Late Merge and Cyclicity

Mariia Privizentseva

Abstract

The goal of this paper is to explore how late Merge that is often used to derive anti-connectivity effects relates to different conceptions of cyclicity such as the Extension Condition, the Strict Cycle Condition, the Earliness Principle, and the Featural Cyclicity. I demonstrate that late Merge can be implemented under any of these restrictions on cyclicity if further assumptions on the architecture of syntax are made. I further investigate the contexts where late Merge becomes possible and show that some of the reviewed models overgenerate while others undergenerate.

1. Background

Late Merge is a theoretical tool used to derive anti-connectivity effects, i.e., cases where despite an expected presence of a syntactic object in a certain position, this syntactic object behaves as though it were absent from this position with respect to a number of effects. One such effect is condition C. It requires R-expressions to be free, that is, not bound by a coindexed syntactic object (see Chomsky (1981)). According to a widely acknowledged point of view, A-moved syntactic objects as well as adjuncts of Ā-moved syntactic objects obviate condition C (see Van Riemsdijk and Williams (1981), Lebeaux (1988, 1990), Fox (1999), Bhatt and Pancheva (2004), Hulsey and Sauerland (2006), Takahashi and Hulsey (2009), Van Urk (2015), Keine and Bhatt (2019), and Gong (2022)). This is illustrated in (1)-(2). The sentence in (1) presents raising to subject, an instance of A-movement. It shows that *John* embedded in the moved constituent and *him* can be co-indexed, thereby obviating a condition C violation that would take place between the pronoun and *John* if the latter were present in the base position of the raised subject.

(1) [These pictures of John$_i$]$_j$ seemed to him$_i$ [__$_j$ to be very good].

Cyclicity, 329–350
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

Examples in (2a-b) show that the empirical picture is different for Ā-moved phrases. As in (1), *John* in (2a) is a complement of preposition *of*, but it is evaluated for condition C in the base position of the displaced constituent in this case. As a result, coreference with the personal pronoun *he* is ruled out. Example (2b) differs in that *John* is an adjunct in the displaced constituent, and it obviates a condition C violation just like a complement of A-moved phrase in (1).

(2) a. ?*[Which pictures of John$_i$]$_j$ did he$_i$ like __$_j$?
b. [Which pictures near John$_i$]$_j$ did he$_i$ look at __$_j$?
(Lebeaux 1990: 320)

The approach that relies on late Merge takes condition C obviation at face value and assumes that syntactic objects showing no connectivity with respect to some position are in fact absent from this position. This means that *John* in examples (1) and (2b) is simply not present in the base position of the noun phrase containing it and is merged late as shown in (3). *John* is therefore never c-commanded by the pronoun, and condition C is respected throughout the derivation.

(3) a. [... [$_{XP}$ XP]]

b. [[$_{XP}$ XP YP] ... ~~[$_{XP}$ XP]~~]
Late Merge

Despite the ability to account for anti-connectivity in a straightforward manner, late Merge is not universally accepted; it is widely criticized for violating cyclicity (cf. Chomsky (2019)). The goal of this paper is to investigate how late Merge can be implemented under different conceptions of cyclicity. I will consider four common views on cyclicity as they are defined below: the Earliness Principle, the Featural Cyclicity, the Strict Cycle Condition, and the Extension Condition.

(4) Earliness Principle (EP):
An uninterpretable feature must be marked for deletion as early in the derivation as possible. (see Pesetsky (1989) and Pesetsky and Torrego (2001))

(5) Featural Cyclicity (FC):
A feature must be checked as soon as possible after being introduced into the derivation. (see Chomsky (1995) and Richards (1999, 2001))

(6) Strict Cycle Condition (SCC):
Within the current domain δ, no operation may affect solely a proper subdomain γ that is dominated by δ. (see Chomsky (1973, 1995, 2019) and Müller (2011, 2014) for this formulation)

(7) Extension Condition (EC):
A syntactic derivation can only be continued by applying operations to the root of the tree. (see Chomsky (1993, 1995) and Adger (2003: 75) for this formulation)

When considering their relation to late Merge, I will assume an approach to syntax under which all instances of Merge, including late Merge, are driven by features. Following Heck and Müller (2007), I will indicate features that trigger Merge as [•F•] and features that trigger Agree as [∗F∗]. I will show that late Merge can, in principle, be incorporated under all approaches to cyclicity if further assumptions are made. In particular, ordering of features allows a delayed discharge of merge features and thereby makes room for late Merge under EP and FC. SCC and EC impose more rigid restrictions on Merge, but they can be circumvented if movement involves Merge of a copy to the workspace (see Nunes (2004) and Heck (2016, 2023)).

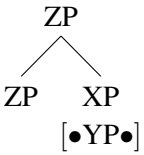
In what follows, I will start with the EP/FC in section 2, then turn to the SCC/EC in section 3, and summarize in section 4.

2. Late Merge and the Earliness Principle / Featural Cyclicity

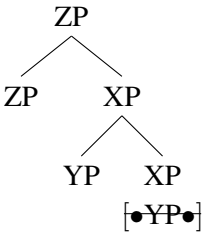
While the EP and FC were proposed independently from each other, they impose identical restrictions on syntax and require syntactic operation-triggering features to be discharged as early as possible. In syntax, where features on syntactic objects are not ordered with respect to each other, this means that active features will be discharged when their target is available. This leaves no room for late Merge: Targets for external Merge are usually available without restrictions so that corresponding selection features will be discharged as soon as their host enters the derivation. Thus, a derivation in (8)-(9), where the

selection feature waits some steps before it is discharged and thereby gives raise to late Merge, is excluded by the EP/FC.

(8) Step 1: Merge XP and ZP



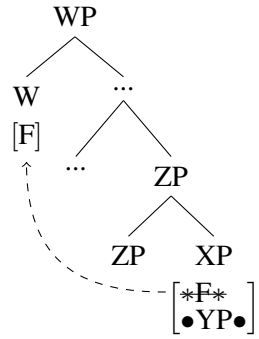
(9) Step 2: Late Merge of YP



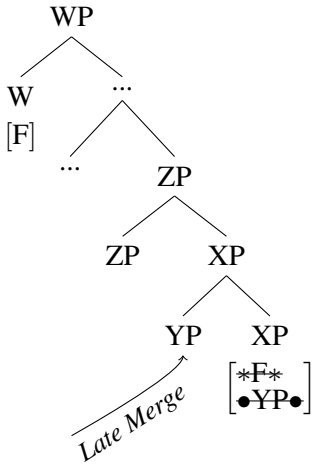
The state of affairs is different if unsatisfied syntactic features are assumed to be ordered (see Stabler (1997) and Müller (2011)), and if Merge and Agree features can be interleaved. In that case, only one feature appears on the top of the stack and can be active. Features ordered after it must wait until this feature is discharged. This introduces an additional condition on the discharge of Merge and Agree features and allows late Merge to be implemented in a way compatible with the EP/FC.

Consider the sample derivation below. In (10), the Merge feature is ordered after an agreement feature that does not find its goal in the c-commanding domain. Assuming the possibility of upward Agree (see Wurmbrand (2012), Zeijlstra (2012), and Bjorkman and Zeijlstra (2019) among others), the agreement probe waits several steps of the derivation until its Goal enters the derivation and is then discharged by probing upwards. After this, a new active feature may appear on the top of the stack. In (11), this is a Merge feature, so that late Merge takes place.

(10) Step 1: Upward Agree



(11) Step 2: Late Merge of YP



Note that late Merge as in (11) is compliant to the EP/FC: These principles require an active syntactic feature to be discharged as soon as possible, but the ordering of the Merge feature after the Agree feature ensures that the former cannot be discharged earlier in the derivation. Merge can (and following the EP/FC must) take place after an Agree feature is deactivated. In result, the syntactic model that assumes feature ordering and is restricted by the EP/FC enables late Merge.

Late Merge as in (11), however, is peculiar and differs from most proposed cases of late Merge in that a phrase within which late Merge applies remains deeply embedded. Being developed for deriving anti-connectivity effects of moved phrases, late Merge typically occurs in configurations where a constituent, within which late Merge takes place, itself moves up to the specifier of the highest projection, see the derivation in (12a) for sentence (12b) repeated from above.

(12) a. $[\ [_{DP}\ DP\ \underline{PP}\ ]_j\ ...\ [\ \sout{_{DP}\ DP\ ]_j}\ ]$

Late Merge

b. $[_{DP}$ Which pictures $[_{PP}$ near John$_i$ $]\ ]_j$ did he$_i$ look at $[_{DP}$ ~~which pictures~~ $]_j$?

Late Merge as in (11) that applies without movement of a constituent targeted by late Merge can have syntactic effects. Consider the two derivations schematized in (13) and (14). In (13), WP binds into YP before WP moves out of the local binding domain αP. The derivation in (14) differs in that YP is merged late, after WP has moved out of αP. As a result, a material in YP cannot be bound by WP.

(13) WP binds YP before movement: [WP ... [$_{\alpha P}$ ... WP ... X YP]]
2. Move *1. Bind*

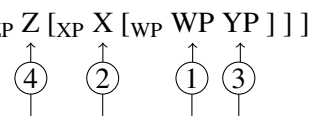
(14) Late Merge counterfeeds binding: [WP ... [$_{\alpha P}$ ... WP ... X YP]]
1. Move *2. Late Merge*

To the best of my knowledge, such data are rare if existent (see Costa (2000) for one example) and are not analyzed via late Merge. Instead, WP would be most likely analyzed as being first-merged outside of αP and further compelling arguments would be required to postulate a base position within αP from which WP cannot bind. If attested, however, such data could provide an argument in favor of the EP/FC and the implementation of late Merge suggested in the section.

Another group of cases where late Merge without movement of the constituent targeted by late Merge (as in (11)) can have an effect in the derivation arises if a derivational definition of c-command in (15) is assumed.

(15) Derivational C-Command
X c-commands all and only the terms of the category Y with which X was concatenated by Merge or Move in the course of the derivation. (see Epstein et al. (1998))

In that case, there can be no c-command relation between a late merged syntactic object and any above material that is introduced in the derivation before it. Compare the structure in (16), where YP is late merged after X but before Z is introduced. As a result, despite showing the same structural relations in the final representation, according to (15) X does not c-command YP, while Z does.

(16) 

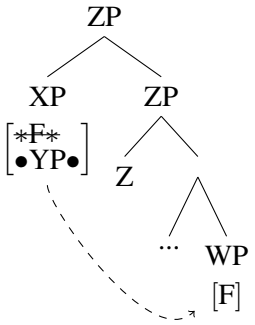
Epstein et al. (1998) use derivational c-command to account for data as in (17), where personal pronoun *she* can be co-referent with *Mary* and thus should not c-command it at any stage of the derivation, while *every student* must c-command *he* to allow for bound variable interpretation.

(17) [$_{DP}$ Which paper [$_{CP}$ that he$_i$ gave to Mary$_j$]] did every student$_i$ think that she$_j$ would like _ ?

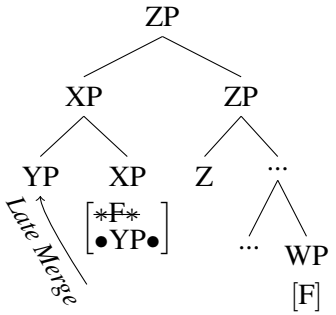
Given derivational c-command all c-command relations are as required if the relative clause is late merged after the pronoun *she* enters the derivation but before the noun phrase *every student* is introduced. This application of late Merge, thus, does not involve movement of the host and is, in this respect, identical to late Merge that was shown to be possible under the EP/FC plus upward Agree. One complication however comes from the fact that the feature responsible for Merge of the relative clause in (17) must be ordered after an agreement probe. This agreement probe must be checked by a syntactic object above *she* but below *every student* to ensure the correct timing of late Merge,. It is not immediately clear what probe it could be in this case.

The model relying on the EP/FC becomes more restrictive if search only applies downwards. This implies that only agree features on specifiers and heads of a topmost projection can be discharged. Consequently, a merge feature that is shielded by an agree feature earlier in the derivation can occur on the top of the stack only after its projection has moved to the specifier position; see (18).

(18) Step 1: Downward Agree



(19) Step 2: Late Merge of YP

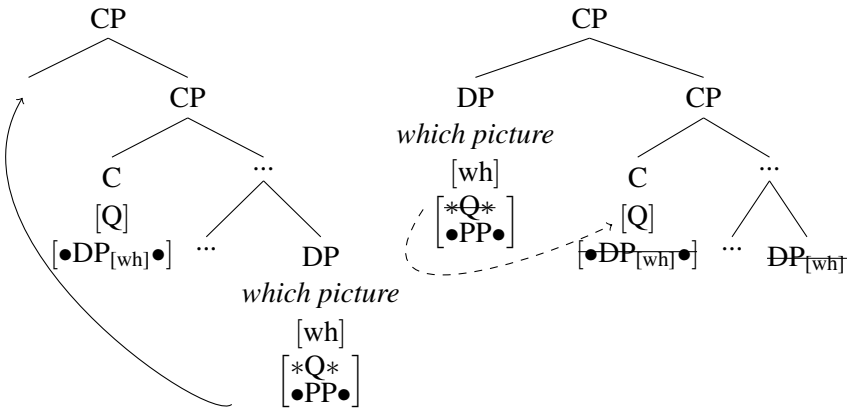


The XP targeted by late Merge in this derivation may be either first merged as a specifier of ZP or moved to this position. In the first case, the presence or absence of the late Merge of YP has no further effect. In the second case, the derivation is the one that is typically proposed to account for the anti-connectivity effects, so let us see how it applies to the actual data. I will start with adjuncts of Ā-moved wh-phrases as in (20) .

(20) Which pictures near John$_i$ did he$_i$ look at _ ?

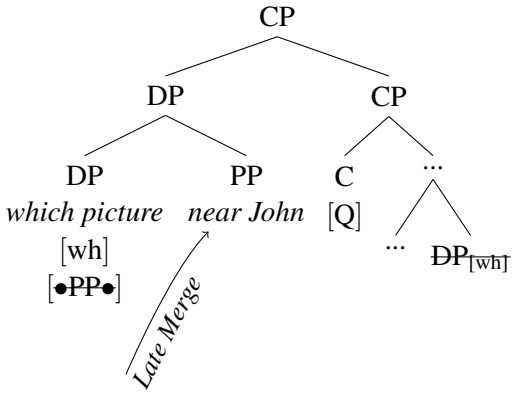
As discussed earlier, since *John* allows co-reference with the personal pronoun c-commanding the base position of the wh-phrase, the adjunct *near John* must be merged only after movement of the wh-phrase. Its delayed Merge is derived if corresponding Merge feature [•PP•] follows an active agreement probe that can be discharged only in the landing position of the wh-phrase. I assume that there are two features building up a wh-dependency: the [•DP$_{[wh]}$•] feature on the C head and the [∗Q∗] probe on a wh-phrase. The first feature is satisfied by movement of the wh-phrase to the specifier of the C head, which in turn creates a context for the discharge of the agreement probe on the wh-phrase (see (21)-(22)). Crucially, note that the [∗Q∗]-probe cannot be checked before movement of the wh-phrase to Spec,CP, because Agree is assumed to apply only downwards.

(21) Step 1: Wh-movement (22) Step 2: Agreement



The [•PP•] responsible for the Merge of the adjunct *near John* is ordered after the [∗Q∗] probe. Therefore, it occurs on the top of the stack and can be subsequently discharged only after wh-movement and checking of [Q] (see (23)). This derives late Merge of the adjunct.

(23) Step 3: Late Merge of the adjunct

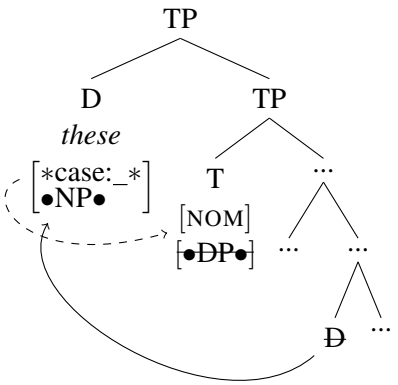


I will next turn to anti-connectivity attested for arguments of A-moved phrases as in (24).

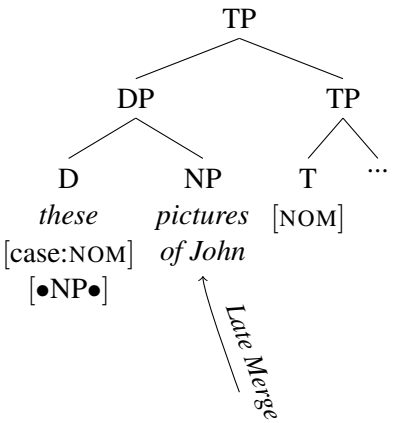
(24) [These pictures of John$_i$]$_j$ seemed to him$_i$ [_$_j$ to be very good].

To account for these data, it is assumed that the whole noun phrase *[pictures of John]* is late-merged (see Takahashi and Hulsey (2009)) and only the D head *these* is present in the base position. As in the previous derivation, late Merge is ensured here by the ordering of the corresponding selection feature ([•NP•] in this case) after the agreement probe that is discharged only in the landing position of the DP. The unchecked case feature plays the role of this agreement probe in the current derivation; see (25)-(26).

(25) Steps 1-2: Move and Agree



(26) Step 3: Late Merge



Interestingly, the distribution of late Merge under the EP/FC plus downward agree is similar to the distribution of late Merge in a model where the application of Merge is restricted by yet another cyclicity principle: the Peak Novelty Condition (PNC). The PNC was introduced by Safir (2019), and unlike EP/FC, which require the earliest possible discharge of active features, it imposes a restriction on the effect each application of merge must have in the derivation.

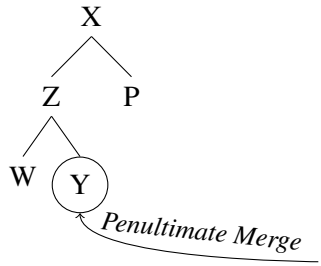
(27) Peak Novelty Condition
After every instance of Merge, M_i, the undominated node U of the resulting structure immediately dominates a node that U did not immediately dominate before M_i. (see Safir (2019: 292))

The PNC is satisfied by a regular Merge to a root node because a completely new root node is created and one of the two nodes it immediately dominates is

introduced in the course of this merge step. The PNC also rules in a so-called penultimate Merge. In this case, a new syntactic object is not merged to a root node, but to a node immediately dominated by the root. Following Safir (2019), such Merge satisfies the PNC because it also changes the identity of a node dominated by the root. Penultimate Merge is illustrated in (28)-(29).

(28) Before Merge

(29) Penultimate Merge



The possibility of penultimate Merge relies on the assumption that node W in (28) is different from node Z in (29). While this is automatically the case according to Safir (2019), the identity of node Z depends on the approach to labeling and the relation between W and Y. For instance, a widely adopted projection by selection labeling algorithm states that the label of a newly created syntactic object is determined by the syntactic object that selects (see Chomsky (1995), Adger (2003) as well as Stabler (1997)). As a result, if W selects for Y in (28), Z is equal to W, the identity of a node dominated by the root remains the same, and the PNC is not fulfilled. Thus, not all instances of penultimate Merge are automatically included under the PNC.[1]

This technical issue notwithstanding, the distribution of Merge under the PNC is similar to the one under the EP/FC plus downward Agree in that in addition to the regular Merge with a root node, Merge can target a node immediately dominated by the root node. This allows us to derive anti-connectivity effects via late Merge.[2]

However, both the PNC and the EF/FC plus downward Argee might be too restrictive to account for all attested cases of anti-connectivity. In particular, Sportiche (2019) has most recently argued that some of the data require late

[1]A possible objection would be that even though the labels of W in (28) and Z in (29) are the same, the two nodes cannot be identical and differ at least in a number of active merge features.

Merge to an arbitrarily deeply embedded node within the moved phrase. He provides example (30) as evidence:

(30) [Whose criticism of [Mary's rendition of (...) the claim [that you [formulated (...) the hypothesis [that Henri [visited the villages near Picasso$_i$'s estate]]]]]$_k$ did he$_i$ endorse __$_k$?

This example is peculiar in that an adjunct which does not show connectivity is embedded into several complements of an Ā-moved phrase. Following earlier empirical conclusions (see Lebeaux (1988, 1990), Fox (1999), Takahashi and Hulsey (2009), as well as most recently Stockwell et al. (2021, 2022)), Ā-moved phrases and their complements obligatorily show connectivity and therefore should not be targeted by late Merge. Takahashi and Hulsey (2009) account for this by imposing further restrictions on late Merge. First, they suggest that nouns must be present in a position where they get case. This excludes late Merge of the noun phrase *[criticism of ...]* to a displaced wh-operator *whose* in the example above. Second, late Merge is restricted by interpretability at LF. Derivation remains interpretable if a late-merged syntactic object is an adjunct, because adjuncts are attached by Predicate Modification, or if it is a restrictor of a moved operator/determiner, because restrictors are supplied to lower copies of an operator by the Variable Insertion operation in any case (see Trace Conversion proposed by Fox (1999)). Late Merge of a complement renders the structure uninterpretable and is therefore excluded.

All in all, independently of an account, if only adjuncts of a wh-moved phrase can be late merged, then for late Merge to derive anti-connectivity in (30), it must apply unboundedly deep within the displaced wh-phrase. Such applications are prohibited by the EP/FC plus downward Agree as well as by the PNC, and thus both models undergenerate. One possible solution would be to reconsider the original observation that anti-connectivity affects only adjuncts of Ā-moved phrases (see Adger et al. (2017), Bruening and Al Khalaf

[2]Note that the distribution of Merge under the EP/FC plus only downward Agree and under the PNC are not completely identical. They differ in that under the former model, late merge on a deeper level of embedding is not excluded by definition and is, in fact, possible for multiple specifiers of one head; cf. structures [$_{XP}$ WP [$_{XP}$ RP [$_{XP}$ YP X]]] and [$_{ZP}$ [$_{XP}$ WP [$_{XP}$ RP [$_{XP}$ YP X]]] [$_{ZP}$ UP Z]], where merge of YP after both WP and RP is allowed under the EP/FC plus only downward Agree but not under the PNC.

(2019), and Wierzba et al. (2020) for resent research casting doubts on earlier empirical results).

To sum up, in this section, I have shown that syntax governed by the EP/FC in combination with feature ordering allows late Merge to be implemented. I have shown that the model overgenerates and permits late Merge in seemingly unattested configurations if Agree can apply upwards. The model is more restrictive if only downward Agree is possible. In that case, it can account for most cases of anti-connectivity but potentially undergenerates or requires some other restrictions on application of late Merge to be reviewed.

3. Late Merge and Strict Cycle Condition / Extension Condition

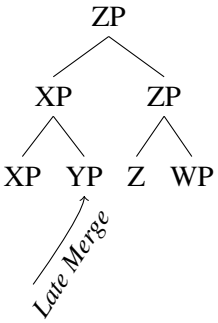
In this section, I will turn to the two stronger notions of cyclicity: the SCC and the EC. The definitions I will rely on here are repeated below.

(31) Strict Cycle Condition (SCC):
Within the current domain δ, no operation may affect solely a proper subdomain γ that is dominated by δ. (see Chomsky (1973, 1995, 2019) and Müller (2011, 2014) for this formulation)

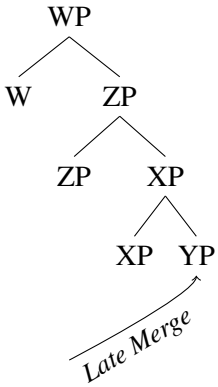
(32) Extension Condition (EC):
A syntactic derivation can only be continued by applying operations to the root of the tree. (see Chomsky (1993, 1995) and Adger (2003: 75) for this formulation)

Similarly to the EP and the FC, the SCC and the EC impose essentially identical restrictions on syntax and prohibit operations that apply not to the root of the exisiting structure. They exclude late Merge as in (33) and (34), because it involves Merge to a node XP properly included in the root domain.

(33) Late Merge of YP

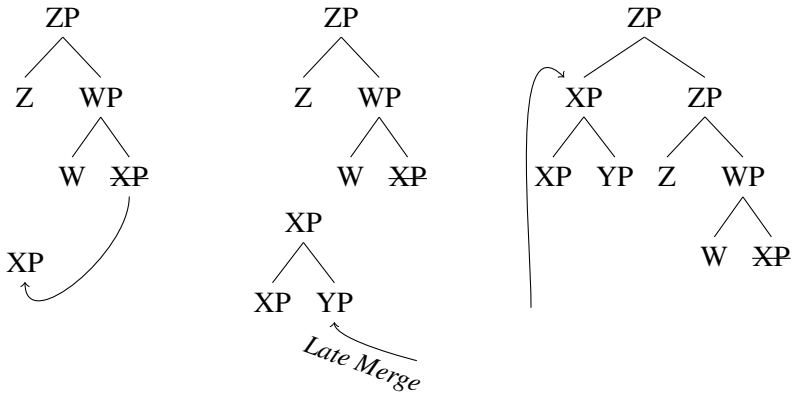


(34) Late Merge of YP



Nevertheless, syntax restricted by the SCC/EC can, in fact, incorporate late Merge if additional assumptions are made. As noted in the previous section, since late Merge is used to derive anti-connectivity effects, it applies inside constituents that undergo movement. This opens up the possibility to circumvent the SCC/EC by assuming that n the course of movement phrases are first copied to the workspace. In the workspace, they can be merged with further syntactic objects without violating strict cyclicity. Such approach to late Merge was pursued in Nunes (2004) as well as by Heck (2016, 2023). It is schematized below in (35)-(37) in the most general form. In this derivation, XP is the syntactic object that moves, but instead of merging with the root node directly, it is first copied to the workspace as shown in (35). After this, XP ceases to be in a proper subdomain of the main tree structure and becomes a root of another tree. Thus, it can be merged with another syntactic object YP without violating the SCC/EC; see (36). Finally, XP is merged back into the main structure.

(35) Step 1 (36) Step 2 (37) Step 3



Since the SCC/EC do not force the quickest possible discharge of active features, the delayed checking of the active merge feature responsible for Merge of YP in the derivation above is, by itself, in line with these conditions and can be ensured by principles like Procrastinate (see Chomsky (1993, 1995)). Alternatively, ordering of features can be used to avoid early checking of Merge probes. In that case, a Merge feature must be ordered after an Agree feature that, in turn, finds its goal only later in the derivation but, notably, before movement of a constituent targeted by late Merge. I will pursue this second option here.

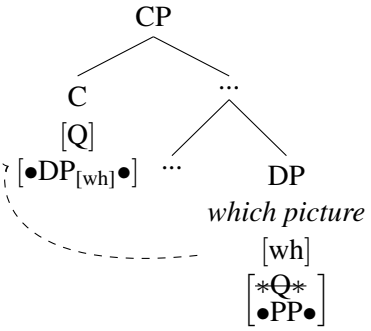
Let us see how this applies to the core cases of anti-connectivity discussed above. Again, I will start with the adjuncts of Ā-moved wh-phrases in (38) repeated from (2b).

(38) [Which pictures near John$_i$]$_j$ did he$_i$ look at $\_\__j$?

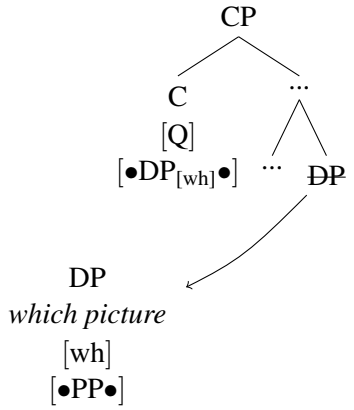
Similarly to the previous section, I assume that wh-dependency involves [•DP$_{[wh]}$•] on the C head and [∗Q∗] on the wh-phrase, but in this derivation the agreement probe must be discharged via upward agreement before movement as shown in (39). After deletion of the agreement probe, the Merge feature occurs on the top of the stack, but it cannot trigger Merge right away because the operation would require Merge to the non-root domain and is therefore excluded by the SCC/EC. In the next step of the derivation shown in (40),

[•$DP_{[wh]}$•] on the C head attracts the wh-phrase, but instead of merging directly to Spec,CP, the wh-phrase is first merged to the workspace (see (40)).

(39) Step 1: Agree

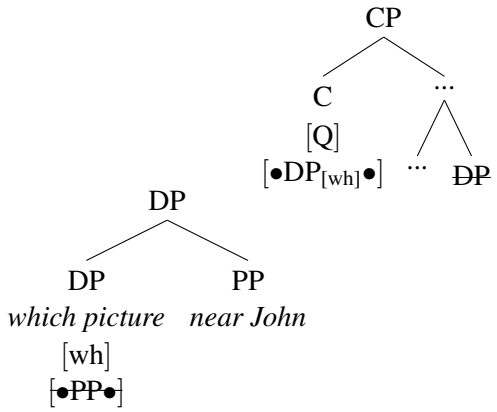


(40) Step 2: Merge



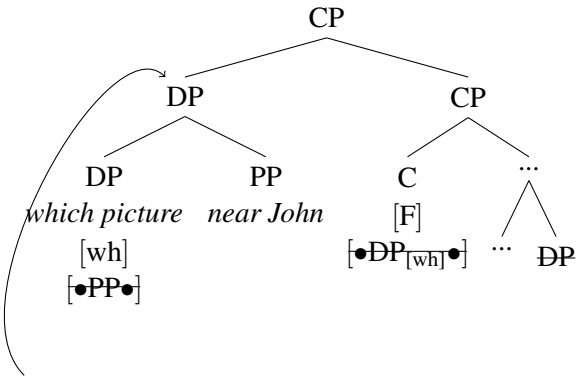
After this, the active Merge feature can be discharged without violating the SCC/EC (see (41)).

(41) Step 3: Late Merge



In the final step in (42), the wh-phrase with the late merged adjunct is merged in Spec,CP.

(42) Step 4: Merge into Spec,CP

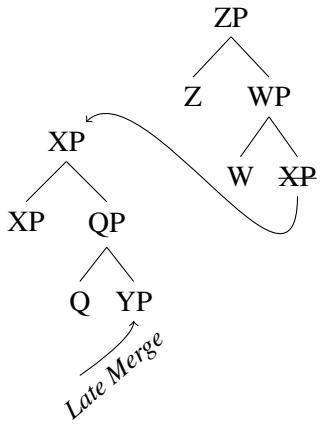


Anti-connectivity for arguments of A-moved phrases, as in (43), is derived in the same vein with the only differences being that the whole NP is late merged, and it is a case probe that can ensure a delayed discharge of the merge feature.

(43) [These pictures of John$_i$]$_j$ seemed to him$_i$ [__$_j$ to be very good].

Interestingly, this implementation of late Merge imposes restrictions on the distribution of late Merge analogous to those discussed in the previous section for the model with the EP/FC plus downward Agree. In particular, late Merge can apply to a moved syntactic object itself, but not to a node deeper embedded into the displaced constituent, as in (44). In this case, late Merge would need to apply to a node properly contained in the phrase copied to the workspace, and this violates the SCC/EC.

(44) Impossible late Merge



As a result, if earlier consensus that late Merge cannot take place in a position above the case assignment position and cannot apply to complements is to be preserved, late Merge compatible with the SCC/EC cannot account for the data in (45), where the adjunct that shows no connectivity to the base position is embedded in several complements of an Ā-moved phrase.

(45) [Whose criticism of [Mary's rendition of (...) the claim [that you [formulated (...) the hypothesis [that Henri [visited the villages near Picasso$_i$'s estate]]]]]]$_k$ did he$_i$ endorse __$_k$?

To sum up, in this section, I have discussed the implementation of late Merge that is compatible with the SCC/EC (see Nunes (2004) and Heck (2023)). It requires movement through the workspace, where a displaced phrase is not in the subdomain of the main structure, but a root of its own tree.

4. Summary

Late Merge is often used for deriving anti-connectivity effects, but is at the same time extensively criticized as being inherently counter-cyclic. This paper shows that it can, in fact, be implemented in a way fully compatible with all major concepts of cyclicity. In each case, however, additional assumptions are necessary. To incorporate late Merge into models regulated by the EP/FC, it is required to assume that features on syntactic nodes are organized in ordered

stacks. Models restricted by the SCC/EC require to assume that movement involves copying of a syntactic object to the workspace. At the same time, despite the general possibility to implement late Merge, none of the considered models by themselves predict late Merge in exactly those contexts where it is needed to derive anti-connectivity. Coupling with widely assumed restrictions on late Merge related to case and adjunct/complement status also does not automatically yield a correct distribution.

References

Adger, David (2003): *Core Syntax: A Minimalist Approach*. Oxford University Press, Oxford.

Adger, David, Alex Drummond, David Hall and Coppe van Urk (2017): 'Is there Condition C Reconstruction?', *North East Linguistics Society* **47**, 131–145.

Bhatt, Rajesh and Roumyana Pancheva (2004): 'Late Merger of Degree Clauses', *Linguistic Inquiry* **35**(1), 1–45.

Bjorkman, Bronwyn M. and Hedde Zeijlstra (2019): 'Checking Up on (ϕ-)Agree', *Linguistic Inquiry* **50**(3), 527–569.

Bruening, Benjamin and Eman Al Khalaf (2019): 'No argument–adjunct asymmetry in reconstruction for Binding Condition C', *Journal of Linguistics* **55**(2), 247–276.

Chomsky, Noam (1973): Conditions on Transformations. *In:* S. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1993): A minimalist program for linguistic theory. *In:* K. Hale and S. J. Keyser, eds, *The view from Building 20: Essays on linguistics in honor of Sylvain Bromberger*. MIT Press, Cambridge, MA, p. 1–52.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge.

Chomsky, Noam (2019): Lectures. Paper presented at University of California, Los Angeles, April 29 – May 2.

Costa, João (2000): Word order and discourse-configurationality in European Portuguese. *In:* J. Costa, ed., *Portuguese syntax: New comparative studies*. Oxford University Press, Oxford, p. 94–115.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima and Hisatsugu Kitahara (1998): *A derivational approach to syntactic relations*. Oxford University Press, New York, Oxford.

Fox, Danny (1999): 'Reconstruction, Binding Theory, and the Interpretation of Chains', *Linguistic Inquiry* **30**(2), 157–196.

Gong, Zhiyu Mia (2022): Issues in the Syntax of Movement: Cross-Clausal Dependencies, Reconstruction, and Movement Typology. Doctoral thesis, Cornell University, Ithaca, NY.

Heck, Fabian (2016): Non-Monotonic Derivations. Habilitation, Universität Leipzig, Leipzig.

Heck, Fabian (2023): Towards a Unified Explanation of Apparent Cases of Countercyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeitsberichte*, Institut für Linguistik, Universität Leipzig.

Heck, Fabian and Gereon Müller (2007): Extremely local optimization. *In:* E. Brainbridge and B. Agbayani, eds, *WECOL 34: Proceedings of the 34th Western Conference on Linguistics*. Fresno, California State University, p. 170–183.

Hulsey, Sarah and Uli Sauerland (2006): 'Sorting out Relative Clauses', *Natural Language Semantics* **14**(2), 111–137.

Keine, Stefan and Rajesh Bhatt (2019): Secondary strong crossover in Hindi and the typology of movement. *In:* M. Baird and J. Pesetsky, eds, *NELS 49: Proceedings of the Forty-Ninth Annual Meeting of the North East Linguistic Society*. GLSA, University of Massachusetts, Amherst, MA, pp. 125–134.

Lebeaux, David (1988): Language acquisition and the form of grammar. PhD thesis, University of Massachussetts at Amherst.

Lebeaux, David (1990): 'Relative Clauses, Licensing, and the Nature of the Derivation', *North East Linguistics Society* **20**(2), 318–332.

Müller, Gereon (2011): *Constraints on Displacement. A Phase-Based Approach*. Benjamins, Amsterdam.

Müller, Gereon (2014): *Syntactic Buffers*. Vol. 91 of *Linguistische Arbeitsberichte*, Universität Leipzig, Amsterdam.

Nunes, Jairo (2004): *Linearization of Chains and Sideward Movement*. Cambridge, MA.

Pesetsky, David (1989): Language-Particular Processes and the Earliness Principle. Ms., MIT, Cambridge, MA.

Pesetsky, David and Esther Torrego (2001): T-to-C movement: causes and consequences. *In:* M. Kenstowicz, ed., *Ken Hale: a life in language*. MIT Press, Cambridge, MA, pp. 355–426.

Richards, Norvin (1999): Featural Cyclicity and the Ordering of Multiple Specifiers. In: *Working Minimalism*. The MIT Press, Cambridge, MA.

Richards, Norvin (2001): *Movement in Language*. Oxford University Press, Oxford.

van Riemsdijk, Henk and Edwin Williams (1981): 'NP-Structure', **1**(2), 171–218.

Safir, Ken (2019): 'The A/Ā Distinction as an Epiphenomenon', *Linguistic Inquiry* **50**(2), 285–336.

Sportiche, Dominique (2019): 'Somber Prospects for Late Merger', *Linguistic Inquiry* **50**(2), 416–424.

Stabler, Edward P. (1997): Derivational minimalism. *In:* C. Retoré, ed., *Logical aspects of computational linguistics*. Springer, Berlin, p. 68–95.

Stockwell, Richard, Aya Meltzer-Asscher and Dominique Sportiche (2021): There is reconstruction for Condition C in English questions. *In:* A. Farinella and A. Hill, eds, *NELS 51: Proceedings of the Fifty-First Annual Meeting of the North East Linguistic Society*. Vol. 2, GLSA, University of Massachusetts, Amherst, MA, pp. 205–214.

Stockwell, Richard, Aya Meltzer-Asscher and Dominique Sportiche (2022): Experimental evidence for the Condition C argument-adjunct asymmetry in English questions. Ms., Christ Church, Oxford; Tel Aviv University; University of California, Los Angeles. To appear in NELS 52: Proceedings of the Fifty-Two Annual Meeting of the North East Linguistic Society.

Takahashi, Shoichi and Sarah Hulsey (2009): 'Wholesale Late Merger: Beyond the A/Ā distinction', *Linguistic Inquiry* **40**(3), 387–426.

van Urk, Coppe (2015): A uniform syntax for phrasal movement: A case study of Dinka Bor. Doctoral thesis, Massachusetts Institute of Technology, Cambridge, MA.

Wierzba, Marta, Martin Salzmann and Doreen Georgi (2020): An experimental investigation of reconstruction for Condition C in German Ā-movement. Ms., Universität Potsdam & University of Pennsylvania. To appear in Proceedings of the Annual Meeting of the Chicago Linguistic Society (CLS) 56.

Wurmbrand, Susi (2012): The syntax of valuation in auxiliary–participle constructions. *In:* J. Choi, E. A. Hogue, J. Punske, D. Tat, J. Schertz and A. Trueman, eds, *Proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL 29)*. Cascadilla Press, Somerville, p. 154–162.

Zeijlstra, Hedde (2012): 'There is only one way to agree', *The Linguistic Review* **29**(3), 491–539.