] [$_{\alpha}$ what [$_{VP}$ <[$_{V}$ think]> [$_{CP}$ <what> [$_{CP'}$ that Mary bought <what>]]]]] (= (34b))
b. ... [$_{vP}$ [$_{v\text{-}V}$ seem] [$_{\beta}$ a man [$_{VP}$ <[$_{V}$ seem]> [$_{TP}$ <a man> [$_{TP'}$ be <a man> in the garden]]]]] (= (35b))

As seen, the relative positioning of *what* in (36a) and *a man* in (36b) is very the same as what we saw in (34b) and (35b).[7] The ungrammaticality of the sentences in question should be explained under this scenario as well.

As depicted in (36), each instance of movement therein may go cyclically: in (36a) *what* may visit the edge of the embedded CP′, and in (36b) *a man* may visit the edge of the embedded TP′. This cyclic movement is reminiscent of so-called improper movement (Chomsky 1973, 1977, 1981, 1993; Fukui 1993; May 1979; inter alia), informally defined as in (37):

---

[7] This relative positioning entails that the matrix verb gets into morpho-phonological shape at the matrix v but not the matrix V. This matter concerning verbal mopho-phonology can be couched in terms of the traditional assumption that V moves to v in the syntactic component. The syntactic representation in (36) (and the following) is based on this assumption. However, head-to-head movement like this has been considered highly disputable since the advent of the minimalism in the enterprise of generative grammar. Chomsky (2001) casted serious doubts on the syntactic treatment of head-to-head movement for the following reasons: (i) it is an acyclic operation (i.e., it does not extend the derivation), and (ii) it does not have any semantic effect. He then offered an alternative treatment for this kind of operation by suggesting that the concatenation of heads should be manipulated in the morpho-phonological component (see also Boeckx and Stjepanovic´ 2001). This morpho-phonological treatment does not have arguments against it, however. See Embick and Noyer 2001, Matushansky 2006, and Zwart 2001 for instance. In this paper, I do not afford to enter this disputable terrain anymore. So I will remain agnostic as to which treatment is theoretically and empirically more adequate.

(37) Movement is improper if it goes from an A′-position to an A-position.

Following Chomsky (1993), a certain position is defined as an A-position if it is related with a lexical category, and otherwise as an A′-position. It is then uncontroversial to classify the edge of a lexical category as an A-position and the edge of a functional category as an A′-position. According to this division of A/A′-position, the final and the intermediate landing sites of cyclic movement in (36) can be identified as A-positions and A′-positions, respectively. The final landing sites are the edges of the matrix VPs, which are standard lexical categories, and the intermediate landing sites are the edges of the embedded CP′ and the embedded TP′, both of which are functional categories. Hence, the journey of each instance of cyclic movement in (36) passes through the A′-position and terminates at the A-position. The definition of (37) then decides cyclic movement of the sorts at hand as improper.

One might then be tempted to blame the ungrammaticality of (34b) and (35b) under the scenario in (36) on the illicitness of improper movement. However, I think to decide right now to do so is a hasty conclusion. For, the definition of (37) is merely a description, i.e., it does not explain why movement from an A′-position to an A-position is illicit. To explain the ungrammaticality at issue in a principled way, I think, we should seek an alternative.

Before doing so, I want to add another point germane to improper movement. As is well known, improper movement has so far been frequently related to Japanese scrambling in the traditional literature (Fukui 1993; Kawamura 2004; Sakai 1994, 1996; inter alia). The relevant examples are presented below:

(38) a. John-ga Mary-ni [Ken-ga kodomatati-o home-ta-to] it-ta.
John-NOM Mary-DAT Ken-NOM children-ACC praise-PAST-C say-PAST
'John said to Mary that Ken praised children.'
b. Kodomotati-o/-wa John-ga Mary-ni [Ken-ga home-ta-to] it-ta.
c. *John-ga kodomotatj-o/-wa Mary-ni [Ken-ga home-ta-to] it-ta.

In Japanese, scrambling can operate not only inside a single clause, but also across clauses (Mahajan 1990; Miyagawa 1997; Nemoto 1993; Saito 1992; Tada 1990; inter alia). Conventionally, the former type of scrambling is referred to as clause-internal scrambling, the sort which we have so far treated, while the latter type is referred to as long-distance scrambling, the sort exemplified in (36b), which is derived from (36a) by scrambling the embedded object phrase *kodomotati* out of the embedded clause. The scrambled/extracted phrase in (36b) may be accompanied by the topic marker *-wa*. With this marker attached to it, it tends to be interpreted as a topic, but without as a focus. This interpretive trait is indicative of it that long-distance scrambling is akin to A′-movement such as topicalization, focalization, and the like. Therefore, long-distance scrambling is also known as A′-scrambling, vis-a-vis clause-internal scrambling as A-scrambling (see the references cited above). Putting aside the A/A′-distinction of Japanese scrambling for the moment, what is of interest to us at the present is the ungrammaticality of (38c). Therein, the scrambled/extracted phrase *kodomotati* breaks off its journey of long-distance scrambling at the position between the matrix subject phrase *John* and the matrix indirect object phrase

*Mary*. Its positioning relative to these two neighbors indicates that it lingers within the matrix VP, as schematized below (with irrelevant details omitted):

(39) [$_{TP}$ John-ga [$_{TP}$ [$_{vP}$ <John> [$_{vP'}$ [$_{\alpha}$ kodomotati-o [$_{VP}$ Mary-ni [$_{VP'}$ [$_{CP}$ <kodomotati-o> [$_{CP'}$ Ken-ga <kodomotati-o> home-ta-to]] <[$_{V}$ iu]>]]] <v>]] [$_{T\text{-}v\text{-}V}$ it-ta]]]

Provided that long-distance scrambling is implemented in a cyclic fashion, it can be thought that, as depicted in (39), the scrambled/extracted phrase *kodomotati* in (38c) stops by the edge of the embedded CP′ on the way to the matrix VP. If so, we can decide long-distance scrambling in (38c) as improper movement, because it passes through the edge of the embedded CP′ as an A′-position and then halts at the edge of the matrix VP as an A-position. This might tempt one to blame the ungrammaticality of (38c) on the illicitness of improper movement. However, once again, to do so is premature. Repeatedly, the definition in (37) is not an explanation but rather an explanandum. That is, the illicitness of movement from an A′-position to an A-position is a theory-internal matter to be explained.

It is obviously desirable to seek an uniform explanation for both the ungrammaticality of (34b) and (35b) under the derivational scenario in (36) and the ungrammaticality of (38c), since all these sentences involve improper movement in common. However, we have to do so without saying that the sentences are ungrammatical merely because improper movement is illicit. I will show from now on that this ungrammaticality is blamed on the illicitness of the site where the moved phrases in question are transferred, but not of the movement involved per se. To this end, I will demonstrate that the ungrammaticality at hand constitutes a piece of empirical evidence showing that our second prediction is correct.

To begin with, from (36) and (39) we can abstract a common syntactic configuration as below:

(40) [$_{\alpha}$ XP, [$_{ZP}$ $Z_L$, [$_{WP}$ $W_F$, ... [$_{YP}$ $Y_L$ <XP>] ... ]]]

In (40), XP is meant to stand for the moved phrases in the sentences, viz., *what* in (36a), *a man* in (36b), and *kodomotati* in (39). It has two occurrences of itself, which each are distributed at the site inside YP, on one hand, and at the edge of ZP, on the other hand. YP and ZP are both lexical, as indicated by the subscript L attached to their heads. YP corresponds to the embedded VPs in (36) and (39), within which the moved phrases are base-generated. ZP corresponds to the matrix VPs in the sentences, at whose edges the moved phrases halt. Between these two lexical projections intervenes WP, which is a functional category as indicated by the subscript F attached to its head. This node is dominated by ZP and dominates YP. It corresponds to the embedded CPs in (36a) and (39), on one hand, and to the embedded TP in (36b), on the other hand.

The configuration in (40), in fact, is nearly analogous to (28b), which we discussed in Section 4.3 and found ineligible to converge. Remember that the derivation in (28b) crashes due to the illicitness of the landing site of the moved phrase therein. The very same can be said to (40). Therein, the moved phrase XP halts at the edge of ZP as a lexical category. XP and ZP compose the complex SO α, which is unlabelable as it stands. To fix the label of α in an unambiguous way, XP has to undergo Transfer. This operation makes XP inaccessible to further

syntactic operations, hence preventing it from participating in Labeling for α. ZP, on the other hand, is still accessible, hence its head and only this head determines the label of α in terms of c-commanding any other accessible heads within α. α is then labeled as Z(P). The outcome of transferring XP at the edge of ZP should attain LC-Compliance, since ZP is a lexical category. This requires XP to be lexically interpreted at that site. If this requirement is met, the derivation in (40) will successfully end in converging: not only can α be ensured labelability, but also XP can stand legible at the site where it is transferred. However, it is desperately hopeless to obtain this result in the sentences under discussion. The moved phrase in each sentence is interpreted as the internal argument of the embedded VP, but not of the matrix VP. To rephrase it in terms of the abstract configuration in (40), XP is lexically related with YP, but not ZP. This means that XP cannot be semantically interpreted at the edge of ZP in compliance with LC. Namely, if it is transferred at that site, LC is inevitably violated, which causes the derivation to crash. Even though XP is brutely forced to be lexically related with ZP, only a bizarre interpretation obtains (for example, *kodomotati* in (38c) is absurdly interpreted as the internal argument of the matrix verb *iu* in spite of its natural interpretation as the internal argument of the embedded verb *homeru*). This unwanted situation might be obviated if XP is withheld from the target of Transfer. However, this option yields another unwanted situation, i.e., α is left unlabelable throughout the derivation. In the end, (40) is not eligible to converge, insofar as XP is lexically related with YP, or unless it is so with ZP. This leads us to not only a principled but also a unified explanation for the ungrammaticality of (34b) and (35b) under the derivational scenario in (36) and the ungrammaticality of (38c). The ungrammaticality of these sentences thus convincingly demonstrates that a sentence is ungrammatical with involving movement such that it halts at the edge of a lexical category with which the moved phrase is not lexically related. This proves our second prediction empirically correct.

It should be stressed here that our explanation for the ungrammaticality in question does not hinge on the illicitness of improper movement. Rather, it illuminates why improper movement is illicit. As defined in (37), movement is improper when it starts from an A-position, crosses over an A´-position, and eventually halts at another A-position. This route of improper movement is exactly as schematized in (40). In general, the launching site of movement is always an A-position. Moreover, as is also generally accepted, the lexical interpretation of a certain phrase can never be duplicated. Namely, once a certain phrase receives a lexical interpretation at an A-position, it cannot receive another one at any other A-position. However, in terms of our approach, when a moved phrase halts at an A-position, it has to be lexically interpreted at that site: otherwise, the derivation would violate LC. Accordingly, our approach deduces that improper movement always incurs either the duplication of lexical interpretation or the violation of LC. The illicitness of improper movement is thus explained in a principled way under our approach.

Note that CLIDO does not save the derivation in (40) from violating LC. We have already noted the reason for this on the way of simulating (28b). To recap it by using (40) above because of its configurational similarity to (28b), CLIDO does not allow (the higher occurrence of) XP at the edge of ZP to access a lexical relation borne by (the lower occurrence of) XP with YP, because ZP does not count as an extended lexical projection of YP due to the intervention of the functional projection WP. Therefore, if XP is transferred at the edge of ZP, LC is inevitably violated unless XP is lexically related with ZP. Moreover, even if such relation is established anyhow, the result will be the duplication of XP's lexical interpretation, because XP already bears a lexical relation with

YP and receives a lexical interpretation based on this relation. The duplication of the lexical interpretation of a single phrase, as stated above, should be in principle ruled out. CLIDO is thus helpless to obviate the violation of LC in (40).

Notably also, our account of the illicitness of improper movement has no commitment to assert that this type of movement always comes in a successive-cyclic fashion. Namely, we can explain the ungrammaticality of the sentences in question even though each "offending" movement therein goes from the inside of the embedded VP to the edge of the matrix VP in one fell swoop without visiting the edge of the embedded CP′ or the embedded TP′ just as abstractly schematized in (40) above. Crucial to our account is that movement form an A-position to another one always becomes "improper", i.e., offends against LC-Compliance, if it crosses over at least one boundary of a functional projection, rather than visiting at least one edge of a functional category. Therefore, our account can be maintained regardless of whether improper movement is assumed to apply successive-cyclically or in one fell swoop.

### 5.3. Interim summary

In conclusion, it turns out that our two predictions are both empirically borne out. The first one is that a sentence should be ungrammatical if it involves movement halting at the edge of a functional category with which the moved phrase does not enjoy an Agree-dependency. We proved this prediction correct by considering the ungrammaticality of (34b) and (35b). In these sentences, the moved phrases (*what* in (34b) and *a man* in (35b)) halt at the embedded CP and the embedded TP, respectively, but they do not enjoy any Agree-dependency with these two functional categories. We further demonstrated that the ungrammaticality of (34b) and (35b) under another scenario leads us to verify our second prediction, viz., that a sentence should be ungrammatical if it involves movement halting at the edge of a lexical category with which the moved phrase bears no lexical relation. The scenario supposes that the moved phrases in the sentences are extracted out of the embedded CP and the embedded TP, respectively, and eventually placed at the edges of the matrix VPs each. The latter sites are where the moved phrases cannot not receive any lexical interpretation. The ungrammaticality of the Japanese sentence in (38c) was likewise shown to constitute a piece of empirical evidence showing that our second prediction is correct.

The English sentences in (34b) and (35b) under the said scenario and the Japanese sentence in (38c) all involve improper movement, which traditionally has been argued to be literally illicit. The reason for the illicitness of improper movement, however, has ever in the literature received a mere descriptive account but not a principled explanation. We succeeded to illuminate the reason in a principled way by considering the sentences in question. Namely, the proof of our second prediction amounts to a principled explanation for the illicitness of improper movement. This consequence, I think, deserves considerable attention. For one thing, our explanation does not hinge on any ad hoc assumption. For the other thing, it is highly amenable to the current minimalist model of grammatical architecture. Recall that this model presumes that movement is unrestrained as a variant of Merge. Namely, we can explain the illicitness of improper movement without prohibiting this sort of movement per se. In these respects, I hope, the consequence we obtained will significantly contribute the development of the current minimalist theorizing.

### 5.4. Additional predictions

The proof of our second prediction or our explanation for the illicitness of improper movement yields at least two additional predictions. The first one is that, once movement comes across the boundary of a functional category, its halting site must be an A′-position, i.e., the edge of another functional category. (41i) below schematizes this. A caveat is in order that at this halting site the moved phrase must be functionally interpreted. The second one is that, if (28b) or (40) above is truncated into (41ii) below by pruning off WP, this truncated configuration should be eligible to converge. This is due to CLIDO, which we have already simulated by using (28a) in Section 4.3.

(41) i. $[_{\alpha}$ XP, $[_{FP}$ $F_F$, $[_{ZP}$ $Z_L$, $[_{WP}$ $W_F$, ... $[_{YP}$ $Y_L$ <XP>] ... ]]]
ii. $[_{\alpha}$ XP, $[_{ZP}$ $Z_L$, ... $[_{YP}$ $Y_L$ <XP>] ... ]] (≒ (28a))

In the following subsections, we will examine these two additional predictions in this order. Specifically, in Section 5.4.1. it will be shown that the first additional prediction is borne out by the fact that long-distance scrambling in Japanese always exhibits A′-properties, and in Section 5.4.2. it will be shown that a chunk of evidence proving the second additional prediction correct comes from clausal reduction in Japanese.

#### 5.4.1. The first additional prediction and long-distance scrambling in Japanese

To verify the first additional prediction, we have to return to the A/ A′-distinction of Japanese scrambling. As stated before, Japanese scrambling is divided into two sub-types, one being A-scrambling and the other A′-scrambling. Long-distance scrambling in Japanese is said to fall under the latter type. The reason for this is because this sort of scrambling always exhibits A′-properties like wh-movement, focalization, and topicalization, all of which are grouped under A′-movement. This sharply contrasts with clause-internal scrambling in Japanese, which is known to exhibit A-properties (but not always). Several tests have so far been devised to discern the A/A′-division of movement. Among them, the binding test is popular. As is well known, A-movement, but not A′-movement, feeds anaphoric binding.[8] When this test applies to clause-internal vs. long-distance scrambling in Japanese, its result is as shown below:

(42) *Clause-internal scrambling*
a. **[Otagai-no$_i$ yuuzin]-ga karera-o$_i$ home-ta.*
each.other-GEN friend-NOM they-ACC praise-PAST
'*Each other's friends praised them.'
b. *Karera-o$_i$ [otagai-no$_i$ yuuzin]-ga home-ta.*

(43) *Long-distance scrambling*
a. **[Otagai-no$_i$ yuuzin]-ga [Mary-ga karera-o$_i$ home-ta-to] it-ta.*
each.other-GEN friend-NOM Mary-NOM they-ACC praise-PAST-C say-PAST
'*Each other's friends said that Mary praised them.'
b. **Karera-o$_i$ [otagai-no$_i$ yuuzin]-ga [Mary-ga home-ta to] it-ta.*

(42) and (43) exemplify clause-internal scrambling and long-distance scrambling, respectively. In (42), the

[8] However, I will remain agnostic as to why this is so, or why A-movement exhibits distinct properties from A′-movement.

subject phrase contains the reciprocal anaphor *otagai* within it, and the object phrase is coreferential with *otagai*. (42a) is ungrammatical simply because the reciprocal anaphor is not properly bound by the object phrase coreferential with it. This ungrammaticality is obviated by scrambling the object phrase over the subject phrase as in (42b). In (43), analogously to (42), the reciprocal anaphor *otagai* is contained within the subject phrase in the matrix clause. However, the object phrase coreferential with *otagai* is generated within the embedded finite clause. Although (43a) is ungrammatical for the same reason as (42a), it cannot be salvaged from the ungrammaticality by means of scrambling the object phrase over the subject phrase in in the matrix clause as in (43b). This result of binding test indicates that clause-internal scrambling is A-movement (but not always) and long-distance scrambling is A´-movement (and exclusively so).

The A´-status of long-distance scrambling in Japanese implies two things. For one thing, the halting site of this type of scrambling is syntactically distinct from that of clause-internal scrambling as A-movement, despite the superficial identification of these two sites (see (42b) and (43b) for instance, where the scrambled phrases are both placed at the sentence-initial positions). For the other thing, the former site is syntactically identical to the halting site of other sorts of A´-movement like wh-movement, focalization, and topicalization. These two things, taken together, convincingly suggest that long-distance scrambling in Japanese halts at the edge of CompP but not of TP. Note that CompP is treated as a functional category in the standard theory. Hence, the fact that long-distance scrambling in Japanese exclusively exhibits A´-properties meets our first additional prediction. In addition to this, it is expected under our approach that this type of scrambling always requires its target to be functionally interpreted. Recall that movement halts just when Transfer applies to the moved phrase, and that, if its halting site is the edge of a functional category, this instance of Transfer has to attain FC-Compliance as its outcome. In fact, we have already witnessed in (38b) that long-distance scrambling therein places the embedded object phrase *kodomotati* at the sentence-initial position, where this phase receives either a focus or topic interpretation. This interpretive trait of long-distance scrambling contrasts it with clause-internal scrambling, which does not require its target to be either focalized or topicalized. This contrast, too, is exactly as predicted under our approach. As for how the target of long-distance scrambling in (38b) is semantically interpreted either as a focus or as a topic, please refer back to the discussion in Section 3.1, which I refrain from recapitulating here for space.

### 5.4.2. The second additional prediction and Clausal Reduction in Japanese

A lump of evidence which I present in this section to verify the second additional prediction concerns what Sakai (1996) dubs Clausal Reduction (henceforth abbreviated as CR). A relevant example is given below:

(44) *John-ga Mary-ni [karera-o homeru-yoo(-ni)] tanon-da/meizi-ta.*
John-NOM Mary-DAT them-ACC praise-so.that(-NI) ask-PAST/order-PAST
'John asked/ordered Mary to praise them.'

In (44), the matrix verb *tanom(u)* or *meiz(u)* denotes the act of requesting, and the complement it takes is headed by the modal-like affix *-yoo(-ni)*, expressing the requirement that the requestor asks the requestee to fulfil. This affix is glossed as "so.that(-NI)", and the parenthesis notation in the affix and its glass indicates that *-ni* may be optionally omitted. Please note, however, that this is just a terminological matter and that I have no analytical

commitment to it.

The reason why Sakai refers to CR in a construction like (44) is because such construction deceptively behaves as if it were mono-clausal, albeit being superficially bi-clausal. To discern whether a certain construction is mono-clausal or bi-clausal, the negative polarity test is useful among others. In Japanese, as is well known, the negative polarity item *-sika* 'only' is licensed by a negation marker like *-nai* iff the former is c-commanded by the latter and both are contained within one and the same clause. Observe (45) below (where the negative polarity item and the negation marker are double-underlined for emphasis):

(45) a. *John-ga karera-sika home-nakat-ta.*
John-NOM them-only praise-NEG-PAST
'John praised only them.'

b. **Mary-ga [John-ga karera-sika home-ta-to] iwa-nakat-ta.*
Mary-NOM John-NOM them-only praise-PAST-COMP say-NEG-PAST
'Mary said that John praised only them.'

c. *John-ga Mary-ni [karera-sika home-ru-yoo(-ni)] tanoma-nakat-ta/meizi-nakat-ta.*
John-NOM Mary-DAT them-only praise-NONPAST-so.that(-NI) ask-NEG-PAST/order-NEG-PAST
'John asked/ordered Mary to praise only them.'

In (45a), the negative polarity item *-sika*, which is attached to the object phrase *karera*, is (structurally) c-commanded by the negation marker *-nai*, which is attached to the verb *home*. As these two are contained in one and the same clause, the former is licensed by the latter. In (45b), on the other hand, *-sika* and its licenser *-nai* are distributed in different clauses: the former is attached to the object phrase in the embedded clause, while the latter is attached to the verb *home* in the matrix clause. Therefore, *-sika* in (45b) is not licensed, albeit being (structurally) c-commanded by *-nai*. (45b) is hence ungrammatical. (45c) exemplifies a construction where CR is assumed to be involved. Applying the negative polarity test to this construction, the result tells us that the construction is mono-clausal rather than being bi-clausal. Note that (45c) is grammatical, although *-sika* and its licenser *-nai* appear to be distributed in different clauses much as in (45b). This grammaticality can be explained if the two clauses are reduced to one. Provided that this is on track, we can maintain that *-sika* in (45c) is licensed by *-nai* , hence the grammaticality results.

The simplest way to capture the mono-clausality of (44)/(45c) or the CR effect therein is to assume that CompP as a clausal boundary does not exist between the matrix clause and the complement selected by the matrix verb of requesting. In fact, Sakai proposed to assign a sentence like (44)/(45c) the following syntactic structure, where the complement of the matrix verb of requesting is TP but not CompP:

(46) $[_{TP2}$ *John-ga* ... $[_{VP2}$ *Mary*$_i$*-ni*, $[_{VP'2}$ $[_{TP1}$ PRO$_i$ ... $[_{VP1}$ *karera-o*, $[_{V1}$ *home*$]_L]$ ... $[_{T1}$ *-yoo(-ni)*$]_L]$, $[_{V2}$ *tanomu/meizu*$]_L]$ ... $T2_L]$

In (46), for the expository purpose, any irrelevant details are omitted, and some nodes are assigned the relevant labels though they each immediately dominate two phrases. Moreover, the embedded V/VP and T/TP, on one hand, and the matrix V/VP and T/TP, on the other hand, are indexed for the sake of discriminating them visually.

Here, I want to clarify two points with regards to the inside of the embedded TP1. Firstly, the subject phrase therein is represented as PRO, which is a phonologically null pronominal element and controlled by the requestee (*Mary*) in the matrix clause. Secondly, the modal-like affix *-yoo(-ni)* is base-generated at the head of the embedded TP1. That said, these two points are mostly orthogonal to our main concern here.

Of importance to our interest at the present is that in (46) no functional projection like CompP intervenes between the embedded TP1 and the matrix VP2. This means that the spine of extended lexical projection can by definition stretch out from the embedded V1 to the matrix TP2, much as it can from Y to ZP in (41ii). Recall here that Japanese categorially specifies TP as lexical.[9] It is then predicted that in (46) the embedded object phrase *karera* can be scrambled out of the embedded TP1 and halt either at the edge of the matrix VP2 (over *Mary* in the matrix clause) or at the edge of the matrix TP2 (over the matrix subject *John*). Halting at either site is tantamount to phonologically realizing at either site through Transfer. Whichever site Transfer applies, LC has to be saturated as its outcome, since the matrix TP2 and VP2 are both lexical. For attaining the LC-Compliance, *karera* should be lexically interpreted at its halting site. That said, it is obvious that this phrase is not lexically related with either the matrix TP2 or the matrix VP2: it is interpreted as an internal argument of the embedded verb *home*, but it can never receive any such interpretation by relating with any lexical category within the matrix clause. However, *karera* can access its lexical relation borne with the embedded V1 from either the edge of the matrix TP2 or the edge of the matrix VP2 through the spine of extended lexical projection mentioned above. This solely owes to CLIDO. This ensures LC-Compliance with respect to the outcome of transferring *karera* at either edge. Accordingly, whichever edge *karera* halts at, the sentence should be grammatical. Moreover, it is predicted that, if such scrambling takes place, it should feed anaphoric binding just like clause-internal scrambling, because both edges are identified as A-positions. If these two predictions are verified, our second additional prediction proves empirically correct.

In fact, both predictions are borne out as we witness below:

(47) a. **John-ga [otagai$_i$-no yuuzin]-ni [karera$_i$-o home-ru-yoo(-ni)] tanon-da/meizi-ta.*
John-NOM each.other-GEN friend-DAT them-ACC praise-NONPAST-so.that(-NI) ask-PAST/order-PAST
'John asked/order each other's friends to praise them.'

b. *John-ga **karera$_i$-o** [otagai$_i$-no yuuzin]-ni [$t_i$ home-ru-yoo(-ni)] tanon-da/meizi-ta.*

(48) a. **[Otagai-no$_i$ yuuzin]-ga Mary-ni [karera-o$_i$ home-ru-yoo(-ni)] tanon-da/meizi-ta.*
each.other-GEN friend-NOM Mary-DAT them-ACC praise-NONPAST-so.that(-NI) ask-PAST/order-PAST
'*Each other's friends asked/ordered Mary to praise them.'

b. ***Karera-o$_i$** [otagai-no$_i$ yuuzin]-ga Mary-ni [home-ru-yoo(-ni)] tanon-da/meizi-ta.*

The (b)-sentences in (47) and (48) are derived from the (a)-sentences therein by scrambling the embedded object phrase *karera* (bolded for emphasis) out of the complement selected by the matrix verb of requesting. If this scrambled phrase halts at the edge of the matrix VP2, it appears between the matrix subject phrase and the phrase corresponding to the requestee-argument of the matrix verb of requesting, as in (47b). If, on the other hand, it

[9] Sakai concurs with this hypothesis concerning the categorial specification of T in Japanese.

halts at the edge of the matrix TP2, it appears at the initial position of the sentence as in (48b). These sentences are utterly grammatical. Moreover, these instances of scrambling feed anaphoric binding as evidenced by the licitness of the coreferential relation between *karera* and the reciprocal anaphor *otagai* in (47b) and (48b) each. In the former sentence, *otagai* is embedded inside the phrase corresponding to the requestee-argument of the matrix verb of requesting, and in the latter sentence it resides within the matrix subject phrase. In either case, *karera* is licitly coreferential with *otagai*. These facts sharply contrast with the ones concerning long-distance scrambling. As we saw in (38c), long-distance scrambling cannot break off its journey at any intermediate site within a clause higher than another one from which it launches. In addition, as evidenced by (43), this type of scrambling exclusively exhibits A′-properties (e.g., it does not feed anaphoric binding). To this end, we can conclude that the facts witnessed in (47) and (48) concerning CR constitute a chunk of evidence showing that our second additional prediction is empirically correct.

**5.5. Summary**

In this section, we have devoted ourselves to confirming the two main predictions listed below:

(49) i. If movement halts at the edge of a functional category with which the moved phrase enjoys no Agree-dependency agreement, the sentence should be ungrammatical.
ii. If movement halts at the edge of a lexical category with which the moved phrase bears no lexical relation, the sentence should be ungrammatical.

These predictions were shown to be both born out on the empirical ground. In particular, we noted that the second prediction is germane to improper movement in effect, arguing that the proof of its correctness amounts to a principled explanation for why improper movement is illicit.

Furthermore, in the course of proving (49ii) correct, we came by the two additional predictions listed below:

(50) i. Once movement comes across the boundary of a functional category, it must halt at the edge of another higher one, where the moved phrase should be functionally interpreted (see (41i)).
ii. Movement can launch from the inside of a lexical category and halt at the edge of another higher one, iff there is no functional category such that it is dominated by the former lexical category and dominates the latter one (see (41ii)).

We witnessed these two are both likewise born out. (50i) was verified by the fact that long-distance scrambling in Japanese exclusively exhibits A′-properties and that at its halting site the scrambled phrase should be interpreted either as a focus or as a topic. As for (50ii), a chunk of evidence for it obtained from the facts concerning CR in Japanese. It was observed that in a construction involving CR, scrambling can launch from a lexical category inside the complement selected by the matrix verb and halt at the edge of another lexical category inside the matrix clause. Furthermore, this sort of scrambling was found to exhibit the same property as clause-internal scrambling, exactly as predicted.

Eventually, it turned out that all the predictions above are borne out on the empirical ground. This consequence

dramatically corroborates the empirical adequacy of the current approach we have developed throughout the present paper.

**References**


Abels, Klaus. 2007. Towards a restrictive theory of (remnant) movement. In Jeroen van Craenenbroeck & Johan Rooryck (eds.), *Linguistic Variation Yearbook* 7, 53-120. Amsterdam: John Benjamins.

Abels, Klaus. 2009. Some implications of improper movement for cartography. In Jeroen van Craenenbroeck (ed.), *Alternatives to cartography*, 325-359. Berlin: de Gruyter.

Akiyama, Masahiro. 2004. Multiple nominative constructions in Japanese and Economy. *Linguistic Inquiry* 35(4): 671-683.

Akiyama, Masahiro. 2005. On the general tendency to minimize moved elements: the multiple nominative construction in Japanese and its theoretical implications. *The Linguistic Review* 22(1): 1-68.

Bauke, Leah S. & Tom Roeper. 2017. How unlabelled nodes work: Morphological derivations and the subcomponents of UG operations. In Leah Bauke & Andreas Blumel (eds.), *Labels and roots*, 233-262. Berlin: Mouton de Gruyter.

Biskup, Petr. 2017. Labeling and other syntactic operations. In Leah Bauke & Andreas Blumel (eds.), *Labels and roots*, 91-116. Berlin: Mouton de Gruyter.

Blümel, Andreas. 2017. Exocentric root declaratives: Evidence from V2. In Leah Bauke & Andreas Blumel (eds.), *Labels and roots*, 263-289. Berlin: Mouton de Gruyter.

Boeckx, Cedric. 2008. *Bare syntax*. Oxford: Oxford University Press.

Boeckx, Cedric & Sandra Stjepanović. 2001. Head-ing toward PF. *Linguistic Inquiry* 32(2): 345-355.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38(4): 589-644.

Brentano, Franz. 1973. *Psychology from an empirical point of view*. Translated by Antos C. Rancurello, D. B. Terrell, & Linda L. McAlister from *Psychologie vom empirischen Standpunkt* (1874, 1924). Sections V-IX.

Caponigro, Ivano. 2019. In defense of what(ever) free relative clauses they dismiss: a replay to Donati and Cecchetto (2011). *Linguistic inquiry* 50(2): 356-371.

Cecchetto, Carlo & Caterina Donati. 2010. On labeling principle C and head movement. *Syntax* 13(3): 241-278.

Cecchetto, Carlo & Caterina Donati. 2015. *Relabeling*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. Bare phrase structure. In Gert Webelhuth (ed.), *Government and binding theory and the minimalist program: Principles and parameters in syntactic theory*, 383-439. Oxford: Oxford University Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by step: Essay on minimalism in honor of Howard Lasnik*, 89-155. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1-22.

Chomsky Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interface + recursion = language?: Chomsky's minimalism and the view from syntax-semantics*, 1-30. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33-49.

Chomsky, Noam. 2015. Problems of projection: extensions. In Eisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honor of Adriana Belletti*, 3-16. Amsterdam: John Benjamins.

Chomsky, Noam. 2020. Puzzles about phases. In Ludovico Franco & Paolo Lorusso (eds.), *Linguistic variation: structure and interpretation*, 163-167. Berlin: Mouton de Gruyter.

Chomsky, Noam, Ángel J. Gallego, & Dennis Ott. 2019. Generative grammar and the faculty of language: insights, questions, and challenges. *Catalan Journal of Linguistics Special Issue*: 229-261.

Collins, Chris. 2017. Merge(X, Y) = {X, Y}. In Leah Bauke & Andreas Blumel (eds.), *Labels and roots*, 91-116. Berlin: Mouton de Gruyter.

Culicover, Peter. 1970. Syntactic and semantic investigation Doctoral dissertation, MIT, Cambridge, Mass.

Davidson, Donald. 1967a. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81-120. Pittsburgh, PA: University of Pittsburgh Press.

Davidson, Donald. 1976b. Truth and meaning. *Synthese* 17(1): 304-323.

De Vos, Mark. 2007. Toward a definition of SUBJECT in binding domains and subject oriented anaphors. *Stellenbosch Papers in Linguistics* 37: 23-65.

Donati, Caterina. 2006. On wh-head movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 21-46. Cambridge, MA: MIT Press.

Donati, Caterina, & Carlo Cecchetto. 2011. Relabeling heads: a unified account of relativization structures. *Linguistic Inquiry* 42(4): 519-560.

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4): 555-595.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2014. Labeling by minimal search: Implications for successive-cyclic A-movement and the conception of the postulate 'phase'. *Linguistic Inquiry* 45(3): 463-481.

Fanselow, Gisbert. 2004. The MLC and derivational economy. In Arthur Stepanov, Gisbert Fanslow, & Ralf Vogel (eds.), *Minimality effects in syntax*, 73-123. Berlin: de Gruyter.

Fukui, Naoki. 1986. A theory of category projection and its applications. Cambridge, MA: MIT dissertation.

Fukui, Naoki. 1988. Deriving the differences between English and Japanese: a case study in parametric syntax. *English Linguistics* 5: 249-270.

Fukui, Naoki. 1993. A note on improper movement. *The Linguistic Review* 10(2): 111-126.

Fukui, Naoki. 1995. The principles-and-parameters approach: a comparative syntax of English and Japanese. In Masayoshi Shibatani & Theodra Byron (eds.), *Approaches to language typology*, 327-372. Oxford: Oxford University Press.

Fukui, Naoki. 2011. Merge and bare phrase structure. In Cedric Boeckx (ed.), The Oxford handbook of linguistic minimalism, 73-95. Oxford: Oxford University Press.

Fukui, Naoki, & Margaret Speas. 1986. Specifiers and projection. *MIT Working Papers in Linguistics 8*: 128-172.

Grewendorf, Günther & Joachim Sabel 1999. Scrambling in German and Japanese: adjunction versus multiple specifiers. *Natural Language and Linguistic Society* 17(1): 1-65.

Grimshaw, Jane. 2000. Locality and extended projection. In Peter Coopmans, Martin Everaert, & Jane Grimshaw (eds.), *Lexical specification and insertion*, 115-133. Amsterdam: John Benjamins.

Hale, Kenneth. 1980. Remarks on Japanese phrase structure: comments on the papers on Japanese syntax. *MIT Working Papers in Linguistics 2*: 185-202.

Hale, Kenneth. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1(1): 5-47.

Hiraiwa, Ken. 2000. On nominative-genitive conversion. *MIT Working Papers in Linguistics* 39: 66-123.

Hiraiwa Ken. 2001. Multiple Agree and the defective intervention constraint in Japanese. *MIT Working Papers in Linguistics* 40: 67-80.

Hoji, Hajime. 1985. Logical form constraints and configurational structures in Japanese. Seattle, Washington: University of Washington dissertation.

Jackendoff, Ray. 1969. Some rules of semantic interpretation in English. Doctoral dissertation, MIT, Cambridge, Mass.

Jenkins, Lyle. 1972. Modality in English syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Jespersen, Otto. 1924. *The philosophy of grammar*. Chicago: University of Chicago Press (1992).

Kawamura, Tomoko. 2004. A feature-checking analysis of Japanese scrambling. *Journal of Linguistics* 40(1): 45-68.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Keine, Stefan. 2016. Probes and their horizons. PhD diss, University of Massachusetts, Amherst, MA.

Keine, Stefan. 2019. Selective opacity. *Linguistic Inquiry* 50: 13-62.

Keine, Stefan. 2020. *Probes and their horizons*. Cambridge, MA.: MIT Press.

Kitagawa, Yoshihisa. 1986. Subjects in Japanese and English. Cambridge, MA: University of Massachusetts dissertation.

Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik: Linguistische Berichte Sonderheft* 4, 17-53. Opladen: Westdeutscher Verlag.

Koopman Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85(2-3): 211-258.

Kuroda, S.-Y. 1988. Whether we agree or not: a comparative syntax of English and Japanese. *Linguiticae Investigationes* 12: 2-44.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3): 335-391.

Ladusaw, William A. 1994. Thetic and categorical, stage and individual, weak and strong. *SALT* 4: 220-229.

Mahajan, Anoop Kumar. 1990. The A/A' distinction and movement theory. Cambridge, MA: MIT dissertation.

Manzini, M. Rita & Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 19: 413-444.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1): 69-109.

May, Robert. 1979. Must COMP-to-COMP movement be stipulated? *Linguistic Inquiry* 10(4): 719-725.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28(1): 1-26.

Miyagawa, Shigeru. 2001. The EPP, scrambling, and wh-in-situ. In Michael Kenstowicz (ed.), *Ken Hale: A life*

*in language*, 293-338. Cambridge, MA: MIT Press.
Miyagawa, Shigeru. 2003. A-movement scrambling and options without optionality. In Simin Karimi (ed.), *Word order and scrambling*, 177-200. Oxford: Blackwell.
Miyagawa, Shigeru. 2010. *Why agree? Why move? Unifying agreement-based and discourse configurational languages*. Cambridge, MA: MIT Press.
Miyagawa, Shigeru. 2017a. *Agreement beyond phi*. Cambridge, MA: MIT Press.
Miyagawa, Shigeru. 2017b. Topicalization. *Gengo Kenkyu* 152(1): 1-29.
Müller, Gereon & Wolfang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461-507.
Narita, Hiroki. 2012. Phase cycles in service of projection-free syntax. In Angel J. Gallego (ed.), *Phases: Developing the framework*, 125-172. Berlin: Mouton de Gruyter.
Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax*. Amsterdam: John Benjamins.
Neeleman, Ad & Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11(2): 137-189.
Nemoto, Naoko. 1993. Chains and case position: a study from scrambling in Japanese. Doctoral dissertation, University of Connecticut, Storrs.
Pietroski, Paul M. 2002. Function and concatenation. In Gerhard Preyer & Georg Peters (eds.), *Logical form and language*, 91-117. Oxford: Oxford University Press.
Richards, Norvin. 1997. What moves where when in which Language. Doctoral dissertation, MIT, Cambridge, Mass.
Richards, Norvin. 2001. *Movement in Language: Interactions and Architectures*. Oxford: Oxford University Press.
Riemsdijk, Henk van. 1990. Functional prepositions. In Harm Pinkster & Inge Genee (eds.), *Unity in diversity: Papers presented to Simon C. Dik on his 50 birthday*, 229-241. Dordrecht: Foris.
Riemsdijk, Henk van. 1998. Categorial feature magnetism: The endocentricity and the distribution of projections. *Journal of Comparative Germanic Linguistics* 2(1): 1-48.
Riemsdijk, Henk van. 2006. Grafts follow from merge. In Mara Frascarelli (ed.), *Phases in interpretation*, 15-44, Amsterdam: John Benjamins.
Rizzi, Luigi. 2015. Notes on labeling and subject positions. In Elisa Di Domenco, Cornelia Hamann, & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honor of Adriana Belletti*, 17-46. Amsterdam: John Benjamins.
Rizzi, Luigi. 2016. Labeling, the maximality and the head-phrase distinction. *The Linguistic Review* 33: 177-198.
Ross, John R. 1967. Constrains on variables in syntax. Cambridge, MA: MIT dissertation.
Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. Cambridge, MA: MIT dissertation.
Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1(1): 69-118.
Saito, Mamoru. 2016. (A) case for labeling: Labeling in languages without φ-feature agreement. *The Linguistic Review* 33(1): 129-175.
Saito, Mamoru. 2018. Kase as a weak head. *McGill Working Papers in Linguistics* 25 [Special Issue in Honor of Lisa Travis]: 382-391.
Sakai, Hiromu. 1994. Derivational economy in long distance scrambling. *MIT Working Papers in Linguistics* 24:

295-314.

Sakai, Hiromu. 1996. Clause reduction in Japanese. *MIT Working Papers in Linguistics* 29: 193-212.

Tada, Hiroaki. 1990. A/A-bar partition in derivation. Cambridge, MA: MIT dissertation.

Takita, Kensuke, Nobu Goto, and Yoshiyuki Shibata. 2016. Labeling through Spell-Out. *The Linguistic Review* 33(1): 177-198.

Ura, Hiroyuki. 2001. *Checking theory and grammatical function in Universal Grammar*. Oxford: Oxford University Press.

Uriagereka, Juan. 1999. Multiple Spell-Out. In Samuel D. Epstein & Norbert Hornstein (ed.), *Working Minimalism*, 251-282, Cambridge, MA: MIT Press.

Van Riemsdijk, Henk. 1990. Functional prepositions. In Harm Pinkster & Inge Genee (eds.), *Unity in diversity: Papers presented to Simon C. Dik on his 50 birthday*, 229-241. Dordrecht: Foris.

Vercauteren, Aleksandra. 2017. Features and labeling: Label-driven movement. In Leah Bauke & Andreas Blumel (eds.), *Labels and roots*, 69-116. Berlin: Mouton de Gruyter.

Vermeulen, Reiko. 2005. Possessive and adjunct multiple nominative constructions in Japanese. *Lingua* 115(10): 1329-1363.

Williams, Edwin. 2009. There is no alternative to cartography. In Jeroen van Craenenbroeck (ed.), *Alternatives to cartography*, 361-373. Amsterdam: Mouton de Gruyter.

Zushi, Mihoko. 2003. Null arguments: in the case of Japanese and Romance. *Lingua* 113(4-6) [Special Issue: Formal Japanese Syntax and Universal Grammar, The Past 20 Years]: 559-604.

Zushi, Mihoko. 2005. Deriving the similarities between Japanese and Italian: a case study in comparative syntax. *Lingua* 115(5): 711-752.

Zwart, C. Jan-Wouter. 2001. Syntactic and phonological verb movement. *Syntax* 4(1): 34-62.