A-Movement: Successive Cyclic or One Fell Swoop?

Manabu Mizuguchi*

Toyo University

1 Introduction

The successive cyclicity of A-movement has been controversial and there are arguments for and against successive cyclic A-movement. Bošković (2002), for instance, discusses examples such as (1) and (2), arguing that the intended interpretations are possible only if the surface subject moves successive cyclically and is interpreted in intermediate positions, which are marked as underlines in the examples (Bošković 2002:179-180):

(1) a. Mary seems to John [__ to appear to herself to be in the room].
    b. *Mary seems to John [__ to appear to himself to be in the room].

(2) [His$_i$ mother's$_j$ bread] seems to every man$_i$ [__ to be known by her$_j$ to be the best there is].

On the other hand, Boeckx (2000), Castillo et al. (2009) and Epstein & Seely (2002, 2006) among others argue that the NP moves in a single leap in A-movement, claiming that alternative explanations are possible for examples like (1) and (2).

In this paper, I consider how A-movement proceeds in the derivation. Assuming the framework of the recent Minimalist Program, I claim that A-movement is not an "either-or" movement: it can be both successive cyclic and non-successive cyclic. I argue that whether A-movement proceeds successive cyclically or not depends on how Merge applies, showing that the operation plays a key role in determining the successive cyclicity.

2 Merge and A-Movement

A language yields a digitally infinite array of hierarchically structured expressions or syntactic objects (SOs) with systematic interpretations at interfaces with the Conceptual-Intentional system and the Sensory-Motor system. Given this basic property of language,

* I thank the reviewers for comments and suggestions. 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).

I assume that the ingredients of the language faculty are Merge and the interfaces (Chomsky 2010, 2017) and that in search of the simplest account of the language faculty, which is the basic hypothesis in an explanatory theory of language like the Minimalist Program, the properties of language are explained by these ingredients alone. Moreover, principles of Universal Grammar or UG, including Merge, are subject to the Strong Minimalist Thesis, SMT: UG operates in accord with language-independent conditions of computational efficiency (what are alternatively called "third factor principles") (Chomsky 2017:296). Given SMT, Merge is formulated as "simplest Merge," creating a set out of any two SOs (i.e., Merge(α, β) = {α, β}).

Given simplest Merge, Merge applies freely and can produce SOs shown in (3):

(3) a. {χ, {α, β}} b. {α, {α, β}}

In (3a), α and β are merged to form a set, with which another SO χ is merged to form another set; in (3b), on the other hand, Merge embeds the set formed out of α and β under α by taking α and recursively merging it with the set. Notice that the set produced in (3b) is an adjunction structure: β is put asymmetric to α due to α embedding the {α, β} set or β. Chomsky (2004, 2015) proposes a distinct operation called pair-Merge for adjunction, which, as shown in (4), produces an ordered pair, not a set:

(4) <α, β>

Given simplest Merge, Merge yields adjunction when applied in the way shown in (3b) and there is no need to assume pair-Merge in addition to simpest Merge (see Chomsky et al. 2019 for arguments in favor of this conclusion). This argument is endorsed by Tourlakis (2003), who shows that the ordered pair (4) is mathematically equivalent to the set (3b) (see also Fukui 2017 and Omune 2018). In this paper, unless otherwise specified, for ease of distinction and for illustration, I use the term "pair-Merge" for the merge that generates adjunction or the set (3b), and the ordered pair (4) is employed for (3b).

With (3) in place, now consider the derivation of (5):

(5) The student seems to be in the library.

I argue that pair-Merge explains the successive cyclicity of A-movement. Assuming that both C and T are elements of clausal architecture, with the two heads selected and merged in to form clauses (Chomsky 1981; Ormazabal 1995; Pesetsky 1992 among many others), simplest Merge allows two pair-Merge options (6a,b) in the embedded clause:

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

In (6a), T is externally pair-merged to C while in (6b), C is externally pair-merged to T. I

claim that the derivation under (6a) will make A-movement successive cyclic. <C, T> has the properties of C, with T being syntactically invisible and de-activated, since C embeds the {C, T} set, hence T: in other words, it is on a par with C (<C, T> = {C, {C, T}} = C). Given this and given that C has phasehood as one of its properties, <C, T> works as a phase head and its complement is cyclically transferred at the phase level. This means that in order to move out, the NP, as shown in (7), must go through the Spec of <C, T>; otherwise, it will be unavailable to the computation in the higher phase for phase impenetrability and cannot move out from the embedded clause:

(7) [NP [… [$_{\mu}$ t [$_{\delta}$ {C, {C, T}} [$_{\alpha}$ … t …]]]]]

Transfer

If Merge produces (6a), A-movement will be necessarily successive cyclic.

The phase status of a raising complement is endorsed by the impossibility of Quantifier Raising (QR) out of the raising infinitive. Consider (8) (Wurmbrand 2013:620):

(8) a. Mary seems to two women to be expected to dance with every senator.

b. #This soldier seems to someone to be likely to die in every battle.

In (8), a universal quantifier cannot take scope over an existential quantifier (*∀»∃). Wurmbrand (2013) argues that given Scope Economy (Fox 2000), this is due to the phase status of raising infinitives: QR is a syntactic movement (Takahashi 2010) and requires successive cyclic movement in order to raise out into the higher clause; successive cyclic QR to the phase edge, however, violates Scope Economy as it is semantically vacuous, and cannot be operated. ∀ cannot move out of the embedded clause for phase impenetrability. Her argument supports our claim that raising infinitives are phasal if Merge yields (6a).

Now consider A-movement under (6b). In this case, A-movement will proceed in a single leap. <T, C>, in which T embeds the {C, T} set, hence C, is on a par with T, having the properties of T, not C (<T, C> = {T, {C, T}} = T). T, unlike C, is not a phase head. Consequently, α, the complement of <T, C>, is not cyclically transferred and the NP can move to Spec,TP without undergoing successive-cyclic movement:

(9) [NP [… [$_{\delta}$ {T, {C, T}} [$_{\alpha}$ … t …]]]]

One piece of evidence for the non-phase status of <T, C> is shown by long distance agreement in (10), where the matrix T_{ϕ} agrees with the NP *three men* in the embedded clause without movement:

(10) There seem to be likely to be three men here. (Boeckx 2009:2)

Long distance agreement suggests that there are no phase boundaries in between and that α is not rendered impenetrable by cyclic Transfer.

The argument is also endorsed by Standard Arabic. In this language, post-verbal subjects trigger partial agreement (gender agreement) while preverbal subjects (i.e., subjects in Spec,TP) trigger full agreement (gender and number agreement). Consider (11):

(11) a. ʔakal-at (/*ʔakal-na) ṭ-ṭaalibaat-u.
eat.PAST-3.FEM.SG eat.PAST-3.FEM.PL the-students.FEM.PL-NOM
'The students ate.'
b. ṭ-ṭaalibaat-u ʔakal-na (/*ʔakal-at). (Benmamoun 1992:121)

With this in mind, now consider (12):

(12) a. ʔawʃakna (ʔan) tanjaħ(u/a) l-tˁa:liba:t-u.
were.about.to.3.F.PL (C/to) succeed.3.F.SG the-students.F-NOM
b. ʔawʃakat (ʔan) tanjaħ(u/a) l-tˁa:liba:t-u.
were.about.to.3.F.SG (C/to) succeed.3.F.SG the-students.F-NOM
'The female students were about to succeed.' (Alexiadou et al. 2014:2-3)

In (12a), full agreement appears in the matrix clause and it can be considered that the post-verbal NP, unlike in (12b), moves to the matrix Spec,TP but that its copy is pronounced (cf. Alexiadou et al. 2014). Notice that partial agreement emerges in the embedded clause in (12a) as well as in (12b), suggesting that the movement proceeds in a single leap thanks to lack of phase boundaries. (10) and (12) show that phase boundaries can be absent in the raising complement, and the non-phase status is explained by (6b).

Notice that under (6b), one fell swoop A-movement is not forced. Given simplest Merge, Merge can make A-movement successive cyclic even under (6b), which can only be barred by stipulation:

(13) [NP [… [$_{\mu}$ t [$_{\delta}$ {T, {C, T}} [$_{\alpha}$ … t …]]]]]

This derivation, however, is independently ruled out given third factor principles. Bošković (2019) argues that asymmetric relations are maximized in language, proposing MAR, which is traceable to labeling by minimal search (Dadan 2019):

(14) Maximize Asymmetric Relations (MAR) (Bošković 2019:2)

One consequence of (14), he argues, is that Spec positions are avoided as much as possible, which suggests that successive cyclic movement occurs only when it is really necessary, namely, when it is forced by phase impenetrability. In (13), phase impenetrability does not arise for lack of a phase head in the embedded clause, with the result that (13) violates (14).

Also, Chomsky (2019) and Chomsky et al. (2019) argue for Determinacy in terms of efficient computation and propose (15):

(15) Simplest Merge maps WS = [X, Y] onto WS′ = [{X, Y}], reducing its complexity

and avoiding indeterminate rule application.

Derivation is deterministic to the extent that derived structures are unambiguous to rule application in subsequent derivation. (13) will produce (16), in which Internal Merge or movement will ambiguously apply to the boxed NPs, and will be ruled out by Determinacy:

(16) WS = [__ [T [… [$_{\mu}$ NP [$_{\delta}$ <T, C> [$_{\alpha}$ … NP …]]]]]]

Both MAR and Determinacy are properties of language which follow from third factor principles. Given that UG is subject to SMT, (13) is excluded by efficient computation.

3 Movement and Agree Languages

In this section, I show that the analysis of A-movement proposed in this paper is cross-linguistically supported. Alexiadou et al. (2014) (henceforth, AAW) argue on the basis of raising constructions that there are two types of languages: movement languages and Agree languages. In movement languages, the NP must move to Spec,TP because there are phase boundaries between T and the original position of the NP and long distance agreement is blocked; on the other hand, in Agree languages, the NP can agree with T in the in-situ position in the absence of phase boundaries. According to AAW, Greek is an Agree language. Consider (17) (Alexiadou et al. 2014:3-4):

(17) a. Arhise {***pikni**} [na skepazi **i skoni** {**pikni**} ta epipla].
started dense.FEM SUBJ cover.3SG the dust dense the furniture
'The dust started to cover the furniture densely.'

b. Stamatise na perni mono i Maria kakus vathmus.(*stop*»*only*; **only*»*stop*)
stopped SUB get only Mary grades weak
'It stopped being the case that only Maria got bad grades.'

In (17), the NP cannot license a matrix modifier nor can it take scope over the matrix predicate, which AAW argue suggests that the NP agrees with T in situ without movement for lack of phase boundaries. I argue that movement and Agree languages are explained by (6): in movement languages, the higher R, which is realized as V, selects a head with the properties of C, hence (6a), whereas in Agree languages, it selects the one with the properties of T, hence (6b). Notice that selection is required in some form or other in order to constrain head-complement relations. Given this, the resort to it is not at all a stipulation.

Provided that movement and Agree languages are explained by (6) and that whether a language is a movement or Agree language depends on selection, the prediction is that there are languages whose predicates can select both (6a) and (6b). I suggest that English is one such language. In English, as evidenced by (1)/(2) and (10), both successive cyclic

A-movement and long distance agreement are observed, which suggests that both (18a) and (18b) are allowed in the language:

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

Under the proposal in this paper, movement and Agree languages discussed in AAW follow from the way Merge applies in derivation. In (18a) or (6a), phases emerge while in (18b) or (6b), they do not. AAW, on the other hand, propose that the presence or absence of phases is derived by a particular type of selection of the raising complement. On the assumption that the highest projection of a cyclic domain constitutes a phase, which includes TP as well as CP (Bošković 2014), they argue that an obligatory selectional valuation relation between the matrix R and the highest head in the embedded clause voids the phasehood of the raising complement. However, this argument just restates that the raising complement is not a phase. The analysis proposed in this paper answers the question why a particular type of selection of the raising complement leads to phase cancellation: the raising predicate in the matrix clause selects T with C pair-merged to it (i.e., (6b)), which does not work as a phase head and hence does not induce phase impenetrability.

4 Implications

In this section, I discuss three implications of the proposed analysis. The first is that Merge can not only expand the clause; it can also reduce the clause. On the assumption that both C and T are ingredients of clausal architecture, if C and T are merged as in (19), we get a full-sized clause in that there are two sets above the verbal set α in the clausal domain; on the other hand, if C and T are pair-merged as in (20), there is only one set above α and we get a smaller-than-full clause:

(19) [$_{\kappa}$ C [$_{\lambda}$ T [$_{\alpha}$ …]]] (20) [$_{\delta}$ <C, T>/<T, C> [$_{\alpha}$ …]]

The discussion here argues that both clause expansion (=(19)) and clause reduction (=(20)) are explained by Merge: the clause size follows from how heads are merged in derivation.

This implication shows that Merge can derive clause-type differentiation. As noted in the literature, non-finite complements are different from finite ones in that they allow subject extraction (=(21), (22)); likewise, the NP can move out of the raising complement while it cannot out of the control complement (=(23)). Consider the following examples:

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

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

(23) a. *The student was tried to read the book.

b. The professor is likely to come to the conference.

It has been argued that this is due to different clause types; in the raising/ECM complement, CP is missing and lack of CP makes subject extraction possible. To explain the reduction, Chomsky (1986) lexically stipulates that the raising/ECM complement is TP while Chomsky (1981) and Pesetsky (2019) argue that CP is syntactically deleted in the course of derivation. Notice that these assumptions to explain clause-type differentiation are not necessary given the proposal here. Merge can explain clause reduction, hence deriving unique properties of raising/ECM complements.

The second implication is that T can label on its own. Chomsky (2015) claims that T is weak as a label and can label when it has overt/visible Spec, which he argues explains the EPP (Chomsky 1981, 1982). Unless overt/visible Spec,TP is created as in (24), a T-headed set or λ will not be labeled due to label weakness of T, which violates Full Interpretation:

(24) a. [$_{\lambda}$ T [be written the book by the professor]] (λ = ?)

b. [$_{\gamma}$ The book [$_{\lambda}$ T [be written t by the professor]]] (λ = T)

Recall that the ordered pair <T, C> is syntactically on a par with T and has the properties of T (<T, C> = {T, {C, T}} = T). Then if T is weak as a label, δ will not be labeled in the absence of overt/visible Spec in (25):

(25) [NP [… [$_{\delta}$ <T, C> [$_{\alpha}$ … t …]]]] (δ = ?)

As far as the discussion in this paper is correct, it suggests that T can label on its own.

This implication is desirable in that it can eliminate a strong/weak distinction on the labelability of T, which, as has been pointed out in the literature, runs the risk of stipulation (see Gallego 2017, Goto 2017 and Mizuguchi 2017 for problems with such a distinction). For instance, Chomsky (2015) argues that null-subject languages like Italian have strong T, meaning that it can label without overt/visible Spec. Likewise, discussing (26), Epstein et al. (2014) argue that non-finite T, unlike finite T, can label on its own, with α labeled:

(26) There is likely [$_{\alpha}$ to be a man in the room].

All these arguments do not explain, which argues that the distinction is just stipulated.

Labelable T straightforwardly follows without running the risk of stipulation if labelability is explained by interpretability (i.e., Full Interpretation). In my earlier work (Mizuguchi 2017), I proposed (27), which says that heads are labelable if and only if they are without unvalued features:

(27) Heads can label only when they are without unvalued features.

Given (27), T, hence <T, C>, can label since it is free from unvalued features including ϕ.

The final implication is that the proposed analysis can give an explanation to *wager*-class sentences. Consider examples from (28) to (30):

(28) a. *John wagered [Mary to have entered the room].
b. *John admitted [Mary to have entered the room].

(29) a. Mary was wagered [to have won the race].
b. Mary was admitted [to have won the race].

(30) The professor believes [the student to be the most intelligent].

In ECM sentences, the NP can halt in the embedded clause (Lasnik 2001, 2004) while in *wager*-class sentences, it cannot (Bošković 1997; Pesetsky 1992; Postal 1974, 1993). This peculiar behavior of *wager*-class sentences will follow if the relevant predicates select only (6b). Chomsky (2013) argues that one way to label XP-YP structure is by agreement: symmetric XP-YP structure can be labeled as far as X and Y agree. If (6b) is created in the *wager*-class infinitive, labeling will fail for lack of agreement. Recall that <T, C> is on a par with T as C is embedded under T. Since T does not bear ϕ-features of its own but inherits them from C, <T, C> cannot agree with the NP. As shown in (31a), μ, the XP-YP created by the movement of the NP, cannot be labeled, violating Full Interpretation. It follows that the NP cannot halt in the embedded clause or in the Spec of <T, C>. On the other hand, <C, T> is on a par with C and can agree with the NP for having ϕ. If (6a) is created in ECM, then μ can be labeled thanks to agreement and the NP can halt in the embedded clause (Mizuguchi 2019):

(31) a. $[_{\mu}$ NP $[_{\delta}$ <T, C> $[_{\alpha}$ … t …]]] (μ = ?)
b. $[_{\mu}$ NP $[_{\delta}$ <C, T> $[_{\alpha}$ … t …]]] (μ = LABEL)

5 Conclusion

Summarizing the discussion, I have argued in this paper that A-movement can be both successive cyclic and non-successive cyclic; it is not an "either-or" movement and it can be both successive cyclic and non-successive cyclic. I have claimed that whether A-movement proceeds successive cyclically or not depends on how Merge applies. To the extent that the proposed analysis is correct, it has implications for clausal construction, the labelability of T and *wager*-class sentences. In conclusion, this paper demonstrates that Merge plays a key role in syntactic derivation, which supports the minimalist hypothesis that the operation is the core of language.

References

Alexiadou, A., Anagnostopoulou, E., & Wurmbrand, S. (2014). Movement vs. long distance Agree in raising: Disappearing phases and feature valuation. In H.-L. Huang, E. Poole, & A. Rysling (Eds.), *Proceedings of the Forty-Third Annual Meeting of the North East Linguistic Society* (pp. 1–12). University of Massachusetts, Graduate Linguistics Student Association (GLSA).

Benmamoun, E. (1992). *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press.

Boeckx, C. (2000). EPP eliminated. Ms., University of Connecticut, Storrs.

Boeckx, C. (2009). On long-distance Agree. *Iberia: An International Journal of Theoretical Linguistics*, *1*, 1–32.

Bošković, Ž. (1997). *The syntax of nonfinite complementation: An economy approach*. MIT Press.

Bošković, Ž. (2002). A-movement and the EPP. *Syntax*, *5*, 167–218.

Bošković, Ž. (2014). Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry*, *45*, 27–89.

Bošković, Ž. (2019). Generalized asymmetry. Ms., University of Connecticut, Storrs.

Castillo, J. C., Drury, J. E., & Grohmann, K. K. (2009). Merge over Move and the Extended Projection Principle: MOM and the EPP revisited. *Iberia: An International Journal of Theoretical Linguistics*, *1*, 53–114.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures*. Foris.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Praeger.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *The cartography of syntactic structures, Volume 3: Structures and beyond* (pp. 104–131). Oxford University Press.

Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (pp. 45–62). Cambridge University Press.

Chomsky, N. (2013). Problems of projection. *Lingua*, *130*, 33–49.

Chomsky, N. (2015). Problems of projection: Extensions. In E. Di Domenico, C. Hamann, & S. Matteini (Eds.), *Structures, strategies, and beyond: Studies in honour of Adriana Belletti* (pp. 1–16). John Benjamins.

Chomsky, N. (2017). Language architecture and its import for evolution. *Neuroscience and Biobehavioral Reviews, 81*, 295–300.

Chomsky, N. (2019). Some puzzling foundational issues: The Reading program. *Catalan Journal of Linguistics Special Issue, 2019*, 263–285.

Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics Special Issue, 2019*, 229–261.

Dadan, M. (2019). Loss of movement and labeling: Grammatical pressure and diachronic change. In R. Stockwell et al. (Eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics* (pp.109–114). Cascadilla Proceedings Project.

Epstein, S. D., Kitahara, H., & Seely, T. D. (2014). Labeling by minimal search: Implications for successive-cyclic A-movement and the conception of the postulate "phase." *Linguistic Inquiry, 45*, 463–481.

Epstein, S. D. & Seely, T. D. (2002). Rule applications as cycles in a level-free syntax. In S. D. Epstein & T. D. Seely (Eds.), *Derivation and explanation in the minimalist program* (pp. 65–89). Blackwell.

Epstein, S. D. & Seely, T. D. (2006). *Derivations in minimalism*. Cambridge University Press.

Fukui, N. (2017). *Merge in the mind-brain: Essays on theoretical linguistics and the neuroscience of language*. Routledge.

Fox, D. (2000). Economy and semantic interpretation. Ph.D. dissertation, MIT.

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

Goto, N. (2017). Eliminating the strong/weak parameter on T. In M. Y. Erlewine (Ed.), *MIT working papers in linguistics 85*: *Proceedings of GLOW in Asia XI, Volume 2* (pp. 57–71). MIT Working Papers in Linguistics.

Lasnik, H. (2001). Subjects, objects, and the EPP. In W. D. Davis & S. Dubinsky (Eds.), *Objects and other subjects: Grammatical functions, functional categories and configurationality* (pp. 103–121). Kluwer.

Lasnik, H. (2004). The position of the accusative subject in the accusative-infinitive construction. In P. Bhaskararao & K. V. Subbarao (Eds.), *Non-nominative subjects* (pp. 269–281). John Benjamins.

Mizuguchi, M. (2017). Labelability and interpretability. *Studies in Generative Grammar, 27*, 327–365.

Mizuguchi, M. (2019). Optional raising in ECM and labeling of XP-YP. *Studies in Generative Grammar*, *29*, 373–411.

Omune, J. (2018). Reformulating pair-merge of heads. *English Linguistics*, *34*, 266–301.

Ormazabal, J. (1995). The syntax of complementation: On the connection between syntactic structure and selection. Ph.D. dissertation, University of Connecticut, Storrs.

Pesetsky, D. (1992). Zero syntax, Volume 2: Infinitives. Ms., MIT.

Pesetsky, D. (2019). Exfoliation: Towards a derivational theory of clause size. Ms., MIT.

Postal, P. (1974). *On raising: One rule of English grammar and its theoretical implications*. MIT Press.

Postal, P. (1993). Some defective paradigms. *Linguistic Inquiry*, *24*, 357–364.

Takahashi, M. (2010). Case, phases, and nominative/accusative conversion in Japanese. *Journal of East Asian Linguistics*, *19*, 319–355.

Tourlakis, G. (2003). *Lectures in logic and set theory, Volume 2: Set theory*. Cambridge University Press.

Wurmbrand, S. (2013). QR and selection: Covert evidence for phasehood. In S. Keine & S. Sloggett (Eds.), *Proceedings of the Forty-Second Annual Meeting of the North East Linguistic Society* (pp. 619–632). University of Massachusetts, Graduate Linguistics Student Association (GLSA).

Manabu Mizuguchi Toyo University

mizuguchi@toyo.jp

Coyote Papers Volume 22 (2019)