Linguistic Research 36(3), 325-363
DOI: 10.17250/khisli.36.3.201912.001

Ways of solving (counter-)cyclic A-movement in phase theory*

Manabu Mizuguchi
(Toyo University)

Mizuguchi, Manabu. 2019. Ways of solving (counter-)cyclic A-movement in phase theory. *Linguistic Research* 36(3), 325-363. A-movement in phase syntax is executed counter-cyclically under the assumption that movement or Internal Merge applies at the phase level. This poses a non-trivial problem to syntax as the subject infixes into Spec,TP in violation of the No-Tampering Condition. This paper addresses the counter-cyclicity problem and argues that it is only apparent. I propose two ways to solve the counter-cyclicity, claiming that A-movement is made cyclic by Transfer and Merge. I also suggest that Transfer is reducible to Merge, with counter-cyclic A-movement solved by Merge alone. I show that the proposed analysis can explain subject extraction from the embedded clause and argues for vacuous movement in local subject Ā-movement. The EPP effect is reconsidered as one consequence of the proposals and I argue that it is deducible from externalization as well as from labeling. The present paper endorses the assumption that Merge is the core of syntactic derivation and supports the basic hypothesis in the Minimalist Program, which says that language is explained by Merge and interfaces with two other internal systems. **(Toyo University)**

Keywords A-movement, phase, Merge, subject extraction, local subject Ā-movement, externalization, EPP

* Portions of this paper were presented in somewhat different form at Arizona Linguistics Circle (ALC) 12 (University of Arizona), Florida Linguistics Yearly Meeting (FLYM) 2018 (Florida International University), Seoul International Conference on Generative Grammar (SICOGG) 16 and 20 (Dongguk University and Konkuk University), and Western Conference on Linguistics (WECOL) 2017 (Boise State University). I thank the audiences there for comments and questions. I also thank the anonymous reviewers of *Linguistic Research* for a number of constructive comments and helpful suggestions on earlier versions of this paper. As always, all remaining errors and inadequacies are my own. The research reported here was supported by Japan Society for the Promotion of Science (JSPS) under Grant-in-Aid for Young Scientists (B) (#16K16835) and in part under Grant-in-Aid for Scientific Research (C) (#18K00544).

1. Introduction

This paper discusses A-movement in phase syntax. Chomsky (2000) introduces phases as cyclic points of the derivation and claims that CP and *v**P form phases. Phase syntax is distinguished from previous approaches to syntax as regards cyclicity in that phases are derivational stages where phase-internal operations apply. Chomsky (2007, 2008, 2013) argues that all operations except External Merge (EM) apply simultaneously at the phase level. This assumption allows the interfaces to interact with syntax and distinguish copies from repetitions under the Inclusiveness Condition, which bans addition of new objects in the course of structure building by Merge (in particular, indices, bar levels, etc.) (Chomsky 1995: 228). To see this, consider (1) (order, which is determined by externalization, is irrelevant in (1) and elsewhere in the paper):[1]

(1) a. [The student$_4$ [seems [the student$_3$ [to be believed [the student$_2$ [to be [the student$_1$ intelligent]]]]]]]
b. [The student$_4$ [believes that [the student$_3$ [said that [the student$_2$ was likely to visit the country where [the student$_1$ [was born]]]]]]]

In (1), four occurrences of the phrase *the student* are created by Merge and syntactically, they are no different in both (1a) and (1b). However, at the interfaces, they are interpreted as one, forming a chain in (1a) while they are construed as distinct entities in (1b); in other words, they are copies in (1a) (i.e., non-distinct occurrences of the same SO forming a discontinuous object) while they are repetitions in (1b) (i.e., distinct occurrences of the same SO, each of which is independent of the other and receives semantic and morpho-phonological interpretations). The distinction will be possible under the Inclusiveness Condition if movement or Internal Merge (IM) applies at the phase

1 Externalization is the mapping of syntactic objects (SOs) to representations that can be accessed by the Sensory-Motor (SM) system, applying when SOs are transferred to the interfaces (Chomsky 2017, Chomsky et al. 2019); that is, externalization is "spell-out," to use a different term. In this paper, following the recent literature, externalization is used instead of spell-out.

I assume that the counterpart to externalization or the other side of Transfer is interpretation, the mapping of SOs to representations that can be accessed by the Conceptual-Intentional (CI) system.

level together with Transfer and the information is locally available or communicated to the interfaces (or interface components – interpretation and externalization) that the SO is internally merged (but see Chomsky 2015 for a different approach based on phase-level memory). The relevant assumption makes possible proper interpretations of (1a) and (1b) at the interfaces under the presumption that syntax (more precisely, narrow syntax) conforms to the Inclusiveness Condition (or more generally, under simplest Merge – see (5)).

Though the assumption can make the distinction between copies and repetitions possible through syntax-interface interaction at the phase level, it poses a non-trivial problem to A-movement. Consider A-movement to Spec,TP in (2):

(2) The student will read the book.

Given the relevant assumption, the subject *the student* will be internally merged and moved when the C phase is structured as in (3):[2]

(3) $[_{\kappa}$ C $[_{\lambda}$ T_{will} $[_{\alpha}$ the student $[v^*$ [read the book]]]]]

Movement to Spec,TP at the phase level will be counter-cyclic: since λ is embedded in κ, the IM does not apply to the root but the subject tucks in or infixes between CP and TP, which violates the No-Tampering Condition (NTC) (Chomsky 2008: 138):

(4) **The No-Tampering Condition**
Merge of X and Y leaves the two SOs unchanged.

As regards counter-cyclicity or tucking in, Chomsky (2008: 141) is well aware of this problem, noting that it is indeed a literal violation of NTC. He defends counter-cyclic A-movement by saying that the violation is arguably a principled one, hence consistent with the assumption of optimal computation, assuming that the Spec position is the position as close to the probe as possible. In (3), the subject is the goal of T or the ϕ-probe and Spec,TP, to which it moves, is the

2 In this paper, Greek letters are used both for set labels and for general symbols.

position as close to the T probe as possible. Notice, however, that tucking in, as Chomsky himself admits, literally violates NTC; it clearly deviates from the simplest application of Merge. It has been assumed that Merge is subject to the Strong Minimalist Thesis (SMT), operating in accord with general conditions of computational efficiency (or what are often called "3rd factor" principles), which constrain computation in general, and is formulated as simplest Merge (=(5)):

(5) Merge(α, β) = {α, β}

Given (5), Merge will conform to NTC and counter-cyclic movement is an undesirable elaboration of the operation as it does not apply to the root (see Chomsky et al. 2019 for relevant discussion).

The present paper addresses the counter-cyclicity problem raised by A-movement in the phase-based model of syntax. I propose two ways to circumvent counter-cyclicity in A-movement and make the movement cyclic, demonstrating that the counter-cyclicity is only apparent. I show that the mechanism to be employed is the structure-building operation Merge, which yields a digitally infinite array of hierarchically structured expressions or SOs and is considered indispensable in the Faculty of Language, arguing that it plays a key role in solving the counter-cyclicity. The paper endorses the argument in the literature that the operation is fundamental to syntactic derivation or to the Faculty of Language as hypothesized in Chomsky (2000) and his subsequent writings, in particular, Chomsky (2010, 2017). I also argue that the proposed analysis explains subject extraction out of the embedded clause and provides a novel analysis of the Vacuous Movement Hypothesis (i.e., the hypothesis that the subject moves to a single Spec) in local subject Ā-movement. The proposals in this paper will lead to the reconsideration of the EPP effect, which I argue is explained by externalization as well as by labeling.

This paper is organized as follows. Section 2 presents main proposals of the paper and argues that Transfer and Merge can make counter-cyclic A-movement in phases cyclic; in this section, I also suggest that Transfer is reducible to Merge. Section 3 discusses subject extraction from the embedded clause, showing that the proposed analysis can give a Merge-based explanation to it. Section 4 considers cyclic A-movement in local subject Ā-movement, arguing for a novel

analysis of the Vacuous Movement Hypothesis. Section 5 discusses the zero complementizer in embedded and root clauses and reconsiders the EPP effect. Section 6 summarizes and concludes the paper.

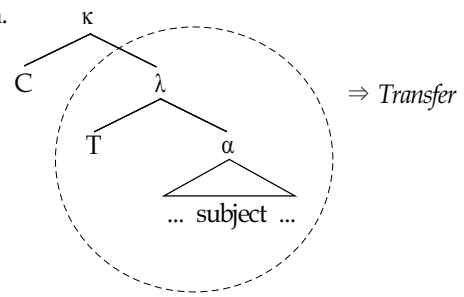
2. "Cyclic" A-movement in phase syntax

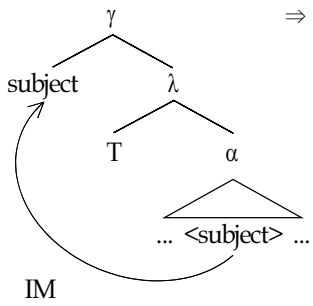
In this section, I argue that counter-cyclic A-movement can be circumvented in two ways. I discuss them one by one, showing that the counter-cyclicity is only apparent.

2.1 Cyclic A-movement by Transfer

Recall that IM applies simultaneously with Transfer at the phase level so that the occurrences created can be identified as copies and can be interpreted as forming a chain, yielding a single interpretation at the interfaces. The standard assumption has been that Transfer cyclically ships off phase-head complements in the course of the derivation, which reconciles cyclic Transfer with cross-phasal operations. Assuming that IM applies in simultaneity with Transfer at the phase level and that what is transferred in the CP phase is TP, I propose that A-movement of the subject to Spec,TP occurs in the transition to the interfaces; in other words, it takes place in the process by which SOs are sent to the interfaces via Transfer, which, through interpretation and externalization, maps SOs to representations accessible to the CI and SM systems (i.e., SEM and PHON). The proposal is illustrated in (6):

(6) a.



b. 

c. (6b) = <SEM, PHON>

In (6), λ is transferred (=(6a)), upon which the subject, in simultaneity with Transfer, is internally merged with λ, which generates Spec,TP (=(6b)); and then, γ as a whole is interpreted and externalized, mapped to SEM and PHON (=(6c)), with the subject in α identified as a copy. Given that the linguistic computation or syntax is the process by which derivations are constructed from lexical items and are mapped to <SEM, PHON> (see, e.g., Chomsky 1995, 2004), SOs leave the hands of syntax and are handed over to the interfaces by Transfer, unable to be further manipulated, when they are mapped to SEM and PHON. IM in A-movement, like IM in Ā-movement, occurs in syntax as the movement takes place before the γ-marked set is mapped to SEM and PHON as in (6c).

The proposed analysis can solve the counter-cyclicity of A-movement: in (6), movement to Spec,TP targets the "root" (that is, transferred TP, which is now not embedded in CP) and does not infix the subject, with the result that it does not violate NTC: thanks to Transfer, TP or λ derivationally turns into a root and A-movement, which applies in simultaneity in Transfer, applies to the root. IM qua A-movement is executed in a way satisfying NTC and keeps to simplest Merge (=(5)). Contra Chomsky (2008), A-movement is an instance of NTC-compliant Merge thanks to the simultaneity of IM with Transfer.[3]

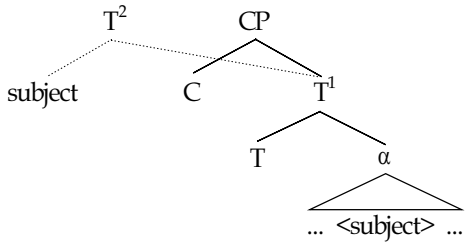
---

3 Given that EM and IM are both instantiations of the same operation Merge, the simultaneity of EM with Transfer will be predicted. One possibility is merge of an expletive (*there* and *it*) as Spec,TP as it does not receive a theta role in the lower position (Bošković 2002; but see also Deal 2009, Nomura 2004). A detailed exploration of the simultaneity of EM with Transfer goes beyond the scope of this paper and is left for future.

As regards counter-cyclic A-movement in phase syntax, Epstein et al. (2012) also discuss the problem and propose a Merge-based solution to it. They claim that in A-movement, Merge applies to produce intersecting set-theoretic objects (7) from the set {C, {T, {*v*P}}}, which happen to share a term T^1; informally shown in graph-theoretic terms, Merge creates a two-peaked/doubly-rooted structure (8):[4]

(7) T^2 = {subject, T^1}, CP = {C, T^1}

(8)

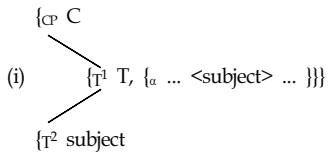


Epstein et al.'s intersecting set analysis, though it may not infix the subject and may not need restructuring to merge it with TP just like (6), is clearly an extension or elaboration of Merge and goes beyond the simplest formulation of Merge given in (5): it does not apply to the root. The analysis requires the manipulation of TP in the already merged {C, {TP}} set and like an infixing, tuck-in movement, expands the structure not at the root but from within, which goes against NTC. As I have argued, the simplest formulation of Merge, which conforms to NTC, suggests that Merge applies to the root and does not allow any manipulation of terms embedded in the set: the Merge which produces (7)/(8) applies to TP embedded inside CP as a target of A-movement, and manipulates it to produce intersecting sets.

---

4 Superscripted numbers are used for expository purposes only and have no theoretical status. (7) can be alternatively represented as (i):

(i)



Moreover, as we can see from (7)/(8), the Merge creating a two-peaked object works in such a way as to bifurcate an existing single set {C, {T, {*v*P}}} into two distinct sets, which intersect with each other. Given that the generative procedure or Merge, as shown in (5), constructs a single object to be mapped onto semantic and phonological representations, integrating two SOs into one, set-bifurcation is an undesirable consequence that follows from Epstein et al.'s analysis.

The problems discussed above do not arise under the solution I have proposed: it does not require any elaboration of Merge or any extension of its application, keeping Merge and its application in the simplest form (=(5)). The IM that produces A-movement is Merge to the root and does not in any way counter-cyclically manipulate TP embedded in {C, {TP}} and merge the subject with it.

2.2 Cyclic A-movement by external pair-Merge

In this section, I propose that A-movement can also be made cyclic by Merge. Given (5) or given that Merge is constrained only by 3rd factor principles, Merge operates freely and can apply asymmetrically as well as symmetrically to create SOs: Merge, applying symmetrically, produces a single set out of any two SOs as in (9a); on the other hand, if it applies asymmetrically, it adjoins one SO to the other and yields an ordered pair (though not imposing any linear order on it) as in (9b):

(9) a. {α, β}
b. <α, β>

Under the assumption of simplest Merge, the two modes of Merge will be available for free, which is empirically supported by substitution and adjunction structures in language. Chomsky (2015) calls the Merge that produces (9a) "set-Merge" and the Merge that generates (9b) "pair-Merge."[5] Keeping in mind

5 As noted in Fukui (2017), the ordered pair <α, β> is mathematically equivalent to the set {α, {α, β}} (a variant of {{α}, {α, β}} in the Kuratowski definition, where α forms a singleton set). See also Tourlakis (2003) for <α, β> = {α, {α, β}}. If so, (9b) can be derived by symmetric Merge or

that set-Merge and pair-Merge are both instantiations of simplest Merge and especially that pair-Merge is nothing other than the set-Merge that yields {α, {α, β}} (i.e., <α, β>) (footnote 5), I use set-Merge and pair-Merge only for the purpose of distinction; that is, the term "pair-Merge" or more specifically, "pair-merge of β to α" is used in this paper to refer to nothing other than the symmetric merge that generates {α, {α, β}} and <α, β> is used instead of {α, {α, β}} for the purpose of illustration.

With this background in mind, in A-movement examples such as (2), suppose that C and T are not set-merged as in (10), where T is set-merged with α, a set headed by *v*/*v** and then C is set-merged with λ, a set headed by T:

(10) a. [$_{\lambda}$ T [$_{\alpha}$ ...]]
b. [$_{\kappa}$ C [$_{\lambda}$ T [$_{\alpha}$ ...]]]

Instead, T is externally pair-merged to C, which generates a head-head amalgam or composite head <C, T>, and then <C, T> is set-merged with α. Notice that pair-Merge, like set-Merge, can be executed both externally and internally as it is an instantiation of Merge.[6] The derived structure is not (10b) but (11b):

(11) a. <C, T>
b. [$_{\delta}$ <C, T> [$_{\alpha}$ ...]]

Under the assumption that Merge is subject to SMT, constrained only by 3rd factor principles, set-Merge and pair-Merge can apply in any order in syntactic workspace to yield (11b), which can only be blocked by stipulation.

Given (11b), the subject is merged with δ and moves to the Spec of <C, T>, which generates (12b) (in (12b) and elsewhere, t represents an occurrence created

---

set-Merge, with Merge applying only symmetrically. This argument is further supported if simplest Merge, as discussed in Chomsky et al. (2019), permits only symmetric Merge, with pair-Merge being a formally distinct operation. See also Oseki (2015) for an attempt to eliminate pair-Merge. However, see, e.g., Park et al. (2019) and references cited therein for empirical arguments for extensions beyond such a narrow conception of Merge.

6 See also Epstein et al. (2016), Mizuguchi (2016a), Nomura (2018) and Sugimoto (2016) among others for external pair-merge of heads. In this paper, unless otherwise mentioned, (pair-)Merge is external (pair-)Merge.

by IM):

(12) a. [$_{\delta}$ <C, T> [$_{\alpha}$ ... subject ...]]
b. [$_{\mu}$ subject [$_{\delta}$ <C, T> [$_{\alpha}$ ... t ...]]]

In (12), when the subject moves, δ is not embedded and movement of the subject can apply to the root: it is not counter-cyclic but cyclic. External pair-merge of T to C can make cyclic A-movement possible.

Notice that in (12b), the subject is internally merged with δ at the phase level, with α cyclically transferred, and that the t in α can be interpreted as a copy. Chomsky (2004, 2015) argues that SOs, when pair-merged to others, get de-activated and become syntactically invisible: SOs, when adjoined, are asymmetric to their hosts and are put on a different plane than their hosts. Recall that pair-Merge yields <α, β> or {α, {α, β}} (footnote 5). When β is pair-merged to α, α embeds {α, β}, hence β; consequently, {α, {α, β}} has the properties of α, with β being de-activated. This means that as schematically illustrated in (13), C with T pair-merged to it is syntactically on a par with C and bears the properties of C, including phasehood; <C, T> thus works as a phase head:

(13)

Also, the Spec of <C, T> constitutes an A-position, with the subject undergoing A-movement to the relevant Spec. It has been argued that ϕ-features play a key role in determining A-positions (Chomsky 2007, Obata 2010, Obata and Epstein 2011, van Urk 2015 among others). For instance, according to van Urk (2015), Ā-movement in Dinka is movement to Spec,CP but shows the typical properties of A-movement. He argues that C in Dinka is a composite probe, having both ϕ-features and those features that drive Ā-movement: movement to Spec,CP in Dinka involves an Agree relation in ϕ-features as well, which morphologically appears on a prefix attaching to the verb/auxiliary (van Urk 2015: Chapter 4). This suggests that movement to the Spec of a ϕ-bearing head

shows A-movement properties, regardless of what head it is. Building on earlier arguments in the literature, I propose (14) as the definition of A/Ā-positions:[7]

(14) The NP is in an A-position if it is merged with an SO headed by a head bearing ϕ-features; otherwise, it is in an Ā-position.

(14) says that a sister relation with a ϕ-bearing SO, which is the basic relation established by Merge(α, β), yields an A-position for the NP (or leads to A-properties of the NP at the CI interface). In the proposed derivation (12), feature-inheritance does not take place since T is externally pair-merged to C and forms a composite head with it. Feature-inheritance is the process by which features on a phase head (C in this case) are transmitted to a non-phase head that it selects (i.e., T). In <C, T>, T is embedded under C and is not available as an independent head in the derivation, with the result that features on C will not be inherited and C keeps ϕ-features. Provided that <C, T> is on a par with C, it carries ϕ-features.[8] This argument is also endorsed by Epstein et al. (2012), who argue that ϕ-feature inheritance occurs for nominative Case valuation by T. The subject moves to merge with δ, a set headed by <C, T>, which is a head bearing the relevant features. Given the definition of A/Ā-positions proposed in (14), the Spec of <C, T> constitutes an A-position and the subject undergoes A-movement to the Spec of <C, T> at the phase level. Under the proposed analysis, the Spec of <C, T> is both an A-position and an Ā-position: it is an A-position as <C, T> bears ϕ-features for under-inheritance; it is also an Ā-position as <C, T> is syntactically on a par with C and bears the properties of C.[9]

I have argued that external pair-merge of T to C provides another way to solve the problem of counter-cyclic A-movement in phase syntax. Given that Merge applies freely, external pair-Merge can go the other way, with C being

---

7 "NP" is used in this paper to refer to a nominal category in general (i.e., NP, DP and phasal NP or *n*P, the head of which (*n*), along with C and *v*, is a phase head).

8 Notice that external pair-merge of T to C derives under-inheritance discussed in Legate (2011, 2014) and Ouali (2008), and a composite probe proposed in van Urk (2015). See section 4 for further discussion on under-inheritance.

9 For the argument that a certain Spec position can be a "mixed" position, see also Bošković (2008), Diesing (1990), Torrego (1984), Uriagereka (1988) and van Urk (2015).

externally pair-merged to T, which yields <T, C>; both <C, T> and <T, C> are straightforward consequences of simplest Merge. In fact, Mizuguchi (2016a) and Sugimoto (2016), developing Epstein et al. (2016), argue for external pair-merge of C to T. Though <T, C>, like <C, T>, may solve counter-cyclic A-movement as the subject can move to the root, the two outcomes will make distinct predictions as regards, in particular, phasehood and the availability of ϕ-features. Notice that <T, C> cannot explain a chain interpretation in A-movement and A-properties of the subject. Consider (15), where the subject moves to the root thanks to <T, C>:

(15) $[_{\tau}$ subject $[_{\omega}$ <T, C> $[_{\alpha}$ ... t ...]]]

In (15), the subject cannot be interpreted as forming a chain, with its occurrences created being interpreted as distinct, since the IM applies at the non-phase level and the interfaces cannot see that the subject forms a chain as Transfer does not apply upon the movement; unlike <C, T>, T with C pair-merged to it is syntactically on a par with T, not C, and does not work as a phase head since T is not a phase head (but see Mizuguchi 2016a for a different proposal). Moreover, the subject cannot be in an A-position given (14) as ϕ on C is de-activated due to its external pair-merge to T; T does not have ϕ-features of its own but inherits them from C (Chomsky 2008). See also Mizuguchi (2016a) and Sugimoto (2016) for empirical arguments for these syntactic effects of <T, C>. Unlike <T, C>, external pair-merge of T to C proposed in the present paper or <C, T>, in addition to solving the counter-cyclicity of A-movement, can explain the chain interpretation and A-properties of a moved NP.[10]

2.3 Summary: Cyclic A-movement by Merge

Summarizing the discussion in this section, I have proposed two solutions to

10 The proposal that two heads (typically, C and T) are bundled into a single head is not new and has been discussed extensively in the literature with various empirical phenomena. See Chou (2018), Erlewine (2017, 2018), Gallego (2014, 2017), Hsu (2016), Martinović (2015) among others and references cited therein. The present paper argues that the bundling of heads or the creation of composite heads can be syntactically executed by Merge. If bundling, as this paper claims, can be taken care of by Merge, other assumptions proposed in the literature to produce composite heads can be wiped out in favor of irreducible Merge.

make A-movement in phase syntax cyclic: cyclic transfer of TP and external pair-merge of T to C. The solutions proposed here do not assume or stipulate any additional mechanisms but rely only on those that are independently motivated: that is, Transfer and Merge. In fact, we can go one step further and argue that cyclic A-movement by Transfer is reducible to cyclic A-movement by Merge. Mizuguchi (2017b) claims that Transfer is reducible to Merge, proposing that it is nothing other than IM: Transfer moves or internally merges the SO to the interfaces.[11] I refer the reader to Mizuguchi (2017b) for details but if Transfer can be reduced to Merge as Mizuguchi claims, then TP or λ will be cyclically transferred by Merge in (6) and Merge plays a key role in solving counter-cyclic A-movement in (6) as well as in (12). Assuming that Mizuguchi (2017b) is correct, I have shown in this section that counter-cyclic A-movement can be circumvented by Merge in phase syntax; counter-cyclicity of A-movement is only apparent.[12]

Given that the two solutions, as I have discussed, follow from independently motivated mechanisms available in Universal Grammar (UG) and especially given that syntactic operations such as Merge and Transfer apply freely, blocked only by stipulation, I argue that both of these solutions are employed by the grammar to resolve counter-cyclic A-movement. For relevant arguments in favor of this conclusion, see section 5.[13]

3. A-movement and extraction of the subject

The solutions I have proposed to make A-movement in phase syntax cyclic can explain the well-formedness of subject extraction discussed in the literature. Consider the following examples:

---

11 See also Ishii (2017) for a relevant proposal. He argues that Transfer is self pair-Merge and that it can be subsumed under Merge.

12 Bearing in mind that Transfer is reducible to Merge, for ease of discussion, I use the term "cyclic A-movement by Transfer" in the rest of the paper.

13 Given that *v**P is a phase along with CP, counter-cyclic A-movement in the *v**P phase or raising-to-object is taken care of in the same way as counter-cyclic A-movement in the CP phase.

(16) a. *Which student does the professor believe [that is the most intelligent]?
b. Which student does the professor believe [is the most intelligent]?
c. Which student does the professor believe [to be the most intelligent]?

(17) a. *The student seems [that is in the library].
b. The student seems [to be in the library].

As we can see from (16) and (17), subject extraction is sensitive to an overt complementizer, being blocked when *that* is present (the so-called "*that*-trace effect" in the literature). The *that*-trace effect (more generally, the comp-trace effect) has been a center of inquiry in generative syntax and various approaches have been proposed in the literature to explain why (16a) and (17a), in contrast with (16b,c) and (17b), are ill-formed (to name but a few, the *that*-trace filter – Chomsky and Lasnik 1977, the ECP – Chomsky 1981, the Nominative Island Condition – Pesetsky 1982, Taraldsen 1980, the EPP – Mizuguchi 2008, label weakness – Chomsky 2015, (anti-)locality – Bošković 2016, Erlewine 2017, Ishii 2004, Pesetsky 2016, 2019, intonational phrase – McFadden and Sundaresan 2018).[14] In this section, I show that given my proposals, the well-formedness of subject extraction in (16) and (17) follows as one consequence of counter-cyclic A-movement, giving a Merge-based analysis to the *that*-trace effect.

Let us start with ill-formed subject extraction in (16a) and (17a), taking (16a) as a representative example. In the derivation, as illustrated in (10), suppose that C and T are set-merged and that (18) is created in the embedded clause, with κ being a phase:

(18) [$_{\kappa}$ that [$_{\lambda}$ T [$_{\alpha}$ is [which student [the most intelligent]]]]]

Given phase syntax, A-movement of the subject to Spec,TP will apply at (18). As I have proposed, movement of the subject to Spec,TP takes place at the phase level in the transition to the interfaces, when λ is transferred and before the derivation is mapped to SEM and PHON (see (6)); otherwise, the IM cannot be executed as it violates NTC. However, this has the effect that the subject will

14 See Pesetsky (2017) for a recent review of the *that*-trace effect and the references cited therein.

become inaccessible to further computation as it is transferred as part of λ when it is internally merged with the set. SOs, once transferred and mapped to SEM and PHON by interpretation and externalization, are kicked out of syntactic workspace and become syntactically invisible and hence, unavailable to further computation. Merge cannot manipulate the subject and it cannot move any further, with the result that it cannot be extracted out of the embedded clause. Consider (19) (in (19) and elsewhere, the shaded part represents <SEM, PHON>; i.e., it has been transferred and is syntactically invisible, hence inoperable):

(19) a. [$_{\kappa}$ that [$_{\lambda}$ T [$_{\alpha}$ is [which student [the most intelligent]]]]]

TRANSFER AND IM

b. [$_{\kappa}$ that [$_{\gamma}$ which student [$_{\lambda}$ T [$_{\alpha}$ is [t [the most intelligent]]]]]]

c. [___ [$_{\kappa}$ that [$_{\gamma}$ which student [$_{\lambda}$ T [$_{\alpha}$ is [t [the most intelligent]]]]]]]

The same analysis applies to (17a), which causes its ill-formedness. Cyclic A-movement by Transfer I have proposed can explain why subject extraction out of the embedded clause is impossible in (16a) and (17a).

Notice that in (18), A-movement to Spec,TP cannot be skipped, though direct movement to Spec,CP as shown in (20), which can circumvent the transfer of the subject, is possible under simplest Merge:

(20) [which student [$_{\kappa}$ that [$_{\lambda}$ T [$_{\alpha}$ is [t [the most intelligent]]]]]]

Though the derivation in (20) is syntactically unproblematic, it will cause labeling failure and will be ruled out at the interfaces for an unlabeled SO. Chomsky (2015: 9-10) argues that T in languages like English is subject to label weakness: it is too weak to serve as a label and cannot label on its own; in order to work as a label, it must have overt or visible Spec,TP, which strengthens T and makes it a labelable head, when the labeling algorithm Label applies to identify λ, a set headed by T. For labeling purposes, A-movement of the subject to the Spec of TP is obligatory and the subject cannot move directly to the embedded Spec,CP. Given (18), whether or not the subject moves to the embedded Spec,TP, subject extraction will be

ill-formed.[15]

Let us now turn to (16b,c) and (17b) and consider why subject extraction is well-formed in the examples unlike in (16a) and (17a). Take (16b) for illustration. I argue that in the relevant examples, the embedded clause or the embedded phase is structured through external pair-merge of T to C: T is pair-merged to C as in (11), and (21) is derived in the embedded clause, with δ being a phase:

(21) [$_{\delta}$ <C, T> [$_{\alpha}$ is [which student [the most intelligent]]]]

At the δ phase, the subject is internally merged. In (21), A-movement of the subject targets δ and applies to the root, with counter-cyclic A-movement solved by the pair-merge. In (21), what becomes syntactically invisible is only α, which is cyclically transferred as the complement of <C, T>, a phase head. The subject moves out of α upon transfer of the set at the phase level; consequently, it is not transferred and is visible at the Spec of <C, T>. When the higher phase is structured, the subject is available to further computation and IM can apply to the NP, with the result that it can move out of the embedded clause. Consider (22):

(22) a. [$_{\mu}$ which student [$_{\delta}$ <C, T> [$_{\alpha}$ is [t [the most intelligent]]]]]
b. [which student [does the professor believe [$_{\mu}$ t [$_{\delta}$ <C, T> [$_{\alpha}$ is [t [the most intelligent]]]]]]]

15 Simultaneous, parallel movement proposed by Chomsky (2008) makes a wrong prediction for subject extraction: in (16a), for instance, as illustrated in (i), upon transfer of λ, the subject moves directly to Spec,CP, which makes subject extraction possible, and at the same time undergoes cyclic A-movement to Spec,TP, thanks to which λ can be labeled:

(i) a. [$_{\kappa}$ that [$_{\lambda}$ T [$_{\alpha}$ be [which student [the most intelligent]]]]]

TRANSFER AND IM

b. [which student [$_{\kappa}$ that [$_{\gamma}$ which student [$_{\lambda}$ T [$_{\alpha}$ be [t [the most intelligent]]]]]]]

c. [which student [$_{\kappa}$ that [$_{\gamma}$ which student [$_{\lambda}$ T [$_{\alpha}$ be [t [the mostintelligent]]]]]]]

Simultaneous, parallel movement cannot explain subject extraction. See also Gallego (2017) and Mizuguchi (2016b) for problems with such movement.

Legitimate subject extraction can be explained by cyclic A-movement by pair-Merge. Notice that μ and δ can be properly labeled when the subject moves out and turns into a copy, which does not form over/visible Spec (Chomsky 2013: 44) and is invisible to the labeling algorithm. Recall from section 2 that <C, T> has the properties of C. Since C, unlike T, is not subject to label weakness and can label on its own, <C, T> can label without overt/visible Spec.

In the discussion so far, I have demonstrated that the well-formedness of subject extraction follows from the proposed solutions to counter-cyclic A-movement. Given that Merge applies freely, the proposed analyses of (16) and (17) will raise the following questions: why aren't C and T externally pair-merged in the embedded clause in (16a) and (17a)? Likewise, why aren't C and T set-merged in the embedded clause in (16b,c) and (17b)? Only a stipulation can block pair-Merge and set-Merge in the relevant examples under simplest Merge, which can make the opposite predictions on the well-formedness of (16) and (17). I claim that externalization at the SM level answers the questions at hand. I argue that C is externalized differently, depending on how C and T are merged in the derivation. I maintain that <C, T> and C are realized as distinct complementizers at the SM level and propose (23) (cf. Sugimoto 2016 for externalization of *to*):

(23) **Externalizations of C**

Syntax	Externalization
$<C, T>_{tensed}$	Ø
$<C, T>_{-tensed}$	*to*
C	*that*

Given the proposal, complementizers indicate how C is merged in the derivation: the null or zero complementizer Ø in English is a spell-out of tensed <C, T> and the English infinitival *to* is an externalization of <C, T> without tense while the overt complementizer *that* is an externalization of C. Notice that finiteness is irrelevant to the application of Merge, hence pair-Merge. In (16a) and (17a), since *that* appears through externalization at the SM level, it indicates that C is merged as an independent head: C and T are set-merged, with (18), not (21), yielded in

the embedded clause. On the other hand, in (16b,c) and (17b), externalizations as Ø and *to* argue that C forms a composite head with T: T is pair-merged to C and (21) is created in the embedded clause.

Notice that the resort to externalization is not a stipulation at all to explain how Merge applies in (16) and (17). Externalization, which is peripheral to syntax or UG, is subject to variation within and across languages and it can be considered a locus of linguistic variation, inducing superficial complexity and variety of language, given the assumption, which is reasonable for scientific and biological reasons, that UG is quite simple at its core and uniform across languages, not subject to parameterization or variation (the Uniformity Principle) (Berwick and Chomsky 2011, 2016, Boeckx 2011, Chomsky 2001, 2017, Miyagawa 2010). Since externalization allows variation, it is not unreasonable to assume that syntactic structures or representations have realizational options, externalized with a variety of lexical items depending on the way syntax or Merge applies in the derivation. Externalization can explain syntactic derivations in (16) and (17) under simple, uniform syntax without constraining Merge.[16]

One reviewer asks what is an underlying principle that explains (23). I suggest that the principle or what is behind the scene is that *that*, Ø and *to* are nothing other than the allomorphs of C: i.e., the forms that are determined by contexts or syntactic contexts in this case and that are without any change in meaning. Recall that C and <C, T> are two distinct heads; <C, T> is an amalgamated or composite head due to pair-Merge while C is not, which is to say that in (10b) and (11b), C is in different syntactic contexts. Moreover, syntactic contexts are distinguished by the tense property of C. Given the suggestion, C, which has allomorphs, is spelled out as different forms when mapped to PHON depending on the syntactic context it is in.

Provided that C has *that*, Ø and *to* as its allomorphs, the next question is

16 Unlike (16b), (i) is ill-formed even though Ø is externalized:

(i) *The student seems [is in the library].

I argue that (i) is ruled out independently. In (i), unlike in (16b), the matrix ϕ-probe cannot agree: the NP *the student* agrees with <C, T> and has its Case feature valued in the embedded clause. Since Case valuation de-activates the NP for ϕ-feature agreement (Chomsky 2000), it cannot agree again with the matrix ϕ-probe, which leaves the features unvalued and violates Full Interpretation.

why *that* is an externalized form of C while Ø and *to* are externalized forms of <C, T>, and not the other way around. I have to admit that at the current stage of research, I have no illuminating answer to this question. I argue, however, that progress in understanding externalization or the mapping to PHON goes hand in hand with progress in understanding syntactic derivations or SOs that are subject to externalization, with one feeding the other and vice versa. I have argued that Merge can freely generate both (10b) and (11b) and that only the latter allows subject extraction. By assuming that *that* is an externalized form of C while Ø and *to* are externalized forms of <C, T>, and not the other way around, the grammatical behaviors of (16) and (17) will fall out as predicted by syntactic derivations, which suggests that externalization of C reasonably works as proposed in (23).

The proposal in this section can deduce, as its consequence, the subject-extraction generalization discussed in Pesetsky (2016, 2019), answering why clause size matters for extraction of the subject. Pesetsky argues for the following generalization:

(24) Subject extraction always entails a smaller-than-full clause.

To explain (24), Pesetsky proposes a deletion operation called Exfoliation, which removes a certain portion of the structure as a by-product of subject extraction. He argues that the operation solves a locality problem that has appeared in the derivational process and makes subject extraction possible. In examples such as (16b,c) and (17b), CP is deleted and removed by Exfoliation, which reduces clause size and makes subject extraction out of the embedded clause possible:

(25) ... $[_{VP}$ V $\underbrace{[_{CP}$ C$]}_{\text{Ø Exfoliated}}$ $[_{TP}$ T $[_{vP}$ *v* [...]]]]] ⇒ ... $[_{VP}$ V $[_{TP}$ T $[_{vP}$ *v* [...]]]]

In (16a) and (17a), on the other hand, CP is not deleted, which blocks subject extraction.[17]

17 The generalization is not new and Pesetsky's proposal has Chomsky (1981) as its precursor. In Chomsky (1981), S'-deletion or more generally, the S'-to-S rule, is proposed to explain subject extraction in examples such as (i):

Under the analysis proposed in this section, in examples in which subject extraction is well-formed, T is externally pair-merged to C and the clause is necessarily reduced to a smaller-than-full clause: there is only one set above α (i.e., δ), not two (i.e., κ and λ) (see (10) and (11)). Clause reduction for subject extraction (that is, the generalization (24)) is explained by Merge and no extra mechanism beyond Merge (such as literal deletion operations like Exfoliation and S′/*that*-deletion – footnote 17, which are syntactically executed and hence tamper with the existing structure, violating NTC) needs to be assumed to explain clause reduction. Moreover, cyclic A-movement by Transfer answers why clause size or reduction of CP matters for subject extraction: unless the clause is reduced by external pair-merge of T to C, the subject will be unavailable to further computation for its transfer in the embedded clause due to A-movement to the Spec of TP. Pesetsky's generalization is deduced in a principled manner under the proposals in this paper.

4. A-movement in local subject Ā-movement

In the last section, I have argued that successful subject extraction from the embedded clause requires cyclic A-movement by external pair-Merge and that the embedded clause necessarily becomes smaller than a full clause. In this section, I consider A-movement in subject Ā-movement in a single clause or local subject Ā-movement and argue that counter-cyclic A-movement in local subject Ā-movement is also solved by external pair-merge of T to C, showing that A-movement in local subject Ā-movement and that in subject extraction are explained in the same way. As instances of local subject Ā-movement, consider (26):

(i) a. Who do you think [$_S$ saw Bill]?
b. Bill was believed [$_S$ to have seen Tom]. (Chomsky 1981: 298, 303)

More recently, Chomsky (2015) proposes *that*-deletion, which deletes C and removes a C-headed set or CP, to account for the contrast between (16a) and (16b). He argues that together with other assumptions, deletion of CP in (16b) allows the embedded TP to be labeled even without overt/visible Spec,TP when the subject is extracted and turns into a copy. See also Bošković (1997), Doherty (2000), and Ishii (2004), who argue that the *that*-less clause in (16b) is TP, as well as Erlewine (2017), who argues that it is CTP, a set headed by the lexically composite head CT.

(26) a. Which professor will give a linguistics lecture?
b. I wonder [which professor will give a linguistics lecture].

In the derivation of (26), suppose that C and T are set-merged and that the subject moves to Spec,TP for A-movement, with counter-cyclic A-movement solved by Transfer. In this case, however, as shown in (27), the subject will be transferred upon its movement at the phase level, which applies to λ when the set is transferred and becomes a derivational root. As discussed, this makes the subject invisible to computation since it is transferred as part of λ upon its merge with the set, with the result that it cannot be manipulated and cannot move onto the Spec of CP from Spec,TP. The subject *wh*-phrase, being an operator, needs to be in Spec,CP, an operator position, for proper interpretation; failure to move to Spec,CP will cause interpretive ill-formedness for the *wh*-phrase at the CI interface:

(27) a. [$_{\kappa}$ C [$_{\lambda}$ T$_{will}$ [$_{\alpha}$ which professor [give a linguistics lecture]]]]
TRANSFER AND IM
b. [$_{\kappa}$ C [$_{\gamma}$ which professor [$_{\lambda}$ T$_{will}$ [$_{\alpha}$ t [give a linguistics lecture]]]]]
c. [___ [$_{\kappa}$ C [$_{\gamma}$ which professor [$_{\lambda}$ T$_{will}$ [$_{\alpha}$ t [give a linguistics lecture]]]]]]

Recall that the subject cannot move in a single leap to Spec,CP, skipping Spec,TP as in (20), which causes labeling failure with λ for weakness of T as a label. Moreover, in (27), C will be externalized as *that* given (23), which, as evidenced by (28), is not spelled out in local subject Ā-movement:[18]

18 Rizzi and Shlonsky (2007) point out an example which corresponds to (28a):

(i) Quel garçon qui est venu? [Québec French]
Which boy QUI has come
'Which boy has come?' (Rizzi and Shlonsky 2007: 142)

Notice, however, that the overt complementizer in (i) is not *que* but *qui*, which can be analyzed as *que* + the expletive *i* (Taraldsen 2002). In (i), labeling failure with λ can be solved by merging the expletive as Spec,TP, which allows the subject to move directly to Spec,CP. For relevant discussion, see Bošković (2016) and Mizuguchi (2018).

(28) a. *Which professor that will give a linguistics lecture?
b. *I wonder [which professor that will give a linguistics lecture].

On the other hand, if counter-cyclic A-movement is solved by external pair-merge of T to C, the problems discussed above do not arise. Consider (29):

(29) a. [$_{\delta}$ <C, T>$_{will}$ [$_{\alpha}$ which professor [give a linguistics lecture]]]
b. [$_{\mu}$ which professor [$_{\delta}$ <C, T>$_{will}$ [$_{\alpha}$ t [give a linguistics lecture]]]]

At the δ phase (=(29a)), A-movement of the subject applies to the root δ and the subject moves to the Spec of <C, T>. Notice that the movement is to an operator position and that the subject can be interpreted properly at the CI interface: recall that <C, T> is syntactically on a par with C and bears the properties of C, including the property [Q] or a Q-feature, which makes possible the interpretation of the *wh*-phrase in the Spec of <C, T> (i.e., through Spec-head or XP-YP − Epstein et al. 2015; also see below). Given that the subject *wh*-phrase must be in Spec,CP for interpretation, counter-cyclic A-movement will be solved by pair-Merge in local subject Ā-movement as well as in subject extraction out of the embedded clause. Finally, since <C, T> is generated, it is not externalized as the overt complementizer *that*; instead, it is externalized as *Ø*.

Notice that the analysis of A-movement in local subject Ā-movement proposed in this paper can explain Spec-to-Spec anti-locality discussed in Erlewine (2016). He argues that Ā-movement that is too short is banned, proposing (30):

(30) **Spec-to-Spec Anti-Locality**
Ā-movement of a phrase from the Specifier of XP must cross a maximal projection other than XP. (Erlewine 2016: 445)

"Crossing" in (30) is defined as (31):

(31) Movement from position α to position β crosses γ if and only if γ dominates α but does not dominate β. (Erlewine 2016: 445)

Spec-to-Spec anti-locality bans movement from Spec,TP to Spec,CP as it crosses only the maximal projection of TP. Evidence for this comes from the emergence of Agent Focus (AF) in subject *wh*-movement in Kaqchikel. Consider (32):[19]

(32) a. Achike *x-ø-u-tëj / ✓x-ø-u-tj-ö ri wäy?
who COM-B3.SG-A3.SG-eat / COM-B3.SG-eat-**AF** the tortilla
'Who ate the tortilla?'
b Iwïr x-ø-u-tëj ri wäy ri a Juan.
yesterday COM-B3.SG-A3.SG-eat the tortilla Juan
'Yesterday Juan ate the tortilla.' (Erlewine 2016: 430)

Erlewine argues that AF appears on the verb when the subject *wh*-phrase skips the position of Spec,TP and the Spec is not created. Compare (32a) with the basic transitive clause VOS in (32b), where the subject moves to the Spec of TP and AF does not appear. The skipping of Spec,TP or Spec,TP-to-Spec,CP anti-locality will be straightforward given the solutions to counter-cyclic A-movement proposed in this paper. Spec,TP-to-Spec,CP movement suggests that C and T are set-merged as in (10). As discussed, if the subject undergoes A-movement to Spec,TP, it will be transferred as part of TP and will become syntactically inaccessible, with the result that it cannot move to Spec,CP from Spec,TP (=(27)); counter-cyclic A-movement in local subject Ā-movement can only be solved by external pair-merge of T to C, which does not produce a set headed by T (i.e., λ) and the subject does not move to Spec,TP. Consequently, the verb is marked with AF. (30) or the ban on movement from Spec,TP to Spec,CP is explained by the proposed solutions to counter-cyclic A-movement.

The proposed analysis of local subject Ā-movement provides a novel analysis of the Vacuous Movement Hypothesis (VMH), according to which the subject moves to a single Spec in its Ā-movement. Two proposals on VMH have been put forward in the literature: vacuous movement to Spec,CP and vacuous movement to Spec,TP. Chomsky (1986) was the first to propose VMH, arguing

19 See also Bošković (2016) for the argument based on Kinande that *wh*-movement to Spec,CP does not proceed via Spec,TP.

Abbreviations in (32): A = Set A agreement; B = Set B agreement; COM = completive aspect; SG = singular; 3 = third person. For details, see Erlewine (2016).

that the operator subject can move only up to Spec,TP at S-Structure; Chomsky (2013) argues that the subject does not move to Spec,CP at all but halts in Spec,TP, which is a criterial position for the subject for Q-feature inheritance (=(33a)). On the other hand, Legate (2011, 2014) and Ouali (2008) claim that it moves only to Spec,CP, skipping Spec,TP (=(33b)):

(33) a. [$_{CP}$ C [$_{TP}$ subject [T [$_{v}$P ... t ...]]]]
b. [$_{CP}$ subject [C [$_{TP}$ T [$_{v}$P ... t ...]]]]

Under the proposed analysis, vacuous movement is due to the fact that T is externally pair-merged to C, which solves counter-cyclic A-movement. This has the effect that the subject *wh*-phrase moves to a single Spec (i.e., the Spec of <C, T>) but not to two Specs (Spec,CP and Spec,TP).

Notice that the proposed version of VMH (=(29b)) is theoretically advantageous over the previous analyses of VMH. Epstein et al. (2015) argue that a *wh*-phrase is properly interpreted at the CI interface if it forms an XP-YP structure with a Q-bearing head (i.e., it is in the Spec of the head). Chomsky (2013) says that Q-feature is inherited from C to T in (33a), which makes possible the interpretation of a subject *wh*-phrase via XP-YP. In the proposed VMH, the subject *wh*-phrase can be interpreted without Q-feature inheritance. Recall that <C, T> is syntactically on a par with C. Since C has a Q-feature, <C, T> is a Q-bearing head, with the result that the *wh*-phrase, which is internally merged with the set headed by <C, T>, can be properly interpreted at CI thanks to the creation of (29b) without Q-feature inheritance. The removal of Q-feature inheritance is desirable in that it requires feature copying in order to leave Q in its original position for selection and labeling (Chomsky 2013: 47), which, Carstens et al. (2016) argue, causes the derivation to crash as ϕ-features are also copied and remain on C without being valued.[20] Moreover, the inheritance in question occurs only in "subject" *wh*-movement, which can be warranted only by stipulation.

The proposed version of VMH, in addition to eliminating Q-feature

20 One solution to the problem is discussed in Chomsky (2015), who proposes C-deletion (see section 3). As I have discussed in that section, C-deletion is not compliant to 3rd factor principles as it violates NTC and is theoretically unfavorable.

inheritance in (33a), can solve a labeling problem which arises under (33b). Recall once again that T is weak as a label and cannot label without Spec; in order to function as a label, it must have overt/visible Spec and {{Spec}, {TP}} must be created (Chomsky 2015). In (33b), the subject moves directly to Spec,CP and Spec,TP is not generated, which causes labeling failure. In (29), T is externally pair-merged to C and a T-headed set is not created in the derivation; consequently, the labeling problem does not arise. In addition, the proposed analysis also accounts for why feature-inheritance does not take place in local subject Ā-movement. Legate (2011, 2014) and Ouali (2008) argue that ϕ-features are not inherited in local subject *wh*-movement or in (33b), with C retaining the features:

(34) $[_{CP}$ subject $[C_{\phi}$ $[_{TP}$ T $[_{vP}$ ... t ...]]]]

As discussed in section 2, when T is externally pair-merged to C, local subject Ā-movement being one such case, T is not available in the derivation as it has become part of C and feature-inheritance does not take place; it follows that C or <C, T> keeps ϕ. Under the proposed analysis, under-inheritance receives a Merge-based, principled explanation.[21, 22]

5. <C, T>, the zero complementizer and the EPP effect

In this paper, I have argued that counter-cyclic A-movement can be solved

21 Mizuguchi (2014) argues that feature-inheritance is preempted by head movement (i.e., internal pair-Merge). This proposal, however, leaves (a copy of) T in the derivation, making feature-inheritance to this T possible.

22 One reviewer asks how the theory proposed here can handle overt *that* observed in the relative clause in examples like (i):

(i) The man [that came here yesterday] is my brother.

In this paper, following Mizuguchi (2008), I assume that *that* in (i) is not a complementizer but a relative pronoun: in the subject relative clause, <C, T> is generated and is externalized as Ø in accord with (23), with *that* as a relative pronoun internally merged as the Spec of <C, T>. With this assumption in place, (i) is not ungrammatical since *that* is not an externalized form of C and null operator movement is not involved; that is, the subject relative clause is derived in the same way as local subject Ā-movement in (26) (=(29)).

by external pair-merge of T to C and that C with T pair-merged to it (i.e., the composite head <C, T>) is externalized as the zero complementizer Ø in English: Ø is not the result of (narrow-syntactic or phonological) *that*-deletion but of pair-Merge; the overt complementizer *that*, on the other hand, is a realization of C. In this section, I discuss how the zero complementizer in embedded and root clauses is explained by the proposed analysis and consider what the proposals made in this paper imply for the EPP effect.

In English, as shown in (35), the embedded clause shows the alternation of *that* and Ø without any change in meaning:

(35) The professor believes (that) the student will win the prize.

Given the analysis proposed in the present paper, Ø in (35) results from pair-merge of T to C. Under simplest Merge, the pair-merge in question can apply externally and internally to yield the zero complementizer. Consider (36):

(36) a. The professor believes [$_{\mu}$ the student [$_{\delta}$ <C, T> [$_{\alpha}$ t [win the prize]]]].
b. The professor believes [$_{\pi}$ <C, T> [$_{\gamma}$ the student [$_{\lambda}$ T [$_{\alpha}$ t [win the prize]]]]].

In languages like English, T-to-C movement does not obtain in the declarative sentence. Given this descriptive generalization, it can be concluded that the zero complementizer in (35) results from T being externally pair-merged to C (=(36a)).[23] On the other hand, in syntactic contexts where T-to-C movement does take place, the zero complementizer can appear thanks to internal pair-merge of T to C. This is evidenced by (37) through (39):

(37) a. I wondered [if they had read the book]. [Belfast English]
b. I wondered [had they read the book].

23 In this case, as discussed, counter-cyclic A-movement is solved by pair-Merge. On the other hand, when *that* is present in (35), (10b) is structured by set-Merge, with counter-cyclic A-movement solved by Transfer. (35) shows that, as discussed in section 2.3, both cyclic A-movement by Transfer and cyclic A-movement by pair-Merge are freely employed to solve counter-cyclic A-movement.

c. *I wondered [if had they read the book]. (Henry 1995: 114)

(38) a. Who did John claim [that he saw]? [Belfast English]
b. Who did John claim [did he see]?
c. *Who did John claim [that did he see]? (Henry 1995: 114)

(39) a. Er sagt, [daß die Kinder diesen Film gesehen haben]. [German]
he says that the children this film seen have
'He says that the children saw this film.'
b. Er sagt, [die Kinder haben diesen Film gesehen].
c. *Er sagt, [daß die Kinder haben diesen Film gesehen].
(Vikner 1995: 66)

(37) and (38) are examples from Belfast English and (39) are examples from German. In (37b), (38b) and (39b), T-to-C movement or internal pair-merge of T to C occurs, obviating overt complementizers in (37a), (38a) and (39a) as <C, T> is produced. Notice that the proposed analysis can explain the fact that overt complementizers block T-to-C movement; that is, overt complementizers and T-to-C movement are in complementary distribution (see (37c), (38c) and (39c)): overt complementizers are externalizations of C, suggesting that internal pair-merge of T to C, which is behind T-to-C movement and yields <C, T>, does not occur in the derivation.

External pair-merge of T to C also explains Ø in the root clause. In English, the overt complementizer does not appear in the root:

(40) (*That) the government may change its policies.

In the root clause as well as in the embedded clause, T is externally pair-merged to C, with <C, T> externalized as the zero complementizer. On the other hand, if set-Merge applies to C and T in the root and (10b) is yielded, which is predicted under simplest Merge, overt complementizers will appear even in the root if T is not internally pair-merged to C as in (37b), (38b) and (39b). This is empirically evidenced by (41) (see also (i) in footnote 18):

(41) a. Que mi gato se enratonó. [Spanish]
that my cat itself enmoused
'My cat got sick from eating too many mice.' (Ross 1970: 270)
b. Ob Johanna den Wagen verkauft hat? [German]
whether Joan the car sold has
'Has Joan sold the car?' (Clahsen and Smolka 1986: 156)

The absence of internal pair-merge of T to C is evidenced by (41b), where the verb does not come to the second position as in (39a) when the overt complementizer is present.

On the other hand, in examples such as (42), where T-to-C head movement does take place, Ø can be externalized thanks to internal pair-merge of T to C just as in (37b), (38b) and (39b):

(42) a. Can John come to the meeting?
b. What will the student read?

Under the proposed analysis of the zero complementizer, T is externally pair-merged to C in *that*-less clauses such as (35), (40) and local subject *wh*-movement discussed in the last section, with counter-cyclic A-movement in a phase solved by the subject being merged with a set headed by <C, T> and moving to its Spec. Recall that the movement is considered an instance of A-movement given the definition of A/Ā-positions in (14), which is repeated below for convenience:

(14) The NP is in an A-position if it is merged with an SO headed by a head bearing ϕ-features; otherwise, it is in an Ā-position.

Thanks to external pair-merge of T to C, ϕ and tense are under-inherited and remain on C, with <C, T> bearing ϕ-features. Consequently, when moved to the relevant Spec position, the subject will be interpreted as being in an A-position, showing A-properties in *that*-less clauses as well. (43) and (44) are explained without any problems:

(43) a. (The professor thinks) the student Ø appears to himself to be intelligent.
b. (People believe) every boy Ø seems to his mother to be intelligent.

(44) a. (John wonders) who Ø appears to himself to be intelligent.
b. (Mary wonders) who Ø seems to his mother to be intelligent.

It follows from (14) that the anaphor can be bound by the moved subject and that weak cross-over does not occur in movement to the subject position, regardless of *that*.

The proposed analysis of the zero complementizer or the proposal to solve counter-cyclic A-movement by external pair-merge of T to C will raise a question on the EPP. Recall once again that T is too weak to label on its own and that it can label if overt/visible Spec,TP is created. This, Chomsky (2015) claims, explains the EPP, which requires that clauses have subjects (Chomsky 1981, 1982), and subject raising to Spec,TP. If counter-cyclic A-movement is solved by external pair-merge of T to C, which realizes the zero complementizer through externalization, the prediction is that the EPP effect will be absent. Recall from section 2 that <C, T> is on a par with C. Since C can work as a label on its own, <C, T>, unlike T, can label even if overt/visible Spec is not yielded. This consequence is discussed in section 3 with (22), where the subject is a copy and is invisible to Label in the embedded clause and yet labeling failure does not arise, with μ and δ labeled successfully. Also, Mizuguchi (2017c) considers languages such as Icelandic and Yiddish and argues that <C, T> is a labelable head. But this is not the case with English examples such as (45), where δ is a set headed by <C, T>, which is indicated by the zero complementizer, but nevertheless the EPP is forced:

(45) a. *The professor believes [$_{\delta}$ will [$_{\alpha}$ be written the paper by the student]].
b. *[$_{\delta}$ May [$_{\alpha}$ be passed the resolutions by the government]].

Failure to create the Spec of <C, T> leads to ill-formedness. Then the question is: why is the EPP forced even when T is externally pair-merged to C in some languages but not in others, and in (45) but not in (22)?

I claim that labeling is one side of the EPP effect; the other side is externalization at the SM level. Mizuguchi (2016b, 2017a) argues that Spec,TP (or the {{Spec}, {TP}} configuration) works as a syntactic instruction to the externalization component that agreement is spelled out canonically at T. Consider (46):[24]

(46) $[_{\kappa}$ C $[_{\gamma}$ NP $[_{\lambda}$ T_{ϕ} $[_{\alpha}$ ... (t) ...]]]] (→ canonical spell-out of ϕ on T)

(46) is a condition imposed on externalization at the SM level, which spells out SOs created by syntax so that requirements of the SM system can be met (see also footnote 1). If the Spec,TP position is not created, ϕ-features on T cannot be canonically externalized, with the result that the derived expression cannot satisfy the SM requirement that ϕ be morpho-phonologically realized.

Mizuguchi argues that (46) is empirically endorsed by languages such as Standard Arabic and a northern Italian dialect of Trentino, in which full-fledged morphological agreement appears when subjects are pre-verbal and move to Spec,TP while non-canonical (i.e., partial/default) agreement emerges when subjects are post-verbal and do not move to Spec,TP. Likewise, Guasti and Rizzi (2002: 175-176) point out that when the subject occupies a surface position in the higher parts of the inflectional system, typically higher than the inflected verb − presumably in the Spec of agreement or higher as in *[Subj* $Infl_{AGR}$ *...]*, the morphological expression of agreement is compulsory (provided that the language is equipped with the relevant morphology); on the other hand, if the subject stays in a VP-internal position or in the lower part of the inflectional system, typically lower than the inflected verb − presumably, lower than the agreement layer as in *[...* $Infl_{AGR}$ *... Subj ...]*, then agreement may or may not be morphologically expressed. Obligatory morphological expression of agreement can be taken as full-fledged or canonical agreement.

In this paper, building upon Mizuguchi's argument, I propose (47), which subsumes (46):

---

24 (46) is in line with Holmberg (2000), Landau (2007), McFadden and Sundaresan (2018), Richards (2016), Sigurðsson (2010) among others, who argue that phonological properties of languages (hence, externalization) are relevant to the EPP effect.

(47) [NP [X_ϕ [... (t) ...]]] (→ canonical spell-out of ϕ on X)

(47) says that syntactically valued ϕ-features of a head X are morpho-phonologically externalized canonically by Spec,XP: that is, (46) is not restricted to T_ϕ but is generalized to any head with ϕ-features (i.e., X_ϕ). This generalization is reasonable considering that what matters in (46) is not a specific category or head but ϕ-features.

Suppose that in English, unlike in languages like Standard Arabic and Trentino, non-canonical agreement cannot be realized as the language is not equipped with the relevant morphology and that only canonical agreement is morphologically available; canonical agreement is the only realizational option for ϕ. Given (47), in languages like English, ϕ on <C, T> will not be externalized or spelled out and the SM interface condition cannot be met, which violates Full Interpretation, unless the NP is externally or internally merged as the Spec of <C, T>, which explains the EPP effect. (47) rules out (45) even though T is externally pair-merged to C and there is no problem with labeling.

The proposal can also explain (22) since copies, though syntactically de-activated and invisible, are indeed visible to the interfaces, hence to externalization; otherwise, copies cannot be interpreted by the interfaces. A copy in the Spec of <C, T> can thus satisfy (47). That copies are visible to the interfaces is evidenced, for instance, by (48) and (49):

(48) a. Caidé aL shíleann Seán [aL iarrann sé ar Nollaig]? [Irish]
what C thinks John C asks he on Noel
'What does John think he asks Noel for?'
b. Deir siad [goN síleann an t-athair [goN bpósfaidh Síle é]].
say they C thinks the father C will-marry Sheila him
'They say that the father thinks that Sheila will marry him.'
(McCloskey 1979: 17, 232)

(49) How much criticism of *himself* does *John* think his wife will tolerate?

In (48a), the complementizer is realized as *aL*, instead of *goN*, in the embedded clause as well as in the matrix clause, where the *wh*-phrase moves to Spec,CP.

Given that *aL* is analyzed as a morphological spell-out of C with its Spec by externalization and that *wh*-movement is executed successive cyclically for phase impenetrability by cyclic Transfer, the fact that *aL* appears in every embedded clause in the path of *wh*-movement argues that copies, along with *caidé* in the matrix Spec,CP, are visible to externalization when C is morphologically realized. In (49), *himself* can be bound only in the intermediate Spec,CP, which suggests that copies are subject to interpretation and hence visible to the interfaces.

The proposal that the EPP effect is reduced to externalization of ϕ-features can deduce forced Spec-head agreement in Bantu languages such as Kirundi and Kilega, where SOs which control subject agreement are in Spec,TP (or the Spec of <C, T> when the complementizer is externalized as Ø under the proposal in this paper), not in any other positions (see, e.g., Kinyalolo 1991, Ndayiragije 1999). Suppose that morphological agreement in these languages is canonical agreement, with no non-canonical (partial/default) agreement available for lack of the relevant morphology as in English. Then whether the head is T or <C, T>, the Spec must be created so that the ϕ-features of the head can be externalized canonically and (47) imposed by the SM system can be satisfied. Though Agree is one instantiation of minimal search and is executed without Spec-head, (47) makes it look as if Spec-head agreement occurs in Bantu languages under Agree theory of feature valuation.[25]

Notice that (47) can predict parametric variation with the EPP effect without any stipulation; more specifically, it predicts that the Spec of ϕP will not be required and the EPP effect will not be observed if morphological agreement is absent or ϕ-features are not at all morpho-phonologically realized by externalization. This prediction is borne out by optional raising-to-object in English (Lasnik 2004) and optional raising-to-subject in languages such as Japanese (Fukui 1995). These cases can be correctly explained by (47) since, as evidenced by (50) and (51), ϕ-features are not at all morpho-phonologically spelled out:[26]

---

25 Merge allows a non-agreeing SO to be merged as the Spec of X_{ϕ}, which can also externalize ϕ canonically under (47). This is independently ruled out: if a non-agreeing SO is merged as the Spec, the resulting XP-YP structure will not be labeled for lack of agreement between X and Y. See Mizuguchi (2017a), who deduces (46), hence (47), from labeling of XP-YP (Chomsky 2013, 2015).

26 In (50), labeling is unproblematic with R or V without overt/visible Spec thanks to its pair-merge

(50) a. I [$_{\phi P}$ like$_{\phi}$ [the book / books]].
b. I [$_{\phi P}$ believe$_{\phi}$ [[the student / the students] to be intelligent]].

(51) [$_{\phi P}$ [Taroo-ga sono hon-o /ni-satsu hon-o t_{yon}] yon-da$_{\phi}$].
Taroo-NOM the book-ACC 2-CL hon-ACC read-PST
'Taroo read the book / two books.'

In (50) and (51), the Spec in question is irrelevant to externalization of ϕ and hence the EPP effect is not forced.

Summarizing this section, the zero complementizer is a spell-out of <C, T>, which is formed by (external and internal) pair-merge of T to C.[27] I have discussed how the distribution of Ø in embedded and root clauses is explained by the proposed analysis. I have also discussed how the EPP effect is explained under the proposed analysis of the zero complementizer or the proposal to solve counter-cyclic A-movement by external pair-merge of T to C, arguing that it follows from canonical externalization of ϕ. The discussion in this section suggests that an XP-YP structure plays an important role not only in interpretation at the CI interface (Epstein et al. 2015) but also in morpho-phonological interpretation or externalization at the SM level.

6. Conclusion

In this paper, I have addressed a problem that is posed in the phase-based model of syntactic derivation: counter-cyclic A-movement. I have claimed that Transfer and Merge make A-movement cyclic in phase syntax, arguing that the counter-cyclicity is only apparent. I have also suggested that Transfer is reducible to Merge, with the solution of the counter-cyclicity reducible solely to Merge. I have demonstrated that subject extraction out of the embedded clause is explained by the proposed solutions to counter-cyclic A-movement; I have also

---

to *v*; <*v*, R> can label alone as it is on a par with *v*, which, unlike R, is labelable on its own. For label weakness of R, see Chomsky (2015).

Abbreviations in (51): NOM = nominative; ACC = accusative; PST = past; CL = classifier.

27 But see also Mizuguchi (2017c, 2018), where it is argued that externalization of <C, T> is subject to variation across languages.

shown that A-movement in local subject Ā-movement is taken care of by external pair-Merge, proposing a new analysis of VMH. The discussion in the paper has led us to reconsider the EPP effect and I have argued that it is explained not only by labeling but also by externalization of ϕ.

The paper has demonstrated that Merge plays a key role in syntactic derivation: counter-cyclic A-movement is solved by Merge. It has also shown that externalization, which applies to derived SOs upon Transfer, interacts with syntax to explain subject extraction and the EPP effect. In conclusion, the present paper endorses the fundamental hypothesis in the Minimalist Program that language is explained by the structure-building operation Merge and interfaces with the CI and SM systems.

References

Berwick, Robert C. and Noam Chomsky. 2011. The biolinguistics program: The current state of its development. In Anna Maria Di Sciullo and Cedric Boeckx (eds.), *The biolinguistics enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Berwick, Robert C. and Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Boeckx, Cedric. 2011. Approaching parameters from below. In Anna Maria Di Sciullo and Cedric Boeckx (eds.), *The biolinguistics enterprise: New perspectives on the evolution and nature of the human language faculty*, 205–221. Oxford: Oxford University Press.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, MA: MIT Press.

Bošković, Željko. 2002. Expletives don't move. In Masako Hirotani (ed.), *Proceedings of the Thirty-Second Annual Meeting of the North East Linguistic Society*, 21–40. Amherst, MA: GLSA Publications.

Bošković, Željko. 2008. On successive cyclic movement and the freezing effect of feature checking. In Jutta M. Hartmann, Veronika Hegedüs, and Henk van Riemsdijk (eds.), *Sounds of silence: Empty elements in syntax and phonology*, 195–233. Amsterdam: Elsevier.

Bošković, Željko. 2016. On the timing of labeling: Deducing comp-trace effects, the subject condition, the adjunct condition, and tucking in from labeling. *The Linguistic Review* 33(1): 17–66.

Carstens, Vicki, Norbert Hornstein, and T. Daniel Seely. 2016. Head-head relations in

Problems of projection. *The Linguistic Review* 33(1): 67–86.
Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.
Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax 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. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures, Vol. 3*, 104–131. Oxford: Oxford University Press.
Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 1–29. Berlin: Mouton de Gruyter.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos. P. Otero, and Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2010. Some simple evo devo theses: How true might they be for language? In Richard K. Larson, Viviane Déprez, and Hiroko Yamakido (eds.), *The evolution of human language: Biolinguistic perspectives*, 45–62. Cambridge: Cambridge University Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.
Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann, and Simona Matteini (eds.), *Structures, strategies, and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins.
Chomsky, Noam. 2017. Language architecture and its import for evolution. *Neuroscience and Biobehavioral Reviews* 81: 295–300.
Chomsky, Noam, Ángel J. Gallego, and Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions, and challenges. Ms., MIT, Universitat Autònoma de Barcelona, and University of Ottawa.
Chomsky, Noam and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8(3): 425–504.
Chou, Tim. 2018. Composite probing and feature split: The case of English locative inversion. In Yoen-Mee Park (ed.), *Proceedings of the 20th Seoul International Conference on Generative Grammar*, 87–108. Seoul: The Korean Generative Grammar Circle.
Clahsen, Harald and Klaus-Dirk Smolka. 1986. Psychological evidence and the description

of V-second phenomena in German. In Hubert Haider and Martin Prinzhorn (eds.), *Verb second phenomena in Germanic languages*, 137–168. Berlin: Mouton de Gruyter.

Deal, Amy Rose. 2009. The origin and content of expletives: Evidence from "selection." *Syntax* 12(4): 285–323.

Diesing, Molly. 1990. Word order and the subject position in Yiddish. *Natural Language and Linguistic Theory* 8(1): 41–79.

Doherty, Cathal. 2000. *Clauses without 'that': The case for bare sentential complementation in English*. New York: Garland.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2012. Structure building that can't be. In Myriam Uribe-Etxebarria and Vidal Valmala (eds.), *Ways of structure building*, 253–270. Oxford: Oxford University Press.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2015. *Explorations in maximizing syntactic minimization*. New York: Routlege.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2016. Phase cancellation by external pair-merge of heads. *The Linguistic Review* 33(1): 87–102.

Erlewine, Michael Y. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34(2): 429–479.

Erlewine, Michael Y. 2017. The anti-locality signature of quirks of subject extraction. Paper presented at Workshop on Quirks of Subject Extraction, National University of Singapore, 10–11 August.

Erlewine, Michael Y. 2018. Extraction and licensing in Toba Batak. *Language* 94(3): 662–697.

Fukui, Naoki. 1995. *Theory of projection in syntax*. Stanford, CA: CSLI Publications.

Fukui, Naoki. 2017. *Merge in the mind-brain: Essays on theoretical linguistics and the neuroscience of language*. New York: Routledge.

Gallego, Ángel J. 2014. Deriving feature inheritance from the copy theory of movement. *The Linguistic Review* 31(1): 41–71.

Gallego, Ángel J. 2017. Remark on the EPP in labeling theory: Evidence from Romance. *Syntax* 20(4): 384–399.

Guasti, Maria Teresa and Luigi Rizzi. 2002. Agreement and tense as distinct syntactic positions: Evidence from acquisition. In Guglielmo Cinque (ed.), *Functional structure in DP and IP: The cartography of syntactic structures, Vol. 1*, 167–194. Oxford: Oxford University Press.

Henry, Alison. 1995. *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford: Oxford University Press.

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31(3): 445–483.

Hsu, Brian. 2016. Unification of feature scattering and m-merger as coalescence. In Christopher Hammerly and Brandon Prickett (eds.), *Proceedings of the Forty-Six Annual Meeting of the North East Linguistic Society*, 133–146. Amherst, MA: GLSA Publications.

Ishii, Toru. 2004. The phase impenetrability condition, the vacuous movement hypothesis, and *that-t* effects. *Lingua* 114: 183–215.

Ishii, Toru. 2017. Transfer and self pair-Merge. In Nicholas LaCara, Keir Moulton, and Anne-Michelle Tessier (eds.), *A schrift to fest Kyle Johnson*, 159–167. Amherst, MA: University of Massachusetts – Amherst, Linguistics Open Access Publications.

Kinyalolo, Kasangati Kikuni Wabongambilu. 1991. Syntactic dependencies and the spec-head agreement hypothesis in KiLega. PhD Dissertation. UCLA.

Landau, Idan. 2007. EPP extensions. *Linguistic Inquiry* 38(3): 485–523.

Lasnik, Howard 2004. The position of the accusative subject in the accusative-infinitive construction. In Peri Bhaskararao and Karumuri Venkata Subbarao (eds.), *Non-nominative subjects, Vol. 1*, 269–281. Amsterdam: John Benjamins.

Legate, Julie A. 2011. Under-inheritance. Paper presented at the 42nd North East Linguistic Society, University of Toronto, 11–13 November.

Legate, Julie A. 2014. *Voice and v: Lessons from Acehnese*. Cambridge, MA: MIT Press.

Martinović, Martina. 2015. Feature geometry and head splitting: Evidence from the morphosyntax of the Wolof clausal periphery. PhD Dissertation. The University of Chicago.

McCloskey, James. 1979. *Transformational syntax and model theoretic semantics: A case study in modern Irish*. Dordrecht: Reidel.

McFadden, Thomas and Sandhya Sundaresan. 2018. What the EPP and comp-trace effects have in common: Constraining silent elements at the edge. *Glossa: A Journal of General Linguistics* 3(1): 1–34.

Miyagawa, Shigeru. 2010. *Why agree? Why move? Unifying agreement-based and discourse-configurational languages*. Cambridge, MA: MIT Press.

Mizuguchi, Manabu. 2008. Derivation, minimalism, and *that-trace* effects. *English Linguistics* 25(1): 56–92.

Mizuguchi, Manabu. 2014. Consequences of feature inheritance for subject movement. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 325–334. Summerville, MA: Cascadilla Proceedings Project.

Mizuguchi, Manabu. 2016a. De-activation of ϕ through pair-Merge. *English Linguistics* 33(1): 108–118.

Mizuguchi, Manabu. 2016b. Simplest Merge, labeling, and A'-movement of the subject. In Miyoko Yasui and Manabu Mizuguchi (eds.), *Phase theory and its consequences: The initial and recursive symbol* S, 41–83. Tokyo: Kaitakusha.

Mizuguchi, Manabu. 2017a. Labelability and interpretability. *Studies in Generative Grammar* 27(2): 327–365.

Mizuguchi, Manabu. 2017b. Deducing Transfer from Merge. In Andrew Alexander Monti (ed.), *Proceedings of the 2017 Annual Conference of the Canadian Linguistic Association*, 1–12. Ottawa, Ontario: The Canadian Linguistic Association.

Mizuguchi, Manabu. 2017c. Interaction of Merge and labeling: Consequences for hyperraising. Paper presented at Western Conference on Linguistics 2017, Boise State University, 20–22 October.

Mizuguchi, Manabu. 2018. Merge, externalization and subject extraction. Paper presented at Arizona Linguistics Circle 12, University of Arizona, 12–14 October.

Ndayiragije, Juvénal. 1999. Checking economy. *Linguistic Inquiry* 30(3): 399–444.

Nomura, Masashi. 2004. Expletives move! In Brian Agbayani, Vida Samiian, and Benjamin V. Tucker (eds.), *Proceedings of the Thirty-Second Western Conference on Linguistics*, 207–220. Fresno, CA: Department of Linguistics, California State University, Fresno.

Nomura, Masashi. 2018. External pair-merge of heads as a core syntactic operation. Paper presented at the 26th Japanese/Korean Linguistics Conference, UCLA, 28 November–2 December.

Obata, Miki. 2010. Root, successive-cyclic and feature-splitting Internal Merge: Implications for feature-inheritance and Transfer. PhD Dissertation. University of Michigan.

Obata, Miki and Samuel D. Epstein. 2011. Feature-splitting internal merge: Improper movement, intervention, and the A/A′ distinction. *Syntax* 14(2): 122–147.

Oseki, Yohei. 2015. Eliminating pair-Merge. In Ulrike Steindl et al. (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, 303–312. Summerville, MA: Cascadilla Proceedings Project.

Ouali, Hamid. 2008. On C-to-T ϕ-feature transfer: The nature of agreement and anti-agreement in Berber. In Roberta D'Alessandro, Susann Fischer, and Gunnar Hrafn Hrafnbjargarson (eds.), *Agreement restrictions*, 159–180. Berlin: Mouton de Gruyter.

Park, Jong Un, Michael Barrie, and Myung-Kwan Park. 2019. External Remerge and linearization in ATB, RNR and PG constructions. *Linguistic Research* 36(2): 141–181.

Pesetsky, David. 1982. Complementizer-trace phenomena and the nominative island condition. *The Linguistic Review* 1(3): 297–344.

Pesetsky, David. 2016. Exfoliation: Towards a derivational theory of clause size. Paper presented at Workshop on Shrinking Trees, Universität Leipzig, 10 October.

Pesetsky, David. 2017. Complementizer-trace effects. In Martin Everaert and Henk C. van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax (second edition)*, 993–1026. Malden, MA: Wiley.

Pesetsky, David. 2019. Exfoliation: Towards a derivational theory of clause size. Ms., MIT.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.

Rizzi, Luigi and Ur Shlonsky. 2007. Strategies of subject extraction. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 125–154. Berlin: Mouton de Gruyter.

Ross, John Robert. 1970. On declarative sentences. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 222–272. Waltham, MA: Ginn and Company.

Sigurðsson, Halldór Á. 2010. On EPP effects. *Studia Linguistica* 64(2): 159–189.
Sugimoto, Yushi. 2016. De-phased by external pair-merge of C to T and its consequences. In Kate Bellamy, Elena Karvovskaya, and George Saad (eds.), *ConSole XXIV: Proceedings of the 24th Conference of the Student Organization of Linguistics in Europe*, 237–253. Leiden: Leiden University Centre for Linguistics.
Taraldsen, Knut Tarald. 1980. *On the nominative island condition, vacuous application and the* that-*trace filter*. Bloomington, IN: Indiana University Linguistics Club.
Taraldsen, Knut Tarald. 2002. The *que/qui* alternation and the distribution of expletives. In Peter Svenonius (ed.), *Subjects, expletives, and the EPP*, 29–42. Oxford: Oxford University Press.
Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1): 103–129.
Tourlakis, George. 2003. *Lectures in logic and set theory, Vol. 2: Set theory*. Cambridge: Cambridge University Press.
Uriagereka, Juan. 1988. On government. PhD Dissertation. University of Connecticut, Storrs.
Urk, Coppe van. 2015. A uniform syntax for phrasal movement: A case study of Dinka Bor. PhD Dissertation. MIT.
Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Manabu Mizuguchi
Professor
Department of Social Psychology
Toyo University
5-28-20 Hakusan Bunkyo-ku
Tokyo 112-8606, Japan
E-mail: mizuguchi@toyo.jp

Received: 2019. 02. 25.
Revised: 2019. 05. 31.
Accepted: 2019. 09. 23.