Challenges for Cyclicity

Gereon Müller*

Abstract

This paper lays out and discusses classical concepts of cyclicity from the point of view of modern grammatical theory (in particular, the minimalist program), focussing on the Cyclic Principle and the Strict Cycle Condition in syntax and morphology. Against this background, the paper addresses two challenges for cyclicity.

First, there is a significant *empirical overlap* between the two constraints, so the question arises whether both are needed. I show that the Cyclic Principle and the Strict Cycle Condition are neither conceptually similar nor reducible to one another as far as their empirical effects are concerned; so they are both required as core constraints in a derivational approach to grammar.

The second challenge is posed by a certain class of phenomena that at first sight seem to call these constraints ensuring cyclicity into question, and that can be grouped under the rubric *counter-cyclic repair*; I argue that these challenges for cyclicity can be overcome by cyclic derivational branching.

1. Concepts of Cyclicity

1.1. Background

The assumption that operations apply cyclically in a derivational approach to grammar was first made in work like Chomsky (1965) (for syntax) and Chomsky and Halle (1968) (for phonology), and it has subsequently been developed and modified in a number of different ways that are often mutually

*For comments and discussion, thanks are due to the participants of the *Cyclic Optimization* Research Unit retreat in Großbothen on May 11, 2022 and the participants of the Research Unit Workshop on (Strict) Cyclicity at Universität Leipzig on June 15, 2022. I am particularly grateful to Doreen Georgi, Daniel Gleim, Fabian Heck, Greg Kobele, Masha Privizentseva, Ezer Rasin, Philipp Weisser, and Jochen Trommer for clarifying matters for me (years ago, in some cases). The research for the present paper was supported by DFG grant MU 1444/15-1 (*Prospects of Inflectional Morphology in Harmonic Serialism*) as part of Research Unit FOR 5175.

Cyclicity, 3–72
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

incompatible. The most general abstract concept of cyclicity that is at the heart of all more specific implementations is arguably (1).[1]

(1) *Cyclicity*:
   a. First, the derivation carries out (a potentially singleton set/list of) operations of type T_1.
   b. Second, the derivation carries out (a potentially singleton set/list of) operations of type T_2.
   c. Then the derivation carries out (a potentially singleton set/list of) operations of type T_1 again.
   d. And so on.

There are various possible instantiations of what T_1, T_2 in (1) stand for, leading to different characterizations of cyclicity with different empirical consequences. Thus, according to the concept of cyclicity proposed in Kiparsky (1982*a*,*b*), which we may refer to as *cyclicity*$_k$, T_1 involves *structure-building* whereas T_2 involves *other operations*. One way of implementing this is pursued in the model of lexical morphology and phonology developed by Kiparsky himself. On this view, T_1 is the system of lexical morphology, and T_2 is the system of lexical phonology. Another option on the basis of cyclicity$_k$, given the premises of minimalist syntax (cf. Chomsky (2001, 2008, 2013)) might be to assume that T_1 involves applications of Merge in a certain local domain (e.g., the *phase*; cf. Chomsky (2001)), and T_2 consists of applications of other operations (Agree, Spell-Out, Delete, ...). This would then lead to a scenario where all Merge operations necessarily precede all other operations in any given phase. Such a strict order of operations is certainly not a priori doomed to fail (and may in fact be tacitly presupposed in a lot of minimalist work), but assuming it to be always present is not uncontroversial either (see, e.g., Assmann et al. (2015), Georgi (2017), Murphy and Puškar (2018) and Fritzsche (2020) for evidence (based on (counter-) feeding and (counter-) bleeding relations) that Agree can precede Merge in phase).

Another view of what T_1 and T_2 stand for in (1) gives rise to a different concept of cyclicity, which we might dub *cyclicity*$_c$ because it is the one proposed by Chomsky (1965, 1975) for syntax (and by Chomsky and Halle

[1] "T" is supposed to be reminiscent of transformations (or sets of transformations), since this was the formal device triggering the grammatical operations in question at the time when cyclicity was discovered; also see below.

(1968) for phonology).[2] Here, T_1, T_2 involve operations taking place in a given *cyclic domain* (i.e., a *cyclic node*), where T_1 precedes T_2. After the cyclic domain is finished, the derivation moves to the next cyclic domain (typically bottom-up), and again applies T_1 before T_2; etc. This kind of interaction of operations is known as the classical *transformational cycle*.

In the standard conception (as it is laid out in Chomsky (1965)), the transformational cycle is an ordered list of transformations that the derivation goes through within a cyclic node; and after finishing the list, the derivation moves up to the next (structurally higher) cyclic node; etc. The order among transformations is justified by sequential interactions: feeding (T_1 creates the context in which subsequent T_2 can apply), counter-feeding (T_1 would create the context in which T_2 can apply but applies too late to actually do so), bleeding (T_1 destroys the context in which subsequent T_2 can apply), and counter-bleeding (T_1 would destroy the context in which T_2 can apply but applies too late to actually do so).[3] Originally, the relevant cyclic nodes were the clause (more specifically, S, or, in current notation, TP) and the nominal domain (i.e., NP or DP, depending on which of the two categories is viewed as the top-most maximal projection of nominal categories). In addition, AP has also sometimes been argued to be a cyclic node (see, e.g., Chomsky (1975)).

1.2. Cyclic Nodes

Interestingly, these cyclic nodes could subsequently also be argued to play an important role in grammatical building blocks that are not directly related to cyclicity. For one thing, cyclic nodes are the relevant *bounding nodes* in the original definition of a general constraint on movement, viz., the *Subjacency Condition* (cf. Chomsky (1973, 1977)). For another, these nodes are also relevant for the definition of *cyclic command*, also known as *kommand* (a

[2]The concept of *cyclicity*$_c$ historically precedes the concept of *cyclicity*$_k$; Trommer (2020) accordingly refers to it as *proto-cyclicity*. Also cf. Kobele (2023) on this distinction.

[3]Incidentally, a major part of any analysis employing the transformational cycle was to determine the exact order of the transformations, and long lists were usually presented to this effect in this kind of approach. See, e.g., Huber and Kummer (1974, 351) for an early proposal for German. (The longest list of transformations for a single language that I am aware of is proposed in Ross (2012) for English. However, the list is thematically ordered; Ross makes no attempt to suggest a (full or partial) ordering of the more than 200 transformations in this paper that would determine the sequence of application, and that would thus be intrinsically motivated by (counter-) feeding and (counter-) bleeding relations.)

predecessor of c-command that it got replaced by in Reinhart (1976, 1983)). Furthermore, these nodes exhibit (morphological, phonological, semantic, or syntactic) *reflexes of successive-cyclic movement*. Let me briefly address these three issues in turn.

The definition of the Subjacency Condition proposed in Chomsky (1973, 1977) is given in (2).

(2) *Subjacency Condition*:
In a structure α ... [$_\beta$ ... [$_\gamma$ ... δ ...] ...] ..., movement of δ to α cannot apply if β and γ are bounding nodes.

As noted, it is exactly the *cyclic nodes* DP (NP) and TP (S) that Chomsky assumes to qualify as *bounding nodes* here. He remarks that he "will tentatively suppose that condition [(2)] is a general property defining cyclic application of transformations" (see Chomsky (1973, 248)). In current terminology, this implies that movement from an embedded clause must be successive-cyclic, as in (3a). Movement from an embedded clause cannot skip a potential specifier SpecC because otherwise two cyclic nodes of type TP would be crossed by a single movement step, as is the case with wh-islands (created by wh-movement of another wh-phrase in the interrogative embedded clause) as in (3b) in English, given that it can be ensured that the embedded C can only have one specifier here.

(3) a. [$_{DP_1}$ Which book] do [$_{TP_2}$ you think [$_{CP}$ t'_1 [$_C$ (that)] [$_{TP_4}$ John read t_1]]] ?
b. ?*[$_{DP_1}$ Which book] do [$_{TP_2}$ you wonder [$_{CP}$ [$_{PP_3}$ to whom] C [$_{TP_4}$ John gave t_1 t_3]]] ?

That said, strictly speaking one might expect CP (rather than TP) to act as both a cyclic node and a bounding node since CP (rather than TP) represents the maximal clausal projection, and wh-movement on the embedded cycle clearly targets a position beyond TP (viz., SpecC). Rizzi (1982) has in fact argued that this null hypothesis is indeed correct for a language like Italian (where the analogue of (3b) is grammatical, at least if the long-distance-moved item is a relative pronoun, and only the crossing of *two* wh-CPs – or an extraction from a CP contained in a DP – can trigger a violation of the Subjacency Condition). However, such a choice of bounding nodes would obviously undermine the account of the illformedness of (3b): Movement crosses two TP nodes here,

but only one CP node. This, then, might be taken to indicate that the conflation of cyclic nodes with bounding nodes is not an entirely innocuous one.

Next, the concept of cyclic node has also been argued to play a role for binding theory. Since Chomsky (1981), it is customary to assume that disjoint reference effects as in (4ac) (vs. (4bd)) in English follow from *Principle C*, according to which non-pronominal DPs must not be *c-commanded* by a co-indexed DP.

(4) a. *He$_1$ always wore dark glasses because John$_1$ was famous
b. Because he$_1$ is famous, John$_1$ always wears dark glasses
c. *He$_1$ always depressed John$_1$
d. The portrait of his$_1$ mother always depressed John$_1$

The definition of c-command standardly makes use of the primitive notions of next branching node and dominance (or Merge; cf. Epstein et al. (1998)), and is therefore as such unrelated to cyclicity. However, there is an earlier approach to the pattern in (4) according to which the relation of *cyclic command*, or *kommand*, in (5) is the relevant concept (see Wasow (1972), Lasnik (1976), and Fanselow (1983), among others).

(5) *Kommand*:
A node A kommands a node B iff A and B are not in a dominance relation and the first cyclic node dominating A also dominates B.

On this view, the illformedness of (4a) and (4c) (as well as the wellformedness of (4b) and (4d)) follows from the Disjoint Reference Constraint in (6).

(6) *Disjoint Reference Constraint*:
A pronoun A cannot be coreferent with a DP B if A precedes B and A kommands B.

In (4a) and (4c), the pronoun precedes and kommands the co-indexed proper name: The first cyclic node dominating the pronoun is the clause, which also dominates the proper name. In contrast, in (4b) and (4d), the pronoun does not precede and kommand the proper name: The pronouns do not kommand the co-indexed proper names in (4b) and (4d) because the first cyclic nodes dominating the pronouns are a clause and a DP, respectively, that do not dominate the proper name.

Since Reinhart (1976, 1983), the approach based on the Disjoint Reference

Constraint that works with cyclic nodes has widely been assumed to have been falsified, and to have successfully been replaced by the approach based on Principle C that relies on c-command. One of the core arguments comes from instances of PP fronting in English as in (7ab). These data do not show a disjoint reference effect even though the pronoun precedes the co-indexed proper name and would seem to kommand it, too (the minimal clause or DP dominating the pronoun also dominates the proper name).

(7) a. Near him$_1$, Dan$_1$ saw a snake
b. In her$_1$ bed, Zelda$_1$ spent her sweetest hours

However, as shown by Bruening (2014), closer inspection reveals that a case can be made for an approach along the lines of the Disjoint Reference Constraint after all, assuming that *phases* are assumed as cyclic domains (and an appropriate theory of reconstruction is adopted).

Finally, cyclic nodes (or, in current approaches, phases; see below) exhibit various kinds of reflexes that suggest that they have been targetted by intermediate movement steps in the course of long-distance extraction. These reflexes can be phonological in nature (see, e.g., Clements et al. (1983) and Korsah and Murphy (2020)). They can also be morphological (see McCloskey (1979, 2002), Collins (1993, 1994), Chung (1994, 1998), Cole and Hermon (2000), Fanselow and Ćavar (2001), Schneider-Zioga (2005), Lahne (2009), van Urk (2015), and Georgi (2017), among others). They can be syntactic, in the sense that a syntactic operation is triggered that would otherwise be unexpected (see, e.g., Barss (1986), Epstein et al. (1998), Müller (1999*a*), and Barbiers (2002)), or they can be semantic (cf., e.g., Fox (2000) and Nissenbaum (2000)).

The morphological reflex of Modern Irish complementizer variation (depending on whether cyclic movement to a SpecC position has taken place or not) that is investigated in McCloskey (1979, 2002) is one of best-known instances of reflexes of successive-cyclicity. The regular form of declarative C is *go*; see (8a). However, if SpecC is targetted by movement, C takes the form *aL*; see (8b). Importantly, if movement takes place from an embedded clause, the complementizer also emerges as *aL* rather than as *go*; see (8c). This systematic morphological change can thus be viewed as an instance of movement-related morphology, and strongly suggests that long-distance movement applies cyclically.

(8) a. Creidim gu-r inis sé bréag
I-believe C:*go*-PAST tell he lie
'I believe that he told a lie.'

b. Céacu ceann$_1$ a dhíol tú t_1 ?
which one C:*aL* sold you
'Which one did you sell?'

c. an t-ainm OP$_1$ a hinnseadh dúinn a bhí t_1 ar an áit
the name C:*aL* was told to us C:*aL* was on the place
'the name that we were told was on the place'

2. The Cyclic Principle and the Strict Cycle Condition

2.1. The Cyclic Principle

So far, the discussion of concepts of cyclicity has been somewhat informal, especially as far as the exact nature of cyclicity$_c$ is concerned. The gist of cyclicity$_c$ is that (i) within each cyclic domain (or cyclic node), operations apply in the presence of triggers, based on some extrinsic order (obligatorily or optionally, depending on the nature of the trigger), (ii) cyclic domains are hierarchically ordered from bottom to top, and (iii) the derivation moves to a higher cyclic domain only after it has carried out the operations applicable in a given cyclic domain. Focussing for now on the grammatical components of syntax and morphology, a more precise standard definition of cyclicity$_c$ is given in (9) as the Cyclic Principle (cf. Chomsky (1965) and Perlmutter and Soames (1979), among many others).

(9) *Cyclic Principle* (standard version):
When two operations can be carried out, where one applies to the cyclic domain D_x and the other applies to the cyclic domain D_{x-1} included in D_x, then the latter is applied first.

A crucial observation going back to McCawley (1984, 1998) is that since the Cyclic Principle predicts orders among operations if they take place in different cyclic domains, the size (and, hence, the number) of cyclic domains will have an interesing effect: The smaller the cyclic domains are (i.e., the more cycles there are), the more orders are predicted. This way, the extrinsic order in (i) above can eventually be dispensed with since it emerges as a

subcase of the statement in (iii). To see this, consider first some possible assumptions about what a cyclic domain is.

(10) *Candidates for cyclic domains*:

a. Every classical cyclic node (CP, DP) is a cyclic domain.[4]
b. Every phase (CP, DP, vP) is a cyclic domain.
c. Every phrase is a cyclic domain.
d. Every projection is a cyclic domain.

Suppose next (just for the sake of the argument) that there are eight operations O_1,...O_8 that can take place in a CP domain, such that O_1 targets SpecC, O_2 targets C, O_3 targets SpecT, O_4 targets T, 0_5 targets Specv, O_6 targets v, O_7 targets SpecV, and 0_8 targets V, as schematically illustrated in (11).

(11) $[_{CP_0}$ ... $[_{CP_1}$ XP($\leftarrow O_1$) $[_{C'}$ C($\leftarrow O_2$) $[_{TP}$ YP($\leftarrow O_3$) $[_{T'}$ T($\leftarrow O_4$) $[_{vP}$ WP($\leftarrow O_5$) $[_{v'}$ v($\leftarrow O_6$) $[_{VP}$ ZP($\leftarrow O_7$) $[_{V'}$ V($\leftarrow O_8$) CP_2]]]]]]]]

If cyclic domains in the sense (9) are the classical cyclic nodes CP and DP, as in (10a), no orders are predicted among the operations O_1,...,O_8 (however, it can be ensured that O_1,...,O_8 precede all operations outside of CP_1 in (11), in the higher CP_0 domain, and follow all operations in the embedded CP_2 domain). Next, if phases (CP, DP, vP) are the cyclic domains, as in (10b), the Cyclic Principle guarantees that O_1–O_4 follow O_5–O_6 (plus, the consequences regarding operations in CP_0 and CP_2 are as before). Third, if every phrase is a cyclic domain, as in (10c), the Cyclic Principle has the consequence that O_1, O_2 follow all other operations (and precede operations in CP_0); O_3, O_4 follow O_5–O_8; and O_5, O_6 follow O_7, O_8 (which in turn follow operations in CP_2). Finally, if every projection is a cyclic domain, as in (10d), this imposes a complete order: Operations in CP_2 precede O_8, which precedes O_7, which in turn precedes O_6, and so on, until O_1 takes place; and then the derivation turns to operations on the CP_0 cycle, following the same fine-grained order there. It should be emphasized that this is an important result since it shows that by adopting the Cyclic Principle it becomes possible to reduce a significant amount of orders among operations to a simple independent factor, viz., the question of what the current cyclic domain is.

As a matter of fact, if one adopts the incremental (bottom-up) derivational

[4]However, recall the qualification regarding TP noted above.

approach to syntax based on structure-building via iterated Merge operations that is developed in Chomsky (2001) and much subsequent work, the notion of cyclic domain in (10d) is a very natural one (it is also the one that McCawley (1984) adopts); on this view, each Merge operation creates a new cyclic domain. From this perspective, it also becomes possible to substantially simplify the definition of the Cyclic Principle, which in the standard form in (9) qualifies as a transderivational constraint (cf. Müller and Sternefeld (2001) and Graf (2013), among others) because it necessitates a comparison between two derivations and, for this reason, can only punish a counter-cyclic order of operations by looking at an alternative, cyclic, derivation (i.e., non-application of an operation on the D_{x-1} cycle cannot be precluded locally, at this stage of the derivation). The simpler version of the Cyclic Principle that becomes available under these assumptions is given in (12).

(12) *Cyclic Principle* (simpler version):
An operation must apply as soon as its trigger is present.

This formulation makes it clear that two other principles that have sometimes been adopted in derivational approaches to grammar are simply rephrased versions of the Cyclic Principle. First, the *Earliness Principle* proposed in slightly different versions in Pesetsky (1989) and Pesetsky and Torrego (2001) essentially amounts to (12).[5] Second, *Featural Cyclicity* (see Richards (1999,

[5]The original version of the Earliness Principle in Pesetsky (1989) postulates as cyclic domains complete syntactic structures at different levels of representation, and thus does not by itself yet predict any order among, say, operations producing surface structure representations; cf. (ia). However, the version of the Earliness Principle in (ib), which goes back to Pesetsky and Torrego (2001), is already much closer to (12) (assuming that uninterpretable features serve as triggers for operations, and marking such a feature for deletion is the result of applying the operation in question).

(i) a. *Earliness Principle* (Pesetsky (1989)):
Satisfy filters as early as possible on the hierarchy of levels: (DS >) SS > LF > LP (level of language-particular rules).
b. *Earliness Principle* (Pesetsky and Torrego (2001)):
An uninterpretable feature must be marked for deletion as early in the derivation as possible.

Finally, the interpretation of the Earliness Principle in Chomsky (2001, 15) is basically identical to (12), and it seems fair to state that this is how the Earliness Principle is normally understood nowadays.

2001), Preminger (2018), and Börjesson and Müller (2020), among others) also emerges as a version of the Cyclic Principle. According to this constraint, active features that can trigger operations must do so as soon as possible; i.e., they cannot wait and become embedded by further structure-building.

2.2. The Strict Cycle Condition

In addition to (a constraint like) the Cyclic Principle, a second constraint demanding cyclicity has often been adopted, so as to ensure that after one cycle is completed and the derivation has moved to the next cycle, the representation attained in the first cycle can only be modified very selectively. This constraint is known as the *Strict Cycle Condition*. A standard definition is given in (13).[6]

(13) *Strict Cycle Condition* (SCC):
Once a cyclic domain D_x has been affected by an operation, no subsequent operation may exclusively affect a cyclic domain D_{x-1} that is a proper subdomain of D_x.

The Strict Cycle Condition has first been proposed for syntactic dependencies in Chomsky (1973, 243).[7] Subsequently, it has been actively employed in many syntactic analyses (see, e.g., Chomsky (1995, 2001, 2004, 2008, 2019), Perlmutter and Soames (1979), Barss (1986), Freidin (1992, 1999), Pullum (1992), Collins (1997), Kitahara (1997), Roberts (1997, 2021), Bošković and Lasnik (1999), Müller (2011), Abels (2012), Collins and Stabler (2016), and Gallego (2020)); and it has been tacitly presupposed in many more.[8] According to the Strict Cycle Condition, it is prohibited to exclusively modify an embedded part of a cyclic structure generated earlier. The degree of strictness of the constraint depends on how "cyclic domains" are understood. According to the most restrictive concept, every projection is a cyclic node; if this is the case, the domains for the Cyclic Principle (based on (10d)/(12)) and the Strict Cycle Condition are identical. On this view, every legitimate

[6] A more radical option that has also been pursued is to assume that after the derivation has moved to a new cycle, the structure generated so far cannot be modified *at all*. This follows if a constraint like *Bracket Erasure* or *Multiple Spell-Out* is adopted; more on these below.

[7] In (13), several minor adjustments to the original definition of Chomsky (1973) have been carried out in order to ensure maximal compatibility with the Cyclic Principle in (9).

[8] The text following on pp. 12–14 partly contains material that is an extended version of the corresponding text in Müller (2021).

operation must involve the current – i.e., top-most – cyclic domain – the current root domain, in standard minimalist approaches employing incremental structure-building Merge. Under these assumptions (and focussing only on Merge operations for now), the Strict Cycle Condition in (13) can in principle be reformulated as in (14), in roughly the same way that the Cyclic Principle in (9) could be simplified as in (12) (cf. Chomsky (1995)).

(14) *Extension Condition*:
Every structure-building operation must extend the current root.

In what follows, if nothing else is said, I will presuppose these most restrictive versions of the Cyclic Principle and the Strict Cycle Condition.

2.3. Convergence

2.3.1. Freezing Effects

Among many other things, the Strict Cycle Condition has been argued to be indispensable in a derivational account of freezing effects (Wexler and Culicover (1980), Browning (1991), Lohndal (2010)), as in (15a) (with DP raising to subject in the passive interacting with wh-movement from DP in English) and (15b) (with VP topicalization interacting with wh-movement from VP in German).

(15) a. *Who$_1$ was [$_{DP_2}$ a picture of t_1] painted t_2 by Mary ?
b. *Was$_1$ denkst du [$_{CP}$ [$_{VP_2}$ t_1 gelesen] hat keiner t_2] ?
what$_{acc}$ think you read has no-one$_{acc}$

Given that extraction from XP is possible only if XP is a complement (cf. the *Condition on Extraction Domain*; see Huang (1982), Chomsky (1986), Cinque (1990)), the illformedness of (15ab) can be derived if movement of DP_2 and VP_2 precedes extraction of DP_1 (because DP_2/VP_2 occupies a specifier when DP_1 extraction takes place; cf. Browning (1991)); but the reverse, counter-cyclic sequence of movement operations where DP_1 extraction takes place when DP_2/VP_2 is still in situ, in a complement position, must also be excluded, and this is accomplished by the Strict Cycle Condition (cf. Collins (1994)): When DP_1 is moved to its target, criterial, SpecC position in (15ab), this defines CP as the current cyclic domain. Consequently, subsequent movement of XP_2 to a lower position (SpecT and embedded SpecC, respectively) affects

solely a proper subdomain of the root CP that is the current cyclic domain, and the Strict Cycle Condition is violated. Similarly, if it is assumed that the embedded wh-phrase in SpecC is responsible for the wh-island effect in a sentence like (3b), repeated here as (16), the Strict Cycle Condition is needed to ensure that it is already present when movement of the other wh-phrase to the matrix domain takes place.[9]

(16) ?*[$_{DP_1}$ Which book] do [$_{TP_2}$ you wonder [$_{CP}$ [$_{PP_3}$ to whom] C [$_{TP_4}$ John gave t_1 t_3]]] ?

But wait. Closer inspection suggests that the Cyclic Principle could already suffice to account for freezing effects. By assumption, there is a local trigger for XP_2 movement on T in (15a), and on embedded C in (15b); such a trigger can, e.g., take the form of designated structure-building features like [•F•] – more specifically, [•D•] for EPP-driven movement to SpecT in (15a), and [•wh•] for movement to SpecC in questions in (15b).[10] Such movement-inducing features are present on a head (T and embedded C, in the cases at hand) when the head enters the structure. Therefore, the Cyclic Principle demands that these features immediately give rise to XP_2 movement in the derivations under consideration, i.e., before the maximal projection of the head is embedded under something else. Postponing a satisfaction of the demands of these [•F•] features until a higher $C_{[•wh•]}$ head has triggered DP_1 movement (while XP_2 is still in situ, in a complement position, so that the CED can be respected) is thus not an option under the Cyclic Principle.[11]

[9]Unless, that is, one assumes that there is a trace, or copy, of the long-distance moved wh-phase in the SpecC position. A stipulation demanding a restriction to a single specifier position for C will then block subsequent, counter-cyclic movement of PP_3 (*to whom*) in (3b)/(16) without recourse to the Strict Cycle Condition, and this may then in turn lead to a violation of some visibility requirement for interrogative C.

[10]Cf. Heck and Müller (2007, 2013) for the specific notation; and Pesetsky and Torrego (2006), Abels (2012), Georgi (2014, 2017), Müller (2014, 2022), Stabler (2013), Assmann et al. (2015), van Urk (2015), Collins and Stabler (2016), Zyman (2018), Longenbaugh (2019), Newman (2021), and Šereikaitė (2021), among others.

[11]Note that this reasoning does not necessitate the assumption that the features [•D•] (on T in (15a)) and [•wh•] (on embedded C in (15b)) are "present" (in the sense of (12)) at an early stage in the generation of the respective TP and CP structures. Even if these features are accessible very late on the respective cycles (e.g., because features are ordered on lists and the features triggering these movements are lowest-ranked), the Cyclic Principle ensures that they will have to be discharged by movement at a stage of the derivation (viz., within TP/CP) that precedes the stage where [•wh•] on (matrix) C becomes active for DP_1.

2.3.2. Late Merge

As a second example illustrating that the Cyclic Principle and the Strict Cycle Condition can yield identical effects, consider the concept of Late Merge. (17b) exhibits a *reconstruction* effect, which is fully expected if Principle C of the binding theory must be respected at every stage of the derivation: Before wh-movement of DP_2, the subject DP *he*$_1$ c-commands *John*$_1$, which is part of a CP that is contained in an object DP; thus, Principle C is fatally violated at an early stage, and a reconstruction effect obtains (in the sense that for the purposes of Principle C, *John*$_1$ behaves as if it were c-commanded by *he*$_1$, which it is not anymore on the surface). In contrast, (17a), which differs minimally from (17b) in that CP is a relative clause rather than a complement clause, does not seem to violate Principle C. Thus, the wellformedness of (17a) indicates *anti-reconstruction*: *John*$_1$ is not c-commanded by *he*$_1$ on the surface (as before), and this time it does not behave as if it were either.

(17) a. [$_{DP_2}$ Which claim [$_{CP}$ that John$_1$ made]] was he$_1$ willing to discuss t$_2$?
b. *[$_{DP_2}$ Which claim [$_{CP}$ that John$_1$ was asleep]] was he$_1$ willing to discuss t$_2$?

A standard analysis of the anti-reconstruction effect in (17a) relies on the observation that CP in (17a) is an adjunct, whereas CP in (17b) is an argument. The central assumption then is that adjuncts can be merged *late*, i.e., after wh-movement applying to DP_2. So, on this view, there is in fact no stage of the derivation where *he*$_1$ would illegitimately c-command the co-indexed proper name *John*$_1$ in (17a); the relevant parts of the derivation are shown in (18). An anti-reconstruction effect arises here because *John*$_1$ only enters the structure as part of the late-merged CP when the wh-phrase has already left the c-command domain of *he*$_1$. Disjoint reference is thus counter-fed by late Merge of the adjunct.

(18) a. *Pre-movement structure*:
[$_{TP}$ he$_1$ was willing to discuss [$_{DP_2}$ which claim]
b. *Wh-movement*:
[$_{DP_2}$ Which claim] was he$_1$ willing to discuss t$_2$?
c. *Late Merge of adjunct CP*:

> [$_{DP_2}$ Which claim [$_{CP}$ that John$_1$ made]] was he$_1$ willing to discuss t$_2$?

This analysis has been adopted for data like those in (17) by Lebeaux (1988), Speas (1990), Freidin (1994), Chomsky (1995), Epstein et al. (1998), and Fox (2000), among others.

In addition, the approach has been extended to other constructions. Thus, based on this earlier work, Takahashi and Hulsey (2009) develop the concept of Wholesale Late Merge, which has subsequently been employed by van Urk (2015), Bhatt and Keine (2019), and Gong (2022). The basic idea here is that not only is it the case that adjuncts can be merged late; late Merge is in fact assumed to be a general option for all items that do not have *case* (yet). For concreteness, according to Wholesale Late Merge, the NP restriction of a D quantifier does not have to be merged in the base position (only D is merged here); it suffices if it is merged late before the DP has been assigned case, after movement to a case position. Again, an anti-reconstruction effect can be predicted. This approach to anti-reconstruction for Principle C with case-driven A-movement in English is illustrated by the example in (19).[12]

(19) [$_{DP_2}$ Every [$_{NP}$ argument [$_{CP}$ that John$_1$ is a genius]] seems to him$_1$ t$'_2$ to be t$_2$ flawless

Late Merge and Wholesale Late Merge both violate the Strict Cycle Condition. For instance, in (17a), wh-movement of the DP$_2$ to SpecC has made it clear that the root CP is the current cyclic domain when the relative clause CP *that John$_1$ made* is merged late with the head noun (or the NP it projects); but this late Merge operation exclusively affects the more deeply embedded cyclic NP domain (which is a proper subdomain of both the DP$_2$ and root CP domains). Thus, the Strict Cycle Condition classifies the derivation in (17a) as counter-cyclic; and the same goes for the derivation in (19), for analogous reasons.

Given this state of affairs, the question arises of whether the Cyclic Principle

[12]Takahashi and Hulsey (2009) refer to the relevant interaction of operations as *bleeding*, but this does not seem entirely correct since Merge of the NP in (19) does not technically bleed disjoint reference (i.e., a Principle C effect) because there would of course be no such effect if Merge of the NP did not apply. Rather, Merge of the NP in (19) comes too late to feed such an effect (i.e., give rise to a Principle C violation or, in other words, to a disjoint reference effect); i.e., the interaction of operations at hand is one of *counter-feeding*.

also classifies (Wholesale) Late Merge as counter-cyclic. It turns out that under standard assumptions, this is the case. The central observation is that Late Merge cannot possibly be taken to imply that an adjunct is obligatorily merged at the latest possible stage of the derivation. Rather, an adjunct (or case-less NP, under Wholesale Late Merge) can be merged at any stage of the derivation, including at a late stage, as in (17a) (or (19)). Thus, postulating Late Merge for some item α really means that Merge of α applies optionally, not obligatorily at a certain designated point. This assumption is empirically required in view of examples like (20)

(20) $[_{DP_3}$ Which paper $[_{CP}$ that he$_1$ gave to Mary$_2$]] did every student$_1$ like t_3 ?

In (20), the pronoun *he*$_1$ is co-indexed with the quantified DP *every student*$_1$. Hence, *he*$_1$ must be interpreted as a bound variable. Bound variable pronouns require an A-binder in the syntax (otherwise a weak crossover effect will occur); this requirement can be formulated as in (21) (cf. Reinhart (1983), Heim (1989), Mahajan (1990), and Heim and Kratzer (1998), among others).

(21) *Condition on Bound Variable Pronouns*:
A bound variable pronoun must be A-bound.

Consequently, the adjunct (i.e., relative clause) CP in (20) must be merged in DP_3 *before* wh-movement of DP_3 to SpecC takes place, so as to ensure that *every student*$_1$ can bind *he*$_1$. Furthermore, requirements for c-command ('reconstruction') and for a lack of c-command ('anti-reconstruction') can hold in a single sentence, and lead to intricate predictions as to the exact place of Merge applying to items for which Late Merge is assumed to be an option. Consider, e.g., the example in (22) (cf. Epstein et al. (1998) for extensive discussion).

(22) $[_{DP_3}$ Which paper $[_{CP}$ that he$_1$ gave to Mary$_2$]] did every student$_1$ think t'_3 that she$_2$ would like t_3 ?

Here, the availability of a bound variable interpretation of *he*$_1$ necessitates the assumption that the adjunct CP is merged in a position c-commanded by the matrix subject *every student*$_1$; and the absence of a Principle C effect with *Mary*$_2$ that would be induced by *she*$_2$ can be taken as evidence that the adjunct CP is merged in a position where DP_3 is not c-commanded by the

embedded subject. As shown in (23), against the background of the Late Merge approach, this suggests that the relative clause CP in (22) is merged when DP_3 has undergone an intermediate movement step to the embedded SpecC position.

(23) a. *Pre-movement structure*:
she$_2$ would like [$_{DP_3}$ which paper]
b. *Intermediate wh-movement*:
[$_{DP_3}$ which paper] that she$_2$ would like t$_3$
c. *Late Merge in intermediate position*:
[$_{DP_3}$ which paper [$_{CP}$ that he$_1$ gave to Mary$_2$]] that she$_2$ would like t$_3$
d. *Binding by matrix subject*:
every student$_1$ did think [$_{DP_3}$ which paper [$_{CP}$ that he$_1$ gave to Mary$_2$]] that she$_2$ would like t$_3$
e. *Criterial wh-movement*:
[$_{DP_3}$ which paper [$_{CP}$ that he$_1$ gave to Mary$_2$]] did every student$_1$ think t$'_3$ that she$_2$ would like t$_3$

Thus, given that Late Merge is to be understood as optional Merge at any point of the derivation (and not as an instruction to merge some item at the final stage of a derivation), it is clear that the Cyclic Principle is directly relevant, and that it is at variance with the Late Merge derivations underlying (17a), (19), and (22): According to the formulation of the Cyclic Principle in (9), if Merge of the adjunct clause can apply at an early stage, it has to apply at that stage, even if this means that a fatal violation of some constraint (Principle C, in the cases at hand) will then subsequently occur.[13]

[13]As observed by Privizentseva (2023), there is a caveat to this conclusion, though: To see this, suppose that features that trigger syntactic operations (i.e., probe features and structure-building features) are all located on a single stack, with only the top-ranked feature accessible at any given stage of the derivation. Suppose, furthermore, that probe features can be satisfied by upward Agree with a c-commanding goal (cf. footnote 26 below). Then, in a scenario where an upward agreeing probe feature dominates a structure-building feature for adjunct Merge, the latter operation may have to wait for a longer period in the derivation before it can become active and trigger Merge – it depends on the top-ranked probe feature to find a goal (and, perhaps, undergo movement in addition). Hence, Merge can be significantly delayed. (Also cf. Fritzsche (2023) for an analogous solution to the problem raised by instances of "late" Agree for the Cyclic Principle.) This kind of approach to late Merge may raise questions with respect to optionality (as discussed in the main text), which may or may not be resolved by postulating

2.4. The Cyclic Principle vs. the Strict Cycle Condition

2.4.1. The Issue

At this point, obvious questions arise: In what sense do the Cyclic Principle and the Strict Cycle Condition differ? Is there a lot of redundancy? Are the constraints both needed, or can one of them be dispensed with in favour of the other?

The first thing to note in this context is that the Cyclic Principle and the Strict Cycle Condition are conceptually very different kinds of constraints. Both presuppose a derivational approach to grammar, where at least some operations are temporally ordered. And to a significant extent, both constraints succeed in excluding "counter-cyclic" derivational steps; accordingly, both can significantly contribute to deriving an order among operations.[14] However, the underlying conceptual justification of the two constraints is not identical: The Cyclic Principle demands that an operation must apply as soon as it can, whereas the Strict Cycle Condition demands maximal stability of linguistic objects created in the course of the derivation. In a nutshell, the two strategies can be identified as "Do it now!" vs. "Leave everything intact!".[15]

These conceptual differences notwithstanding, it has sometimes been

---

different feature lists to be freely available; but it would certainly be in accordance with the Cyclic Principle. Consequently, under these assumptions, Late Merge would in fact qualify as an instance of asymmetries between the two cyclicity constraints that I address in the following section. – That said, under additional assumptions (by postulating non-monotonic derivations), the problem that is raised by late Merge for the Strict Cycle Condition can in principle also be gotten rid of; see Heck (2023) and case study 5 below.

[14]Given an identical conception of the nature of cyclic domains, it can be verified that the Strict Cycle Condition has the same consequences for deriving order as the Cyclic Principle for the operations in (11).

[15]In this context, it is interesting to take into account the discussion of the extra-linguistic relevance of the Strict Cycle Condition in Pullum (1992, 227&230). Pullum basically advances an evolutionary motivation: "Complex structures in language are assembled from well-formed parts which may be modified in the process of being concatenated [...] but retain much of their structural integrity [...] The only way to make a complex object that exhibits stability in the face of disruptions and accidents is to give it a hierarchical structure." Also cf. Chomsky (2007, 2008, 2013) for the No Tampering Condition, which (implicitly) incorporates Pullum's (1992) assumptions about the origins of strict cyclicity and demands that changes to existing structures are to be minimized, and ideally avoided in toto. (Still, the existence of movement makes it impossible to assume that structures can be left completely unchanged in the course of the derivation; independently of whether movement leaves a copy, a trace, nothing, or gives rise to multidominance by providing an additional mother for moved item, *some* change will have

claimed that only one of the two constraints needs to be postulated. For instance, Jacobson and Neubauer (1974) and Pullum (1979) argue that the Strict Cycle Condition is not needed if the Cyclic Principle is adopted. In what follows, I will present some arguments showing that this is not the case. Closer scrutiny reveals that there can in principle be derivations that respect the Cyclic Principle, but that are excluded by the Strict Cycle Condition, as well as derivations that respect the Strict Cycle Condition and are excluded by the Cyclic Principle. Let me start with the former.[16]

2.4.2. Case Study 1: Equi and There-Insertion in Classical Transformational Grammar

The first case study is a historical one, designed to show the incorrectness of Pullum's (1979) claim that the Strict Cycle Condition is superfluous given the Cyclic Principle against the background of the (then standard) assumptions about syntax made in that very study.[17] Thus, suppose first that control constructions are brought about by a designated transformation called *Equi NP Deletion*, according to which the subject of an embedded infinitival clause is deleted under identity with a matrix subject in the presence of a control ('equi') predicate; cf. (24a). Second, there is a transformational rule of *There-Insertion* that inserts an expletive *there* into an otherwise empty subject position; cf. (24b).[18]

(24) a. $[_{CP_1}\ [_{DP_1}$ Some students] try $[_{CP_2}$ ~~$[_{DP_1}$ some students]~~ to be in the lecture hall]]

b. $[_{CP_1}$ There are some students in the lecture hall]

The two operations can interact: Equi NP Deletion, by assumption, fully removes the lower DP (i.e., it does not just affect phonological features), and leaves an empty subject position. Consequently, Equi NP Deletion could in

occurred; see Branigan (2013).) Similar considerations also underly Watanabe's (1995) account of strict cyclicity effects based on a general Avoid Redefinition strategy.

[16]Note, though, that the following case studies first and foremost serve the purpose of illustration; nothing here should be taken to imply the correctness of the the proposals at hand.

[17]I am grateful to Philipp Weisser (p.c.) for pointing out this counter-argument.

[18]Of course, the use of DP and CP labels here is strictly speaking anachronistic; but the conclusions below hold in exactly the same way if labels like NP and S were adopted here.

principle feed subsequent There-Insertion. However, (25) shows that this is not the case.

(25) *[$_{CP_1}$ [$_{DP}$ Some students] try [$_{CP_2}$ there to be in the lecture hall]]

Importantly, a derivation giving rise to (25) is not excluded by the Cyclic Principle: On the initial CP_2 cycle, *there* cannot be inserted yet because there is still a full DP *some students* in the subject position. Subsequently, the derivation moves to the CP_1 cycle, carries out Equi NP Deletion, and, as a consequence of this, There-Insertion affecting solely the embedded cyclic domain can now apply on the CP_1 cycle. This is fully in accordance with the Cyclic Principle: There was no earlier stage of the derivation where *there* could have been inserted into the embedded subject position; therefore, such insertion respects the Cyclic Principle (in either (9) or (12)). In striking contrast to the Cyclic Principle, the Strict Cycle Condition rules out this derivation: The application of Equi NP Deletion unambiguously shows that the cyclic domain CP_1 has been affected; but subsequent application of There-Insertion exclusively affects the embedded cyclic domain CP_2, in violation of (13).

The same kind of conclusion can be drawn for a number of other interactions of operations in classical transformational grammar; cf. Perlmutter and Soames (1979). Rather than going trough these further counter-arguments based on premises which are not maintained in current approaches, I would like to present two arguments based on current minimalist approaches to syntax showing that the Strict Cycle Condition cannot simply be reduced to the Cyclic Principle.

2.4.3. Case Study 2: Intermediate Movement Steps

It is generally assumed that operations like wh-movement, which can in principle apply in an unbounded fashion in many languages, are subject to locality constraints like the Subjacency Condition in (2) (see Chomsky (1973, 1977)) or the Phase Impenetrability Condition in (26) (see Chomsky (2001)).

(26) *Phase Impenetrability Condition* (PIC):
The c-command domain of a head X of a phase XP is not accessible to operations outside XP; only X and its specifier(s) are accessible to such operations.

These locality constraints state that certain kinds of intermediate positions *must* be used in the course of long-distance movement; however, this, in and of itself, does not yet ensure that they *can* be used by intermediate movement steps. As a matter of fact, in an approach to syntax where all movement operations must have a featural trigger, it has long been recognized as a problem how intermediate movement steps as they are required by locality constraints can be triggered. It would seem that assuming the embedded C head *that* in (27a) to *always* be equipped with the relevant feature attracting the wh-phrase to the intermediate SpecC position in (27a) is problematic, given the many environments where an embedded C does not have to attract some wh-phrase to an intermdiate landing site, as in (27b). Given that the feature in question leads to ungrammaticality if it is not satisfied by attracting a wh-phrase (or some similar item involved in long-distance movement), ordinary embedded declarative clauses without any movement, as in (27b), should be prohibited throughout.

(27) a. What$_1$ did Mary say [$_{CP}$ t$'_1$ that John wanted t$_1$] ?
b. Mary said [$_{CP}$ that John wanted a book$_1$]

Various solutions to this problem have been proposed. A first solution might be to postulate that the relevant feature for intermediate wh-movement steps to declarative C heads is only optionally present on C, and not obligatorily present on this functional head. On this view, embedded C in (27a) can choose to either bear the feature, in which case a legitimate derivation can ensue, or not, in which case the derivation will crash, and ungrammaticality will arise. Similarly, embedded C in (27b) can choose to either bear the feature, leading to ungrammaticality, or not, which can then give rise to a legitimate derivation. This proposal is not innocuous, though; it will lead to a multitude of illegitimate derivations, and is therefore fundamentally incompatible with the goal of a crash-proof syntax (see Frampton and Gutmann (2002)).[19]

A second proposal designed to ensure that embedded declarative C has this feature in (27a) but not in (27b) is to invoke a concept like that of a *balanced phase* (see Heck and Müller (2003) and Müller (2011)). On this view, every phase must be "balanced" in a technical sense. Essentially, a phase qualifies as balanced iff, for every movement-inducing feature in the

[19]That said, this strategy would arguably employ the same kind of derivational branching that I suggest underlies other instances of apparent counter-cyclicity in section 3 below.

numeration (like, e.g., the relevant wh-feature of an interrogative C which is still waiting in the numeration when an embedded declarative clause is built), there is a distinct potentially available matching feature; and a feature counts as potentially available at the phase level if it is located in the edge domain of a phase (i.e., on a moved wh-phrase), or if it is also still part of the workspace (on some other wh-phrase).[20] The relevant effect – viz., of triggering insertion of a feature attracting a wh-phrase – can then be produced by a separate constraint (called Edge Feature Condition in Müller (2011)) according to which the head of a phase is assigned an edge feature if that is the only way to produce a balanced phase (i.e., when there is no item with a matching feature waiting in the workspace, and no item with the relevant feature in the edge domain of the current phase yet). This feature then shows up in (27a) (where it correctly triggers intermediate movement), but it must be absent in (27b).[21] This approach works, and can also be shown to make some interesting predictions (e.g., as concerns the existence of intervention effects that do not involve minimal c-command by the intervening item), but it looks like a deviation from an optimal design scenario since it requires the derivation to take into account information that is not locally available to it (viz., the workspace).

A third solution to the problem of triggering intermediate movement steps goes back to Preminger (2014). On this view, it simply cannot be ensured that the feature giving rise to intermediate movement steps is present only when it is needed; rather, the relevant feature is *always* present on declarative C; however, it is assumed that it can fail to trigger the operation it is supposed to trigger without giving rise to a crash (as argued independently by Preminger (2014) for probe features that trigger Agree operations).

In view of this state of affairs, let us consider a fourth option of triggering intermediate movement steps, one which is intuitively counter-cyclic. Assuming, as above, that [•F•] features are responsible for movement (or, more generally, all structure-building) operations, suppose that a declarative C head can have a feature like [•wh•] attracting a wh-phrase if it is c-commanded

[20] The workspace of a derivation comprises all trees generated thus far in the derivation, outside of the current tree, and all lexical items in the numeration.

[21] In the original formulation in Heck and Müller (2003), the requirement to have a balanced phase is not assumed to permit feature insertion on phase heads, but rather to directly permit a violation of a general constraint against non-feature-driven movement.

by interrogative C bearing [•wh•].[22] We then end up with the counter-cyclic derivation of (27a) in (28).

(28) *A counter-cyclic derivation*:

a. $[_{CP}\ [_{C}$ that] John wanted what$_1$]
b. Mary say $[_{CP}\ [_{C}$ that] John wanted what$_1$]
c. $[_{CP}\ [_{C_{[\bullet wh\bullet]}}$ did] Mary say $[_{CP}\ [_{C}$ that] John wanted what$_1$]
d. $[_{CP}\ [_{C_{[\bullet wh\bullet]}}$ did] Mary say $[_{CP}\ [_{C_{[\bullet wh\bullet]}}$ that] John wanted what$_1$]
e. $[_{CP}\ [_{C_{[\bullet wh\bullet]}}$ did] Mary say $[_{CP}$ what$_1$ $[_{C_{[\bullet wh\bullet]}}$ that] John wanted t$_1$]
f. $[_{CP}$ what$_1$ $[_{C_{[\bullet wh\bullet]}}$ did] Mary say $[_{CP}$ t$_1'$ $[_{C}$ that] John wanted t$_1$]

At the point where embedded declarative C has been merged with TP (cf. (28a)), no [•wh•] feature can be inserted on this C head because it is not yet c-commanded by an interrogative C bearing [•wh•] intrinsically. Therefore, the derivation continues as sketched in (28b), eventually merging interrogative C in (28c). Only now can the embedded declarative C head become equipped with a [•wh•] feature (cf. (28d)). Subsequently, embedded C triggers intermediate wh-movement in (28e); and finally, the wh-phrase moves to the criterial matrix SpecC position in (28f). Since Merge applying to interrogative C and TP in the matrix cyclic domain is followed by wh-movement to the embedded SpecC, which exclusively affects the embedded cyclic domain, this derivation looks counter-cyclic. However, it is clear that the Cyclic Principle is satisfied throughout: Every operation applies as soon as possible – in particular, embedded wh-movement could not have applied earlier because of a lack of [•wh•] on embedded C.

In contrast to the Cyclic Principle, the Strict Cycle Condition is violated by the derivation in (28): Whereas non-local copying of the [•wh•] feature from matrix C to embedded C in (28d) is arguably in accordance with this constraint (assuming that since the operation involves both matrix C and embedded C, it does not exclusively affect the embedded CP domain), embedded intermediate wh-movement is not: It only affects the embedded cyclic domain. Hence, whatever the merits of the approach to intermediate movement steps in (28), it can be concluded that it violates the Strict Cycle Condition (which therefore would have to be abandoned or modified if the approach were to be maintained), but not the Cyclic Principle.

[22] In addition, there must be a wh-phrase c-commanded by declarative C, and no intervening wh-phrase c-commanded by interrogative C.

2.4.4. Case Study 3: Feature Inheritance

A third example showing that the Strict Cycle Condition can exclude derivations that the Cyclic Principle is compatible with involves the concept of *feature inheritance* suggested in Chomsky (2008), Richards (2007), and much subsequent work. According to this concept, it is initially only the phase heads that have all relevant features driving syntactic operations; a phase head then passes some of them on to the head of its complement. For concreteness, C is assumed to be equipped with ϕ probe features and tense features; cf. (29a).[23] After merging with a TP, C hands these features down to T; cf. (29bc). This feature inheritance operation involves CP, and thus takes place in the cyclic domain CP. After having received the ϕ and tense features from C, T carries out agreement with the subject, via an Agree operation that values the ϕ probe on T with the relevant information from the subject DP and, in return, assigns nominative case to that DP; cf. (29d).[24] This Agree operation takes place wholly wihinin TP.

(29) a. C:{[*#:□*],[*Gen:□*],[*π:□*], [Tns:PAST]}

b. [$_{CP}$ C:{[*#:□*],[*Gen:□*],[*π:□*], [Tns:PAST]} [$_{TP}$ T [$_{vP}$ [$_{DP}$ D:{[#:pl],[Gen:fem],[π:2],[*case:□*]} ⇒

c. [$_{CP}$ C [$_{TP}$ T:{[*#:□*],[*Gen:□*],[*π:□*], [Tns:PAST]} [$_{vP}$ [$_{DP}$ D:{[#:pl],[Gen:fem],[π:2],[*case:□*]} ⇒

d. [$_{CP}$ C [$_{TP}$ T:{[#:pl],[Gen:fem],[π:2], [Tns:PAST]} [$_{vP}$ [$_{DP}$ D:{[#:pl],[Gen:fem],[π:2],[case:nom]}

Again, intuitively, feature inheritance is counter-cyclic. However, as with the two previous case studies, a feature inheritance derivation is not at variance with the Cyclic Principle: Every operation takes place as soon as it can (T cannot undergo Agree with a subject DP on the lower TP cycle since it does not have the required ϕ probe at this point). In contrast, the Strict Cycle Condition is violated by feature inheritance derivations: Merge of C and TP and the transfer of the unvalued ϕ features to T have activated the CP cycle, and subsequent Agree of T and the subject DP exclusively affects the lower TP

[23] Here, # stands for number, Gen stands for gender, π stands for person, and Tns stands for tense; [*F*] signals probe status of the feature F; and □ indicates that there is no value for the feature yet.

[24] The subject DP can either be in SpecT or in Specv, depending on whether T also has an EPP feature or not; here the latter option is pursued.

cycle. Thus, again, the two cyclicity constraints make different predictions. Consequently, if feature inheritance is to be maintained, the Strict Cycle Condition will have to be abandoned or modified whereas the Cyclic Principle can stay as it is.[25] Alternatively, the severe problem with strict cyclicity can be assumed to cast doubt on the legitimacy of feature inheritance as a syntactic concept.

2.4.5. Case Study 4: Movement and Reflexivization

The previous three case studies have focussed on derivations that respect the Cyclic Principle but are incompatible with the Strict Cycle Condition. Let me now turn to derivations that satisfy the Strict Cycle Condition but violate the Cyclic Principle.

A first relevant scenario involves a feeding interaction between wh-movement of some DP and licensing of a reflexive pronoun in that DP; cf. Barss (1986). Consider the English example in (30).

(30) $[_{TP}\ [_{DP_1}$ John] T $[_{vP}\ t_1$ wondered $[_{CP}\ [_{DP_3}$ which picture of himself$_{1,2}$] $C_{[+wh]}\ [_{TP}\ [_{DP_2}$ Bill] T $[_{vP}\ t_2$ saw t_3]]]]]

In (30), the object wh-phrase DP_3 contains a reflexive pronoun *himself*. This reflexive pronoun can be bound by the embedded subject (*Bill*$_2$), which is not particularly remarkable since DP_3 and *Bill* are co-arguments of a predicate (*saw*). However, the interesting observation is that *himself* does not have to take the embedded subject *Bill* as its antecedent; it can also legitimately be bound by the matrix subject *John*. This option is available only as a consequence of wh-movement applying to DP_3, which transports the reflexive pronoun out of the local binding domain of the embedded subject and into the local binding domain of the matrix subject. As illustrated by the illformedness of co-indexation of the reflexive pronoun and the matrix subject in (31), if there is no wh-movement, the search for an antecedent that binds it by the reflexive pronoun is confined to the embedded domain; thus, (31) shows that one cannot possibly argue that the reflexive pronoun in (30) can have a larger binding

[25] Following Richards (2007), Chomsky et al. (2019) conclude that the cyclicity problem with feature inheritance can be addressed by postulating that only phases qualify as cyclic domains (cf. (10b)), not projections, as presupposed throughout this paper (cf. (10d)); also see Kobele (2023).

domain than other reflexive pronouns, or, indeed, that it might qualify as fully exempt from binding domain restrictions.

(31) $[_{TP}\ [_{DP_1}$ John] T $[_{vP}\ t_1$ wondered $[_{CP}$ whether $[_{TP}\ [_{DP_2}$ Bill] T $[_{vP}\ t_2$ saw $[_{DP_3}$ a picture of himself$_{*1,2}$]]]]]]

Basically the same pattern arises if the wh-phrase containing the reflexive is moved to the matrix SpecC position, as in (32).

(32) $[_{CP}\ [_{DP_3}$ Which picture of himself$_{1,2}$] $[_{C_{[+wh]}}$ does] $[_{TP}\ [_{DP_1}$ John] T $[_{vP}\ t_1$ think $[_{CP}\ t'_3\ [_{C_{[-wh]}}$ that] $[_{TP}\ [_{DP_2}$ Bill] T $[_{vP}\ t_2$ liked t_3]]]]]] ?

Reflexive binding is possible from either the in-situ position of DP_3 (indicated by t_3) or the intermediate landing site in the embedded SpecC position (indicated by t'_3); the reconstruction (i.e., counter-bleeding) effect documented here is essentially identical to that seen with simple cases of wh-fronting as in (33).

(33) $[_{CP}\ [_{DP_3}$ Which picture of himself$_1$] does John$_1$ like t_1 ?

In order to see what the consequences of data like (30) and (32) for cyclicity are, let us look at how constraints on the distribution of reflexive pronouns can be implemented in the grammar. In particular, the question is how the requirement that a reflexive pronoun must find a local c-commanding co-indexed antecedent can be derived. In Chomsky (1981), this was ensured by a designated Principle A of the binding theory which basically just stated the restriction. In more recent approaches to reflexivization, it is standardly taken to follow from postulating that reflexive pronouns need to enter an Agree relation (restricted by locality constraints) with some other DP so as to provide a value for some initially unvalued feature (for instance, a binding index); see, e.g., Reuland (2001, 2011), Fischer (2004), Hicks (2009), and Murugesan (2022). In line with this, suppose that the reflexive pronoun in a sentence like (30) originally has an unvalued binding index, as in (34a). The binding index feature can be valued with an index under local Agree within the minimally dominating domain of a phase head, under c-command by a DP that provides it.[26]

[26]This presupposes that Agree can in principle be both downward (as in Chomsky (2001)) and upward (as in Bjorkman and Zeijlstra (2014)); see Baker (2008), Himmelreich (2017), Murphy

Before wh-movement of DP_3, when the embedded vP is built, this binding index can be provided by the embedded subject *Bill*, as in (34b). However, if such valuation in the embedded domain does not apply, another option arises: After wh-movement of DP_3 has taken place to the embedded SpecC position, further structure building on the matrix vP level will provide an alternative antecedent: The reflexive pronoun's binding index can now be valued by the matrix subject *John*; cf. (34d), based on failure to carry out valuation in the embedded domain in (34c).

(34) a. $[_{DP_3}$ which picture of himself$_{*\square *}]$
b. $[_{vP}\ [_{DP_2}$ Bill] saw $[_{DP_3}$ which picture of himself$_2]]$
c. $[_{vP}\ [_{DP_2}$ Bill] saw $[_{DP_3}$ which picture of himself$_{\square}]]$
d. $[_{vP}\ [_{DP_1}$ John] wondered $[_{CP}\ [_{DP_3}$ which picture of himself$_1]\ [_{C'}$ $C_{[+wh]}\ [_{TP}$ T $[_{vP}\ [_{DP_2}$ Bill] saw $t_3]]]]]$

The availability of the two different points in derivations for valuing the binding index of the reflexive pronoun offers a simple and natural account of the phenomenon at hand, viz., that movement can, but does not have to, feed reflexivization. The phenomenon at hand (which has sometimes been referred to as "pit-stop reflexives") has been the subject of intensive investigations from a variety of perspectives, e.g., with respect to the status of locality constraints on reflexivization (see, e.g., Epstein et al. (1998)), with respect to the question of which intermediate positions are targetted by movement, and can thus give rise to extended binding possibilities (see, e.g., Abels (2003) and Abels and Bentzen (2011) on punctuated vs. uniform movement paths), or with respect to its empirical status in the world's languages.[27] From the perspective of the present paper, though, it is a different question that arises: How do the derivational steps involved in movement-feeding-reflexivization scenarios fare with respect to concepts of cyclicity?

On the one hand, it can be observed that the derivational steps sketched in (34b) and (34d) both satisfy the Strict Cycle Condition. The reason is that in

and Puškar (2018), Bárány and van der Wal (2021), and Schwarzer (2022) for arguments to this effect.

[27] In fact, it is not really clear how widespread this phenomenon is. For German, e.g., there would seem to be a general consensus in the relevant earlier literature that movement in fact cannot feed reflexivization; see Frey (1993), Kiss (2001), and Büring (2005). However, based on an experimental empirical investigation, Georgi et al. (2019) show that the phenomenon can be observed in this language, too. I will come back to this in section 3 below.

neither case does the Agree operation leading to valuation of the binding index of *himself*$_{\square}$ exclusively affect a proper subdomain of the current cyclic domain (which is the embedded vP in (34b) and the matrix vP in (34d)).

On the other hand, however, the Cyclic Principle is violated by the derivation deriving (34d) on the basis of (34c): In (34c), the reflexive pronoun could have valued its binding index in the embedded vP, as in (34b); suppressing this Agree operation in the embedded vP domain and delaying it to the matrix vP domain is therefore incompatible with the Cyclic Principle. Thus, again the predictions of the Cyclic Principle and the Strict Cycle Condition do not converge.[28]

Closer inspection reveals that the different predictions of the Cyclic Principle and the Strict Cycle Condition also arise in other constructions involving

[28]It is worth pursuing the question of whether the Cyclic Principle could turn out to be compatible with the derivation based on (34c) and (34d) after all, once different basic assumptions are made. A potentially available solution might be to weaken the Cyclic Principle by reducing the number of cyclic domains (this is the approach pursued in Müller (2022, ch. 1)). So far, I have assumed that every projection is a cyclic domain (cf. (10d)), which is the most restrictive, hence optimal, solution. But suppose now that only maximal projections qualify as cyclic domains (cf. (10c)). Suppose furthermore that vP is a phase, so that, given the PIC, intermediate movement steps of wh-movement must first target Specv.

Now, given the slightly more liberal notion of cyclic domain, the Cyclic Principle does not differentiate anymore between wh-movement to Specv and binding index valuation on the reflexive pronoun by the subject DP in Specv. Consequently, binding index valuation and movement can apply in either order on the vP cycle. If Agree applies first, the reflexive pronoun will invariably be bound by the embedded subject (*Bill*, in the case at hand). However, if wh-movement to an outer Specv position applies first, the reflexive pronoun is not c-commanded anymore by the embedded subject DP, and also does not c-command the embedded subject DP itself, so that neither upward nor downward Agree is available, the reflexive's binding index feature remains temporarily unvalued, and the reflexive must and will find another antecedent in the matrix clause (here: *John*).

In contrast, under the assumption about cyclic domains made throughout the main text (where projections rather than XPs qualify as cyclic domains), the Cyclic Principle continues to block a delay of index valuation in the embedded vP: Binding index valuation applies to a cyclic domain v′ containing the subject DP, which is included in the cyclic domain vP containing the moved wh-DP$_3$ in an outer specifier of v.

Thus, it is in principle possible to reconcile the Cyclic Principle with the existence of reflexivization-feeding movement (also see Fischer (2004) for another attempt to solve the problem with the Cyclic Principle, based on a separate operation of "intensification"). However, this does not in any way affect the conclusion in the main text: Given a uniform (and maximally restrictive) notion of cyclic domain for both the Cyclic Principle and the Strict Cycle Condition, the former is violated by movement feeding reflexivization, whereas the latter is not – so the two constraints ceteris paribus do not make identical predictions.

reflexives, or anaphoric elements more generally. For instance, as noted in Grewendorf (1988), in constructions like those in (35a) and (35b), the anaphoric indirect object (a reflexive pronoun and a reciprocal pronoun, respectively) can be bound *either* by the subject *or* by the direct object.[29]

(35) a. dass [$_{vP}$ [$_{DP_1}$ der König] [$_{v'}$ [$_{VP}$ [$_{DP_2}$ den Sklaven] [$_{V'}$ sich$_{1/2}$
that the king$_{nom}$ the slave$_{acc}$ REFL
im Spiegel zeigt]] v]]
in the mirror shows

b. dass [$_{vP}$ [$_{DP_1}$ die Gastgeber] [$_{v'}$ [$_{VP}$ [$_{DP_2}$ die Gäste] [$_{V'}$
that the hosts$_{nom}$ the guests$_{acc}$
einander$_{1/2}$ vorstellen]] v]]
each other$_{dat}$ introduce

Again, the Strict Cycle Condition and the Cyclic Principle make different predictions. The Strict Cycle Condition is respected in both the derivation where the anaphoric pronoun in (35ab) is valued by the preceding, c-commanding direct object on the VP cycle, and in the derivation where the anaphoric pronoun in (35ab) is valued by the subject on the vP cycle. However, the Cyclic Principle is not compatible with the co-existence of the two derivations; it favours the one where binding index valuation via Agree takes place in the lower VP domain, and thus excludes regular binding by the subject in (35ab).

This conclusion can be generalized: *All* instances of optionality in binding of reflexives and reciprocals will ceteris paribus give rise to a problem for the Cyclic Principle (but not for the Strict Cycle Condition) because one of the two possible antecedents will always be located in a more remote cyclic domain (given that all syntactic structures are binary branching, and that the minimal projection is the cyclic domain for the Cyclic Principle; cf. footnote 28).[30]

[29]Two remarks. First, I assume here, based on the arguments in Müller (1995, 1999*b*), that the order of direct object before indirect object is uniformly the base order of arguments in German; but the main conclusions do not change if that order is assumed to be derived by scrambling, as in Webelhuth (1992). And second, whereas there would seem to be a complete consensus among speakers regarding the availability of binding of the reciprocal by the direct object in (35b), there is some variation among speakers with respect to the legitimacy of binding by the object in (35a), with some speakers actually preferring the reverse scenario, where a direct object reflexive can be bound by an indirect object antecedent (cf. Featherston and Sternefeld (2003)). These qualifications do not affect the point to be made here, viz., that there can be optionality of binding in double object constructions.

[30]Accordingly, these kinds of phenomena have sometimes been taken to indicate that a

To give just one more example from German: The famous case of optional long-distance binding of reflexives (and reciprocals) in German exceptional case marking (accusativus cum infinitivo) constructions (cf. Reis (1976), Grewendorf (1983), Gunkel (2003), and Barnickel (2014)) is an instance of the same pattern: The PP-internal reflexive pronoun *sich* in (36) can be bound by the embedded subject DP *Paul*, in accordance with the Cyclic Principle, or by the matrix DP *Maria*, in violation of this constraint; in turn, the Strict Cycle Condition is respected by both derivations.

(36) dass Maria$_1$ [$_{TP}$ Paul$_2$ [$_{PP}$ bei sich$_{1/2}$] schlafen] lässt
that Maria$_{nom}$ Paul$_{acc}$ with REFL sleep lets
'that Maria lets Paul sleep at her/his place.'

2.4.6. Case Study 5: Object Shift and EPP-Movement in Non-Monotonic Derivations

Vikner (1989) discusses a dilemma arising if one makes the (standard) assumptions that (i) object shift in (continental) Scandinavian languages is movement of a pronoun to an outer Specv position, and (ii) that there is obligatory EPP-driven movement of a subject DP to SpecT in these languages, as in (37).

(37) I går læste$_3$ [$_{TP}$ Ole$_1$ T [$_{vP}$ den$_2$ [$_{v'}$ t$_1$ [$_{v'}$ v [$_{VP}$ uden tvivl
yesterday read Ole it without doubt
ikke t$_3$ t$_2$]]]]]
not
'Yesterday, Ole doubtlessly didn't read it.'

The problem arising with (37) is that it is not obvious how both the Minimal Link Condition in (38) (cf. Fanselow (1990, 1991), Ferguson and Groat (1994), and Chomsky (1995, 2001), among many others, with notation adapted to assumptions about feature-driven movement made above) and the Strict Cycle Condition can be respected in a derivation producing (37).

---

constraint like Principle A of the binding theory is an "anywhere" principle, i.e., a global constraint (in the sense of Lakoff (1970)), where all steps of a complete derivation must be taken into account to determine whether the constraint is violated or respected. See, e.g., Belletti and Rizzi (1988) on reflexivization in Italian psych verb constructions, or, more recently, Privizentseva (2022*a*) on reflexivization in Moksha Mordvin relative clauses with inverse case attraction.

(38) *Minimal Link Condition* (MLC):
In a structure $\alpha_{[\bullet F \bullet]}$... [... $\beta_{[F]}$ ... [... $\gamma_{[F]}$ ...] ...] ..., movement to [•F•] can only affect the category bearing the [F] feature that is closer to [•F•].

Consider first a derivation of (37) where object shift of the pronoun to an outer Specv position precedes subject movement to SpecT, as in (39).[31]

(39) a. *Pre-movement structure*:
$[_{vP}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]$
b. *Object shift*:
$[_{vP}\ DP_2\ [_{v'}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ t_2\ ]]]]$
c. *Merge of T*:
$[_{TP}\ T\ [_{vP}\ DP_2\ [_{v'}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ t_2\ ]]]]]$
d. *EPP-movement of the subject*:
$*[_{TP}\ DP_1\ [_{T'}\ T\ [_{vP}\ DP_2\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ t_2\ ]]]]]]$

Object shift in (39b) is unproblematic from the perspective of the Minimal Link Condition (the subject in Specv does not intervene at this point, in the sense of (38)). However, subject movement to SpecT in (39d) should be blocked by the object in the outer Specv position since the latter is now closer to T than the subject in the lower Specv position. Thus, it seems that the only way for subject movement to comply with the Minimal Link Condition is to postpone object shift until the subject has undergone EPP-movement, as in the derivation in (40). However, in this derivation, the final object shift operation clearly violates the Strict Cycle Condition.[32]

(40) a. *Pre-movement structure*:
$[_{vP}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]$
b. *Merge of T*:
$[_{TP}\ T\ [_{vP}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]]$
c. *EPP-movement of the subject*:
$[_{TP}\ DP_1\ [_{T'}\ T\ [_{vP}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]]]$
d. *Object shift*:
$*[_{TP}\ DP_1\ [_{T'}\ T\ [_{vP}\ DP_2\ [_{v'}\ t_1\ [_{v'}\ v\ [_{VP}\ V\ t_2\ ]]]]]]$

[31] For now, I abstract away from V movement, which is obligatory in object shift environments; see below.
[32] This derivation also violates the Cyclic Principle when T is merged. See below.

This dilemma is what Heck (2016) refers to as Vikner's Puzzle. The solution to this problem advanced in Heck (2016) relies on a concept of non-monotonic derivations, according to which syntactic trees may have to temporarily shrink before growing again. On this view, syntactic movement is not a primitive, homogeneous operation, but rather composed of two steps in the case of XP movement: First, an item is taken from the current tree, and placed in the workspace of the derivation (triggered by a [•F•] feature on some head); and second, the item is subsequently taken from the workspace again, and merged in the target position. An instance of head movement via adjunction, in turn, requires three steps (also cf. Bobaljik and Brown (1997)): The attracting head is put in the workspace; the attracted head then combines with it (which circumvents the c-command problem otherwise existing with head movement as adjunction);[33] and finally the attracting head (now complex) is remerged. Crucially, between the various suboperations of a given complex movement operation, other syntactcic operations can in principle take place. As shown by Heck (2016, ch. 4), a non-monotonic derivation makes it possible to have a derivation of examples like the one in (37) that respects both the Minimal Link Condition and the Strict Cycle Condition. Such a derivation looks as in (41).[34]

(41) a. *Pre-movement structure*:
$[_{vP}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]$

b. *Merge of T*:
$[_{TP}\ T\ [_{vP}\ DP_1\ [_{v'}\ v\ [_{VP}\ V\ DP_2\ ]]]]$

c. *First step of EPP-movement – DP_1 to workspace*:
$[_{TP}\ T\ [_{vP}\ v\ [_{VP}\ V\ DP_2\ ]]]$ $\boxed{DP_1}$

d. *First step of v-to-T movement – T to workspace*:
$[_{vP}\ v\ [_{VP}\ V\ DP_2\ ]]$ $\boxed{DP_1}, \boxed{T}$

e. *First step of object shift – DP_2 to workspace*:
$[_{vP}\ v\ [_{VP}\ V\ ]]$ $\boxed{DP_1}, \boxed{T}, \boxed{DP_2}$

[33] Incidentally, head movement as adjunction is another problem arising from the perspective of cyclicity; in fact, it belongs to the first group of asymmetric phenomena discussed in this section since it is incompatible with the Strict Cycle Condition (under a narrow understanding of cyclic domains) but compatible with the Cyclic Principle (it cannot take place before the attracting head has entered the structure).

[34] Prior V-to-v movement is presupposed throughout but not indicated here. Material in the workspace of the derivation shows up in a $\boxed{\text{box}}$.

f. *Second step of object shift – DP_2 to Specv*:
$[_{vP}\ [_{v'}\ DP_2\ [_{v'}\ v\ [_{VP}\ V\ ]]]]$ $\boxed{DP_1}$, $\boxed{T}$,

g. *Second step of v-to-T movement – v to T in workspace*:
$[_{vP}\ [_{v'}\ DP_2\ [_{v'}\ [_{VP}\ V\ ]]]]$ $\boxed{DP_1}$, $\boxed{T\text{-}v}$,

h. *Third step of v-to-T movement – T-v is remerged*:
$[_{TP}\ [_{T}\ T\ v\ ]\ [_{vP}\ [_{v'}\ DP_2\ [_{v'}\ [_{VP}\ V\ ]]]]]$ $\boxed{DP_1}$

i. *Second step of EPP-movement – DP_1 to SpecT*:
$[_{TP}\ DP_1\ [_{T'}\ [_{T}\ T\ v\ ]\ [_{vP}\ [_{v'}\ DP_2\ [_{v'}\ [_{VP}\ V\ ]]]]]]$

It can be verified that the derivation in (41) respects the Minimal Link Condition; in particular, when T attracts DP_1 to the workspace in (41c), and when v attracts DP_2 to the workspace in (41e), there is no closer, intervening item. Furthermore, the derivation respects the Strict Cycle Condition: At no point of the derivation is there a step that affects a cyclic subdomain of the currenct root domain. However, whereas the Strict Cycle Condition is satisfied by the derivation in (41), the Cyclic Principle is not (as is in fact noted in Heck (2016)): Exactly as in the original counter-cyclic derivation in (40a) (see footnote 32), the Cyclic Principle is violated by merging T in (41b): By assumption, v has a feature triggering object shift, so object shift could in principle apply in the first step, but it is postponed here to (41e), in violation of the Cyclic Principle.

2.4.7. Case Study 6: Cyclicity and Partially Superfluous Extended Exponence

The third and final example illustrating that derivations can respect the Strict Cycle Condition but violate the Cyclic Principle comes from morphological exponence. Many approaches to inflectional morphology recognize the concept of cyclicity.

For instance, Wunderlich's (1997) Minimalist Morphology is a lexical-incremental approach where each instance of morphological exponence involves genuine structure-building; and such structure-building is subject to a cyclicity requirement.

The same goes for the lexical-realizational approach based on Harmonic Serialism that is developed in Müller (2020), which derives inflectional exponence by iterated Merge operations that are subject to cyclicity.

Next, Stump's (2001) inferential-realizational model of Paradigm Functional Morphology treats exponence by a succession of paradigm functions where

(starting with the root) each function maps a given form/property-set pairing (FPSP) in a given rule block to a modified form/property-set pairing in the next rule block, which is then mapped to another property set pairing in a new rule block, and so on, until the ordered list of rule blocks has been exhausted and the final form/property-set pairing is achieved (which then qualifies as the ultimate realization of a given paradigm cell). Since, by definition, a paradigm function in a given rule block thus makes use of the output of the paradigm function in the previous rule block, Paradigm Function Morphology can be viewed as having built in the concept of cyclicity at its very heart.

Finally, cyclicity has also been regularly adopted within the lexical-realizational theory of Distributed Morphology (cf. Halle and Marantz (1993)); see Bobaljik (2000), Adger, Béjar and Harbour (2003), Embick (2010), Kalin and Weisser (2021), and Privizentseva (2022*b*), among others. Unlike what is the case with morphological exponence in the approaches in Wunderlich (1997) and Müller (2020), morphological exponence in Distributed Morphology does not involve genuine structure-building operations; rather, it is brought about by *vocabulary insertion*, a substitution transformation that inserts a morphological exponent into an abstract functional head (a 'morpheme').[35]

In order to find out whether cyclicity is respected in a derivation, it is imperative to determine the currently active cyclic domain at any given step. This is straightforward if the creation of a cyclic domain results from structure-building; for this reason, it was possible to simplify the formulation of the original Cyclic Principle in (9) as in (12). However, things are not quite the same if the whole structure is present to begin with, and cyclicity is supposed to ensure that operations (like, in particular, vocabulary insertion, but also other operations modifying morphemes or exponents as they are envisaged in Distributed Morphology) apply inside-out, from bottom to top. In such a model (which is also the one underlying classical transformational grammar), (9) must be adopted as the formulation of the Cyclic Principle (and analogously (13) rather than (14) as the formulation of the Strict Cycle Condition). Thus, given cyclicity constraints, in an abstract complex head representation like (42a), post-syntactic vocabulary insertion must first apply to the most deeply

[35]There is but one exception: It has sometimes been argued that so-called *dissociated morphemes* can post-syntactically enter morphological structures (counter-cyclically) before morphological exponence; see Halle and Marantz (1993), Embick (1998), and Embick and Noyer (2001).

embedded root and c (i.e., categorizing head) nodes (cf. (42b), where α and β can be inserted in either order), then to X (cf. (42c), with the exponent γ inserted), and finally to Y (cf. (42d), where the vocabulary item δ is inserted).

(42) a. $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X\ ]\ Y\ ]$
b. $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ [_{c}\ \beta\ ]]\ X\ ]\ Y\ ]$
c. $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ [_{c}\ \beta\ ]]\ [_{X}\ \gamma\ ]]\ Y\ ]$
d. $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ [_{c}\ \beta\ ]]\ [_{X}\ \gamma\ ]]\ [_{Y}\ \delta\ ]]$

The references cited above contain a number of arguments for cyclicity in post-syntactic morphological exponence in Distributed Morphology. However, in general it would seem that these analyses are neutral between the Cyclic Principle and the Strict Cycle Condition. From the present perspective, the interesting question is whether asymmetries can be shown to arise between the two constraints currently under consideration. An argument to this effect is given in Grofulović and Müller (2023).

That study sets out to derive, in Distributed Morphology, a particular generalization about *partially superfluous extended exponence* (cf. Caballero and Harris (2012) for the term), i.e., scenarios where a given morpho-syntactic feature in a word is realized by two separate morphological exponents /a/ and /b/, where the morpho-syntactic features inherently associated with /a/ (e.g., $[f_1]$) are a subset of the morpho-syntactic features borne by /b/ (e.g., $[f_1,f_2]$). The generalization to be derived is that in such a situation, the more general of the two exponents, i.e., /a/, must be inserted first, and closer to the stem, than the more specific exponent, i.e., /b/; cf. (43).[36]

(43) *The Partially Superfluous Extended Exponence Generalization*:
If there are two exponents /a/$\leftrightarrow[f_1]$ and /b/$\leftrightarrow[f_1,f_2]$ in a word, /a/ is realized closer to the stem than /b/.

A relevant example of partially superfluous extended exponence involves number marking on nouns in Archi. The ergative plural form of a noun stem

[36]The underlying rationale is that /a/ would in principle emerge as superfluous, and be blocked by economy considerations (of one type or the other), given that /b/ realizes the same features as /a/, and more. Hence, the only chance for /a/ to legitimately occur in the word is to be subject to exponence at an early stage of the derivation (i.e., close to the stem), where /b/ is not yet available. Alternative accounts of (43) that implement the same hypothesis on the basis of other theoretical models are Caballero and Inkelas (2013), Stiebels (2015), and Müller (2020).

qIinn ('bridge') is *qIinn-or-čaj* ('bridge-PL-ERG.PL'). Here, *or* is a pure plural exponent (/or/↔[+pl]); *čaj* is an exponent that realizes both plural and ergative case (/čaj/↔[+pl,erg] – note that the pure ergative case exponent would be *(l)i*); and the more general plural exponent *or* is realized closer to the stem than the more specific ergative plural exponent *čaj*.[37]

The central background assumption made in Grofulović and Müller (2023) is that each instance of extended exponence requires the application of an operation that post-syntactically copies the feature in question that is realized more than once (so-called *enrichment*; cf. Müller (2007)). In interaction with the Cyclic Principle, this assumption then ensures that a derivation that is at variance with (43) will always be ruled out: Suppose that there is a derivation in which the more general exponent is inserted second, after the more specific exponent. In such a derivation, the required additional copy of a morpho-syntactic feature only has a chance to be generated without violating the Cyclic Principle at a point where the feature is already gone as a consequence of earlier insertion of the more specific exponent. Therefore, the required copy can never be generated, and there can be no extended exponence in this scenario.

For concreteness, consider first an abstract legitimate derivation of partially superfluous extended exponence, where the more general exponent is inserted before the more specific one; cf. (44).

(44) a. *Initial structure*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

b. *Root lexicalization*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

c. *Feature copying on X cycle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_1]}\ ]\ Y_{[f_2]}\ ]$

d. *Vocabulary insertion on X cycle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}\ /a/\ ]]\ Y_{[f_2]}\ ]$

e. *Vocabulary insertion on Y cycle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /a/\ ]]\ [_{Y}\ /b/\ ]]$

[37]There are exceptions to the generalization in (43) which I will not be concerned with here. Arguably, most of these exceptions can be insightfully addressed by employing the concept of movement of morphological exponents (so that the generalization holds true of representations before morphological movement); cf. Müller (2020), Gleim et al. (2022), and Grofulović and Müller (2023).

The initial structure arising after complex head formation (either in the syntax or in the post-syntax) is given in (44a); by assumption, the functional morphemes X and Y bear the features $[f_1]$ and $[f_2]$, respectively (this could stand for plural and ergative in the Archi example discussed above). In (44b), the derivation starts by inserting a vocabulary item into the root node (and, perhaps, another vocabulary item into the categorizing head, which is not indicated here and in what follows). In (44c), the derivation moves to the X cycle, and copies the feature $[f_1]$ on X (this is accomplished by a designated $[f_1]$-enrichment rule that can apply in this environment). After this, in (44d), the (more general) morphological exponent /a/↔$[f_1]$ is inserted into X. By assumption, vocabulary insertion leads to a discharge (i.e., removal) of matched features in the syntactic context (see Noyer (1997), Trommer (1999), and Bobaljik (2000)); so one of the two $[f_1]$ features is now gone from the representation. Finally, the more specific item /b/↔$[f_1,f_2]$ is inserted into Y; cf. (44e). To satisfy the compatibility requirement incorporated into the Subset Principle (cf. Halle and Marantz (1993)), /b/ must find both $[f_1]$ and $[f_2]$ in the syntactic environment; it does (the former in X, the latter in the Y head into which it is inserted), and the two context features are deleted.

In contrast, any derivation in which /a/↔$[f_1]$ is inserted after /b/↔$[f_1,f_2]$ will lead to illformedness. Among the derivations that need to be (and can be) excluded is a counter-cyclic one that proceeds as in (45). Here the only difference in the initial representation is that $[f_1]$ is now located on Y, and $[f_2]$ on X; cf. (45a). After root lexicalization in (45b), feature copying takes place, providing a second $[f_1]$ on Y; cf. (45c). Such a second $[f_1]$ must be present on Y because otherwise /a/ can never satisfy the Subset Principle (and /a/, unlike /b/, cannot be inserted into X because it does not have the $[f_2]$ feature that is now located in X). If, subsequently, vocabulary insertion of /b/ into X, and of /a/ into Y, could take place (as in (45de)), the derivation could give rise to an instance of partially superfluous extended exponence that contradicts the generalization in (43) by realizing the more general of the two exponents in the outer position after all.

(45) a. *Initial structure*:
$[_Y\ [_X\ [_c\ \sqrt{\ }\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

b. *Root lexicalization*:
$[_Y\ [_X\ [_c\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

c. *Feature copying on Y cycle* **Cyclic Principle*
$[_Y\ [_X\ [_c\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_1]}\ ]$
d. *Vocabulary insertion into X on Y cycle*:
$[_Y\ [_X\ [_c\ \sqrt{\alpha}\ c\ ]\ [_X$ /b/ $]]\ Y_{[f_1]}\ ]$
e. *Vocabulary insertion into Y on Y cycle*:
$[_Y\ [_X\ [_c\ \sqrt{\alpha}\ c\ ]\ [_X$ /b/ $]]\ [_Y$ /a/ $]]$

However, the instances of vocabulary insertion in (46de) cannot take place because the step in (46c) violates the Cyclic Principle: Feature copying in (45c) clearly applies to the cyclic domain Y (since the $[f_1]$ feature in question is in Y), but the derivation could have inserted /b/ into the smaller cyclic domain X first (which would then discharge $[f_1]$ in Y before a copy can be made, and thus preclude subsequent insertion of /a/). Thus, invoking cyclicity is crucial in this account of the Partially Superfluous Extended Exponence Generalization. What is more, though, it is only the Cyclic Principle that achieves this. The Strict Cycle Condition turns out to be respected by the derivation in (45): After the derivation has affected the cyclic domain Y in step (45c), it does not exclusively affect a proper subdomain of Y at a later step; to wit, vocabulary insertion of /b/ in (45d), while applying on the X cycle, also affects the Y cycle by discharging $[f_1]$ there; and final insertion of /a/ in Y of course both applies to, and affects, the Y cycle. So, we have a third case of a derivation that violates the Cyclic Principle but satisfies the Strict Cycle Condition.

2.4.8. Interim Conclusion

Cyclicity plays an important role in excluding certain derivations that need to be excluded; but, as we have seen, there are two standard cyclicity constraints that yield effects of this type, viz., the Cyclic Principle and the Strict Cycle Condition. As a consequence, the question arises whether one of the two can be dispensed with in favour of the remaining constraint. The answer is that this is not the case: The Strict Cycle Condition can be shown to rule out derivations that are compatible with the Cyclic Principle (cf. case studies 1-3), and the Cyclic Principle can be shown to rule out derivations that are compatible with the Strict Cycle Condition (cf. case studies 4-6).[38] Consequently, at least for

[38]There are more cases of this type. For instance, in the approach to direct/inverse marking in Potawatomi morphology based on morphological movement developed in Andermann

the time being, I will draw the conclusion that *both* constraints are active in derivations in syntax and morphology.

This implies that accounts of phenomena relying on legitimate operations that violate either the Cyclic Principle or the Strict Cycle Condition cannot be maintained (also see Chomsky (2019) for this type of conclusion). This holds, e.g., for Late Merge and Wholesale Late Merge (cf. section 2.3.2); for Feature Inheritance (cf. section 2.4.4); for the standard approach to movement feeding reflexivization and, more generally, for optionality in antecedent choice (cf. section 2.4.5); for non-monotonic derivations (cf. section 2.4.6); and for the concept of head-movement as adjunction of one head to another. However, it can be noted that for most of these cases, alternative accounts that respect both the Cyclic Principle and the Strict Cycle Condition are available. See, e.g., Kuno (1972, 1987), Riemsdijk and Williams (1981), Huang (1993), Fischer (2004), Chomsky (2004), and Bruening and Al Khalaf (2019) on phenomena that have been taken to motivate a concept like Late Merge; or Chomsky (1995), Matushansky (2006), Fanselow (2003), and Georgi and Müller (2010) for some alternatives to head movement as adjunction (which are still fairly conservative in that they do not reconceptualize head movement as phrasal movement).[39]

To end this section, let me briefly address the issue of weaker and stronger versions of cyclicity, as they may arise by modifying the choice of cyclic domain (recall (10)), or in some other way. It can be noted that for the Strict Cycle Condition in particular, both weaker and stronger versions have been proposed.

A weaker version of the Strict Cycle Condition in (13) is the *Peak Novelty Condition* proposed in Safir (2019), which permits operations which are not *massively* counter-cyclic – i.e., which take place reasonably close to the current root domain. A similar type of weaker version of the Strict Cycle Condition in (13) is adopted in Müller (2022), so as to permit a removal of syntactic strucure

(2023), the Strict Cycle Condition (but not the Cyclic Principle) ensures that the operation of exponent removal (of *mɔn* by *nan*) is strictly local (and does not affect more deeply embedded exponents); and the Cyclic Principle (but not the Strict Cycle Condition) guarantees that instances of morphological movement that relocate exponents to an edge of the word proceed bottom-up.

[39]That said, in some cases, the availability of an alternative approach may not be entirely obvious. For instance, this holds for the problem with the Cyclic Principle incurred by optional and movement-induced reflexivization. See below.

(heads or phrases) that is located in the domain (in the sense of Chomsky (1995)) of the head of the current root node. These weaker versions of the Strict Cycle Condition also permit tucking in in the sense of Richards (2001), i.e., movement to a non-highest specifier position of the current root node.[40]

Then again, there are also stronger versions of the Strict Cycle Condition. A relevant concept is *Bracket Erasure* for morphology (see Chomsky and Halle (1968), Pesetsky (1979), Kiparsky (1982*a*)). On this view, after a cycle of structure-building in morphology is completed, phonological operations apply; and at the end of the phonological cycle, all morphological structure is removed, so that a subsequent morphological cycle cannot look into the word generated thus far. A related concept from syntax is *Multiple Spell-Out* (see Uriagereka (1999), Chomsky (2001)): Here, the assumption is that after a phase is completed in the syntax, the complement of the phase head is sent off to the phonological and semantic interfaces; the structure is thereby flattened and/or removed. As noted by Katamba (1993), Bracket Erasure is a stronger version of the Strict Cycle Condition since it does not permit any access to the internal structure of a linguistic object subjected to it. Similarly, Multiple Spell-Out is a stronger version of the Strict Cycle Condition because material properly contained in the spelled-out object can never be accessed anymore, not even by operations that also access structure outside of the spelled-out domain. Furthermore, under the radical, unified approach to cyclicity pursued in Kobele (2023, sect. 4), every operation that accesses a subtree of a current tree is counter-cyclic; thus, on this view, all instances of movement (conceived of as internal Merge) strictly speaking qualify as counter-cyclic operations.

For now, I will leave open the question of whether weaker or stronger versions of the Strict Cycle Condition (or, for that matter, the Cyclic Principle; cf. footnote 28) may ultimately be required, and continue to assume the versions of the cyclicity constraints in (9) and (13). On this basis, I will address a second challenge for concepts of cyclicity: the existence of apparently counter-cyclic repair operations.

[40] Also cf. Streffer (2023) on incorporation from specifier in Turkana, which requires a weaker version of the Strict Cycle Condition, and which in fact distinguishes between the version of the constraint in Safir (2019) and the version of the constraint in Müller (2022).

3. Counter-Cyclic Repair by Cyclic Derivational Branching

3.1. Repair Operations

A repair in grammatical theory is an operation that is normally blocked, but that can take place under special circumstances where the regular output would violate some constraint α. Standardly, the concept of repair is modelled in such a way that the repair operation intrinsically violates some constraint β. This normally suffices to preclude application of the operation, except for specific environments where otherwise constraint α would have to be violated; in this case, as a last resort, β can be minimally violated by the repair so as to satisfy α. This thus presupposes a general violability of constraint β in favour of a compliance with α. Accordingly, faithful implementations of the concept of grammatical repair typically rely on optimality theory, where constraints are assumed to be violable and ranked (cf. Prince and Smolensky (2004)). They do so either explicitly (as in Grimshaw (1997)) or implicitly (as in analyses invoking a concept like "last resort"); cf. Heck (2022). A classical case of repair in syntax is the existence of *do*-support in English root non-subject wh-questions (cf. (46a)) and negation environments (cf. (46b)).

(46) a. What$_1$ did she buy t$_1$?
b. Mary did not buy a book
c. *Mary did buy a book

As argued by Grimshaw (1997), *do* cannot normally appear in (non-emphatic, non-negated) declarative environments (cf. (46c)) because its presence violates a constraint against semantically uninterpretable, expletive items; however, if other, higher-ranked constraints can only be fulfilled in the presence of a finite auxiliary verb, and there is no alternative auxiliary available, *do*-support becomes legitimate.

The repair phenomena I want to address in what follows can all be dealt with in basically this way, by postulating minimal violability of the constraint blocking the repair in favour of a satisfaction of a higher-ranked constraint (or a set of higher-ranked constraints). So, in this respect, the phenomena to be discussed below are all reasonably well-behaved. However, in addition to instantiating repair, they exhibit an interesting property from the perspective of cyclicity: They involve repair operations that look like they must be counter-cyclic because the relevant pieces of information are only provided in later

(i.e., higher) cycles, and are not available at the point where it looks as though the decision about the repair must be made.

In view of this challenge, I would like to propose that such cases of apparently counter-cyclic repair should be reanalyzed as strictly cyclic repair by postulating that a decision can in fact be made at the early stage, i.e., before the trigger for the repair actually shows up. This implies that the initial decision about the legitimacy of the repair has to be a *preliminary* one: After the tentative decision has been made, two alternatives are subsequently being pursued in parallel; eventually, the initial repair (which is strictly speaking unmotivated at the point where it is carried out) will be successful only if it can be motivated at some point.

3.2. Movement and Reflexivization in German

The first instance of seemingly counter-cyclic repair to be discussed here is the case of movement feeding reflexivzation discussed in section 2.4.5 above. Recall from footnote 27 that Georgi et al. (2019) have shown that, in contrast to claims in the earlier literature (cf. Frey (1993), Kiss (2001), and Büring (2005)), the phenomenon is not confined to English but also shows up in German; cf. their examples in (47ab) (which parallel English examples like (30) and (32)).[41]

(47) a. [$_{CP_0}$ Maria$_3$ erzählt [$_{CP_1}$ [$_{DP_2}$ welche Statue von sich$_{3,4}$]
Maria recounts which statue$_{acc}$ of REFL
Anna$_4$ t$_2$ gesehen hat]]
Anna$_{nom}$ seen has

b. Maria$_3$ erzählt [$_{CP_0}$ [$_{DP_2}$ welche Statue von sich$_{3,4}$]
Maria$_{nom}$ recounts which statue$_{acc}$ of REFL
Anna$_4$ denkt [$_{CP}$ t$'_1$ dass du t$_2$ gesehen hast]]
Anna$_{nom}$ thinks that you$_{nom}$ seen have

The illformedness of (48), where there is no movement of the DP containing the reflexive pronoun, shows that the reflexive pronouns are not exempt from

[41] That said, everything that follows will automatically extend to the English data in section 2.4.5. Also note that I will have nothing to say about apparent violations of the Cyclic Principle incurred by reflexivization in double object constructions, as in (35).

finding a local c-commanding antecedent; the domain extension in (47ab) is indeed due to wh-movement.

(48) *[$_{CP_0}$ Maria$_3$ erzählt [$_{CP_1}$ dass Anna$_4$ [$_{DP_2}$ die Statue von
Maria recounts that Anna$_{nom}$ the statue$_{acc}$ of
sich$_3$] gesehen hat]]
REFL seen has

The possibility of binding of the reflexive pronoun by *Maria*$_3$ in (47a), and by both *Anna*$_4$ and *Maria*$_3$ in (47b), looks counter-cyclic: The Cyclic Principle would demand that the reflexive can only be bound by the local subject that is its initial clause-mate. As noted in footnote 28, it may in principle be possible to reconcile these data with the Cyclic Principle by weakening it (such that the cyclic domains are larger). However, it is worth investigating whether an alternative approach is available that does not necessitate such a weakening.

To this end, suppose first that binding of a reflexive pronoun contained in a DP by an antecedent that is located outside of DP is always a repair operation. This follows without further ado if one makes the following five assumptions (i)-(v).

(i) First, binding of a reflexive pronoun involves an Agree operation (cf. Reuland (2001, 2011), Fischer (2004), Hicks (2009), and Murugesan (2022), among others). More specifically, a reflexive (or reciprocal) pronoun has a binding index probe that needs to be valued by (upward) Agree with some DP that can provide a binding index.

(ii) Agree operations are not generally subject to the PIC in (26). This view has been put forward by Bošković (2007) and Keine (2016), among others, for standard Agree operations like those involving ϕ-features.[42] The

[42]This conclusion suggests itself on the basis of very simple data, such as sentences exhibiting agreement with respect to ϕ-features between T and the sole argument of an unaccusative verb dominated by VP, in a language where T does not have an *obligatory* EPP property, like German.

(i) dass [$_{TP}$ [$_{vP}$ [$_{DP_1}$ ihm] [$_{VP}$ t$_1$ [$_{DP_2}$ ein Fehler] unterlaufen ist]] T]
that him$_{dat}$ a mistake$_{nom}$ occurred to is

Here, the presence of the unstressed pronoun *ihm* ('him'), which must show up at the left edge of vP (where it can only be preceded by a nominative DP that has undergone *optional* EPP-driven movement to SpecT; cf. Müller (2001)) signals that the nominative DP$_2$ *ein Fehler* ('a mistake') has remained in its base position, viz., the complement position of an unaccustive

assumption is virtually unavoidable if Agree is to also hold for binding relations involving non-reflexive (and non-reciprocal) pronouns, which can be taken to be intrinsically equipped with an index but may in many environments (e.g., in contexts where they are supposed to be interpreted as bound variables) *have* to enter a binding relation. Given assumption (i), this can be taken to imply that reflexive (and reciprocal) pronouns are defective in that they initially have an *unvalued* binding index probe ([*□*]), whereas other (bound-variable) pronouns are not defective and have a valued binding index probe (like [*1*]).

(iii) An Agree operation involving an unvalued binding index feature *is* subject to the PIC. This is essentially the residue of Principle A of the binding theory (see Chomsky (1981)).

(iv) DP is a phase (cf., e.g., Svenonius (2004) and Matushansky (2005)).

(v) Finally, failure to find an antecedent that might value the binding index probe feature of a reflexive (or reciprocal) pronoun within the minimal DP phase does not (necessarily) lead to ungrammaticality, but may trigger a repair operation.

Before addressing the question what this repair operation might look like, a first consequence arising under these assumptions can be noted: The prediction is that domain extension under movement is available only for a reflexive pronoun that is part of a DP, not for a reflexive pronoun that shows up as an argument of a verb; in these latter cases, a reflexive will find a possible binder within the minimal vP phase.[43] This prediction is borne out; see (49ab), where a reflexive and a reciprocal pronoun argument of a verb undergo topicalization but can never acquire a new antecedent as a result of this movement step.[44]

(49) a. [$_{DP}$ Sich$_{*2,1}$ (selbst)] denkt Maria$_2$ [$_{CP}$ t′ dass Karl$_1$ t
REFL$_{acc}$ self thinks Maria$_{nom}$ that Karl$_{nom}$
einladen will]
invite wants to

verb (see Grewendorf (1989)). Still, an Agree operation can take place between T and DP$_2$, across the vP phase.

[43] In addition, scenarios where the reflexive is the highest argument of the verb are covered by whatever derives the anaphor agreement effect in a language like German; cf. Rizzi (1990) and Murugesan (2019, 2022) (and references cited there).

[44] (49b) is independently somewhat degraded because of the marked status of topicalization of a reciprocal pronoun; but the binding asymmetry is clearly discernible. (The intended, but unavailable, interpretation would be something like "Each student thinks that the professors should not harm the other students.")

b. ??[$_{DP_2}$ Die Studenten] finden [$_{CP}$ [$_{DP}$ einander$_{*2,1}$] sollten [$_{DP_1}$
the students think each other should
die Professorinnen] nicht schaden]
the professors not harm

So, what could the repair operation applying if a DP-internal reflexive does not find an antecedent that might value its index probe consist in? A possible answer suggests itself if one adopts the proposal that Agree is not a primitive operation, but needs to be decomposed into two separate parts (see Arregi and Nevins (2012), Doliana (2013), and Himmelreich (2017)): First, there is *Agree-Link*, which establishes a link between a probe and a goal feature, and can be taken to remove the former feature's probe status (indicated by absence of *); and second, there is *Agree-Copy*, which transfers the value of the goal feature to the unvalued feature that initially had the probe property. Normally, the two suboperations of Agree apply in this order, as illustrated abstractly in (50) for a standard case of reflexivization among co-arguments (with the box notation indicating the link).[45]

(50) a. *Initial representation*:
[$_{vP}$ DP$_1$ [$_{v'}$ v [$_{VP}$ V [$_{DP}$ REFL$_{[*\square*]}$]]]]
b. *Agree-Link*:
[$_{vP}$ $\boxed{\text{DP}_1}$ [$_{v'}$ v [$_{VP}$ V [$_{DP}$ $\boxed{\text{REFL}_{[\square]}}$]]]]
c. *Agree-Copy*:
[$_{vP}$ $\boxed{\text{DP}_1}$ [$_{v'}$ v [$_{VP}$ V [$_{DP}$ $\boxed{\text{REFL}_{[1]}}$]]]]

However, suppose now that Agree-Link fails because the unvalued binding index feature of a reflexive (or reciprocal) pronoun in a DP phase cannot find a suitable goal within the phase, as required by the PIC; cf. (51ab).[46] Now, by assumption, a repair may take place: Agree-Copy applies directly, without sufficient evidence, but, it can be assumed, on the basis of what is known about all (relevant) D items in the numeration. Thus, Agree-Copy values the probe feature of the binding index feature of the reflexive (or reciprocal) with *some*

[45] In fact, an intrinsic order is derived if Agree-Link and Agree-Copy are assigned to two separate levels of representation, as in the original proposal. As will become clear momentarily, in the present context I will not make this assumption.

[46] The binding index of the D that is the head of the DP *dominating* the reflexive/reciprocal is not accessible because this would yield and i-over-i Filter violation; cf. Chomsky (1981).

index from a D in the numeration, without there being a prior PIC-respecting link established by Agree-Link; however, since Agree-Link has not yet applied, the probe status of the binding index feature will be preserved; cf. (51c).[47]

(51) a. *Initial representation*:
[$_{DP}$ D ... [$_{DP}$ REFL$_{[*\square*]}$]]
b. *Failure of Agree-Link*:
[$_{DP}$ D ... [$_{DP}$ REFL$_{[*\square*]}$]]
c. *Agree-Copy as a repair*:
[$_{DP}$ D ... [$_{DP}$ REFL$_{[*3*]}$]]

Of course, given that there is also a D item with the binding index 4 in the numeration, Agree-Copy could also have turned REFL$_{[*\square*]}$ into REFL$_{[*4*]}$; and similarly for all other indices of D items in the numeration. Thus, at this point, *derivational branching* takes place: In the continuation based on REFL$_{[*3*]}$, this valued probe must at some point find, via late Agree-Link, a c-commanding, locally accessible, minimality-respecting DP with a matching goal (where local accessibility is determined by lack of an intervener, and interveners can be defective, i.e., have a different binding index).[48] If it does, as in (52) (cf. (47a)), Agree-Link can finally take place, the probe status is removed from the index of the reflexive pronoun, and the output can be well formed.

[47]Like all instances of feature valuation, binding index valuation intrinsically violates the No Tampering Condition (Chomsky (2007, 2008, 2013); see footnote 15) and the Inclusiveness Condition (Chomsky (1995, 2001)); see Müller (2015). Thus, specific exceptions to these constraints must be envisaged if the constraints are to be adopted. The question arises whether a similar consequence also holds for the Strict Cycle Condition, such that a similar exception would have to be postulated here as well. This is not the case if it is assumed that Agree-Copy always relies on a source index available on some D that is not in an embedded position, as speculated in the main text (i.e., the Copy operation involves material that is either in the current root domain, or in the numeration (or workspace) of the derivation).

[48]This raises the question of what happens in cases of long-distance binding of bound-variable pronouns, given that these are also brought about by an Agree operation. In these cases, intervention does not seem to play a role; cf. (i).

(i) Every boy$_1$ thinks that Mary$_2$ will invite him$_1$ to the party

For present purposes, I will take this to instantiate an irreducible difference between reflexives (and reciprocals) on the one hand, and personal and possessive pronouns (in the languages under consideration) on the other: Agree-Link for the former is subject to minimality, Agree-Link for the latter is not.

(52) [$_{CP_0}$ Maria$_3$ erzählt [$_{CP_1}$ [$_{DP_2}$ welche Statue von sich$_{[3]}$]
Maria recounts which statue$_{acc}$ of REFL
Anna$_4$ t$_2$ gesehen hat]]
Anna$_{nom}$ seen has

If, on the other hand, REFL$_{[*3*]}$ does not find a c-commanding, locally accessible goal that can lead to Agree-Link, ungrammaticality results. This is the case if the DP containing the reflexive pronoun does not move (as in (48); here, *Anna*$_4$ will qualify as a defective intervener), or if it moves but never finds a locally accessible antecedent to establish an Agree-Link operation with.

As noted, an alternative choice of binding index in (51c) would have resulted in REFL$_{[*4*]}$; this valued probe will find a locally accessible antecedent immediately, i.e., without movement; if the DP that it is part of nevertheless moves in a subsequent step, a counter-bleeding effect arises; cf. (53) (which is identical to (52), except for the index chosen for *sich*).

(53) [$_{CP_0}$ Maria$_3$ erzählt [$_{CP_1}$ [$_{DP_2}$ welche Statue von sich$_{[4]}$]
Maria recounts which statue$_{acc}$ of REFL
Anna$_4$ t$_2$ gesehen hat]]
Anna$_{nom}$ seen has

All in all, it can be concluded that a strictly cyclic approach to data like those in (47) seems viable: For a DP-internal reflexive (or a reciprocal) that does not find a binder in this domain, a problem arises, which is repaired immediately (but tentatively) by chosing a value for the binding index feature without having sufficient evidence for it (i.e., by applying Agree-Copy without a prior Agree-Link); and the choice of an index then has consequences for the remainder of the derivation, leading either to successful binding (if a locally accessible DP with the same index is found at some step), or to a crash (if such a locally accessible DP is not found).

3.3. Movement and Resumption in German

As a second case of seemingly counter-cyclic repair, consider instances of resumption in German that show up with certain kinds of movement across islands. If movement of a (phonologically empty) relative operator takes place from a Complex Noun Phrase Condition (CNPC) island in German, landing in the specifier of a relative C item *wo* ('where'), a resumptive pronoun must

show up in the base position of the relative pronoun; cf. (54a). Similarly, under such relativization, a resumptive pronoun is obligatory if movement crosses an adjunct island; cf. (54b).[49]

(54) a. Das ist ein Buch [$_{CP}$ Op$_1$ [$_C$ wo] ich einen Mann getroffen
this is a book where I a man$_{acc}$ met
habe [$_{CP}$ der *t$_1$/es$_1$ gelesen hat]]
have who it read has

b. Das ist ein Buch [$_{CP}$ Op$_1$ [$_C$ wo] ich eingeschlafen bin [$_{CP}$
this is a book where I fallen asleep have
nachdem ich *t$_1$/es$_1$ gelesen habe]]
after I it read have

If no island is crossed in the course of movement, the use of the resumptive strategy is blocked; cf. (55a) (instantiating clause-bound movement) and (55b) (with movement from a restructuring infinitive).[50]

(55) a. Das ist ein Buch [$_{CP}$ Op$_1$ [$_C$ wo] ich t$_1$/*es$_1$ gelesen habe]
this is a book where I it$_{acc}$ read have

b. Das ist ein Buch [$_{CP}$ Op$_1$ [$_C$ wo] ich [$_{VP}$ t$_1$/*es$_1$ zu
this is a book where I it$_{acc}$ to
kaufen] versucht habe]
buy tried have

The confinement of resumptive pronouns to island contexts in German is indicative of a repair operation; these items clearly show up as a last resort. However, they occur in the base position of the movement operation, and the

[49] See Müller (2014, ch. 4) for arguments that there is indeed a moved empty relative operator involved in this construction, that *wo* is truly a complementizer in this environment, and that we are not dealing with intrusive (i.e., meta-grammatical) resumption here, but with proper, grammaticalized resumption (cf. Sells (1984) for the difference, and for tests to determine the status of a given occurrence of resumption as either intrusive or grammaticalized).

[50] Unlike what is the case with the resumptive stratgegy in the presence of islands in (54), which would seem to be fully acceptable and unmarked for most speakers, the co-occurrence of a zero relative operator and a complementizer *wo* in transparent contexts without resumptive pronouns belongs to substandard or regional varieties of German, and its use is often stigmatized. However, the contrast between the versions of the sentences in (55) without a resumptive, and those with a resumptive, is clear even for speakers who do not tolerate the former, stigmatized construction.

island that licenses their occurrence may show up much later, and much higher in the structure. Therefore, it looks like there is a severe problem from the perspective of the Cyclic Principle: At the point where the decision about the absence or presence of a resumptive pronoun must be taken, there is no island yet; and when the island finally comes into being, going back to the lower cyclic domain and realizing the base position of movement by a resumptive pronoun will violate the Strict Cycle Condition.[51]

Again, the question arises what a cyclic alternative could look like. The approach that is developed in Müller (2014) as an answer to this question might suggest itself from the present perspective because it relies on derivational branching. In what follows, I will sketch the outlines of this analysis (see Müller (2014, ch. 4) for a comprehensive account).

The core assumption is that for the first local intermediate movement step to Specv (a position that may or may not ultimately be deeply embedded within in island), the derivation can choose to either leave nothing behind, or leave a copy behind (which is then subsequently spelled out as a pronoun; see Pesetsky (1998)).[52] Next, the information about the creation of a copy is stored on a buffer of the moved item (more specifically, as the value of the movement-related feature of the item, e.g., [rel], for relativization). For concreteness, if an XP_1 has undergone the copying, the movement-related feature (e.g., [rel]) on the moved item is now accompanied by an edge feature [•1•]; so, if a copy has been split off from a category XP bearing index 1, both items bear index 1 as a consequence, and [•1•] also shows up on the moved XP. More generally, for any index *n*, the feature [•*n*•] signals that an XP_n copy has been split of from the moved XP, and that this operation is not costless: Something is now missing from the moved item (as indicated by • •), and this is the item itself (as indicated by the index 1). The two resulting configurations are shown in (56a) (regular movement, no copy) and (56b) (cyclic generation of a copy in the base position); in both these representations,

[51]Would this derivation also violate the Cyclic Principle? To some extent, this depends on the exact nature of the operation that introduces the resumptive pronoun. Assuming that it is the presence of the island that directly triggers resumption with relativization in German, the Cyclic Principle would not be violated by the counter-cyclic derivation sketched in the text. The phenomenon at hand would then qualify as yet another case where the Cyclic Principle and the Strict Cycle Condition do not make identical predictions, in addition to those discussed in section 2.

[52]This deviates from Chomsky (1995) and much subsequent work based on it, where it is assumed that all instances of movement leave copies.

the assumption is that it is an object that starts moving from the base position (i.e., the complement position of V).

(56) a. $[_{vP}\ XP_{1\,[rel]}\ [_{v'}\ \ldots\ [_{VP}\ -\ V\ ]\ v\ ]]$
b. $[_{vP}\ XP_{1\,[rel]:[\bullet 1 \bullet]}\ [_{v'}\ \ldots\ [_{VP}\ XP'_{1}\ V\ ]\ v\ ]]$

Subsequently, if the derivation proceeds on the basis of (56a), i.e., without the copy, and encounters an island at some point, ungrammaticality results; otherwise everything is fine. If, on the other hand, the derivation proceeds on the basis of (56b), i.e., with the copy, and does *not* encounter an island (which, of course, captures the normal state of affairs), ungrammaticaly results eventually because the information on the buffer leads to illformedness in a criterial position; however, if an island is in fact encountered, the incriminating information is deleted, and everything is fine. Why should all of this be the case?

The key to an answer lies in the adoption of the approach to islands developed in Müller (2011). In this approach, it is assumed that in order to satisfy the PIC (cf. (26)) by movement, an edge feature must be available that triggers an intermediate movement step to a specifier position of the phase. Such edge features are not intrinsically present; rather, they are inserted in accordance with the Edge Feature Condition. The Edge Feature Condition ensures that an edge feature can only be inserted on a phase head if this is the only way to produce a balanced phase (see page 23 above). However, an additional assumption made in that approach is that an edge feature can also only be inserted if the phase has some other active feature at this point that may trigger a syntactic operation (a structure-building feature, or a probe feature). Crucially, with typical XPs that are islands, XP is merged as the final operation driven by structure-building features of a phase head. Furthermore, a potential probe feature that the phase head might retain after combining with a last-merged specifier, and that would permit insertion of an edge feature on the phase head, is blocked by the Strict Cycle Condition (given that Agree requires c-command, and that every projection is a cyclic domain).

Consequently, a last-merged specifier of a phase is predicted to be an island: For an item to be extracted from a last-merged XP in a phase, an edge feature would need to be inserted on the phase head, but an edge feature cannot be inserted because the phase head has by now become inactive. Therefore, a fatal PIC violation will arise with extraction from last-merged specifiers of

phase heads; and assuming typical instances of islands (like, in the case at hand, adjuncts and CPs embedded by nouns) to always qualify as last-merged specifiers derives the illegitimacy of extraction from an island.

With this approach to islands as a background, it becomes clear why the presence of [•*n*•] on a moved item within an island makes it possible to circumvent the island: In effect, the moved item brings its own designtaed edge feature, which can be used to bring about an intermediate movement step to the specifier of the next-higher phase head, and thereby circumvent the island effect; cf. the derivation in (57) (based on (56b)):[53] XP_1 has its own designated edge feature resulting from the generation of a copy in the first movement step (cf. (57a)); this feature is used to license extraction from YP to a specifier in the edge domain of π (cf. (57b)); and finally, the phase πP can now be left, in accordance with the PIC (cf. (57c)).

(57) a. $[_{\pi P}\ [_{YP}\ XP_{1\,[rel]:[\bullet 1\bullet]}\ [_{\pi'}\ \pi$ ... XP'_1 ...]]]
b. $[_{\pi P}\ XP_{1\,[rel]}\ [_{\pi'}\ [_{YP}\ [_{\pi'}\ \pi$ ... XP'_1 ...]]]]
c. $XP_{1\,[rel]}$ ... $[_{\pi P}\ t_1\ [_{\pi'}\ [_{YP}\ [_{\pi'}\ \pi$ ... XP'_1 ...]]]]

In contrast, if no resumptive copy has been generated in the base position, the moved item is not provided with a means to circumvent the island effect incurred by the last-merged specifier that it is a part of; cf. the derivation in (58a) (based on (56a)): In (58a), XP_1 does not have an edge feature (there was no copy operation in the first movement step that would be needed for it to arise); therefore, extraction to Specπ is impossible (cf. (58b)); hence, the island YP remains strict at later steps because any extraction from YP (and πP) will now violate the PIC.

(58) a. $[_{\pi P}\ [_{YP}\ XP_{1\,[rel]}\ [_{\pi'}\ \pi$ ...]]]
b. *$[_{\pi P}\ XP_{1\,[rel]}\ [_{\pi'}\ [_{YP}\ t_1\ [_{\pi'}\ \pi$ ...]]]]
c. *$XP_{1\,[rel]}$ ... $[_{\pi P}\ [_{YP}\ t_1\ [_{\pi'}\ \pi$ ...]]]

Next, assuming that designated edge features resulting from early resumption can only be used if all else fails (i.e., if there is no other way to establish an edge feature on a phase head), they will eventually lead to illformedness if the moved item does not require it to permit extraction from a last-merged

[53] Here, π is a phase head, YP is a last-merged specifier in the phase, and XP_1 is the moved item that wants to leave the island YP, and that has made it to YP's edge domain, in accordance with the PIC (if YP itself is a phase).

specifier at some point of the derivation, as with other features that can trigger syntactic operations but fail to do this in a given derivation.[54]

Thus, in a nutshell, it follows that if a copy is made during cyclic, bottom-up structure-building, an island will have to be encountered at some later step, and if no copy is made, there must not be an island higher up in the tree. From a more general perspective, this way, island repair by resumption can be given an analysis that adheres to the Cyclic Principle and the Strict Cycle Condition.

3.4. Global Case Splits in Yurok

A third relevant instance of seemingly counter-cyclic repair involves global case splits. Usually, case splits in the world's languages are local, in the sense that a given type of argument (a subject or an object) may sometimes appear with case marking, and sometimes without case marking, depending on the degree to which it is "prototypical"; this is taken to be a purely syntactic phenomenon (presence vs. absence of case) in Aissen (2002) and much subsequent work.[55] Prototypicality is based on the position of an argument with a given grammatical function on the Hale/Silverstein hierarchies in (59) (cf. Hale (1972) and Silverstein (1986)).

(59) *Hale/Silverstein hierarchy*

a. Person scale: $1 \succ 2 \succ 3$
b. Animacy scale: human $\succ$ animate $\succ$ inanimate
c. Definiteness scale: pronoun $\succ$ proper name $\succ$ definite $\succ$ indefinite specific $\succ$ non-specific

Prototypical subjects are those that align with the areas of these hierarchies located to the left (ideally a subject is first or second person, human, and a pronoun), whereas prototypical objects align with the areas on the right (ideally an object is third person, inanimate, and indefinite non-specific). In

[54] At least, this is the case for German; parametrization with respect to this condition produces resumption that is not confined to island contexts, which is also established for many constructions in many languages.

[55] In contrast, in Keine and Müller (2015) we show that the phenomenon may ultimately often be morphological in nature since the relevant alternations do not always have to be between zero exponence and non-zero exponence; in some cases, the alternation is between two non-zero exponents, i.e., there is a choice among two (or more) allomorphs realizing one and the same case. These complications do not have to concern us in the present context, though.

the ideal, or close-to-ideal states of affairs, there is often no case marking; but deviations from an ideal state of affairs are often signalled as such, and give rise to differential subject marking and differential object marking (and the stronger the deviation, the more likely this case marking is). Differential subject and object marking can thus be viewed as repair operations. On this view, case is normally unmarked in the languages exhibiting case splits, but a case split occurs, leading to case-marking, if the argument is marked, i.e., non-prototypical.

This is illustrated for differential object marking in Hindi (cf. Mahajan (1990), Stiebels (2002), Butt and King (2004), and Keine (2007), among many others). A maximally typical (indefinite, non-human) object is not case-marked; cf. (60a); it is case-marked by *-ko* if it bears features that are unexpected for objects (like definite interpretation); cf., e.g. (60b).

(60) a. Nadya-ne gaṛi-∅ cɑla-yi
Nadya.F.SG-ERG car.F.SG-NOM drive-PERF.F.SG
hɛ
be.PRES.3SG
'Nadya has driven a car.'

b. Nadya-ne gaṛi-ko cɑla–ya
Nadya.F.SG-ERG car.F.SG-ACC drive–PERF.M.SG
hɛ
be.PRES.3SG
'Nadya has driven the car.'

Aissen (2003) has come up with an optimality-theoretic analysis that incorporates this insight: For objects, in the languages under consideration, there is a high-ranked constraint ensuring that the object is not case-marked; this is the normal state of affairs. However, when the object has atypical features (i.e., features corresponding to the left regions of the Hale/Silverstein hierarchies), an even higher-ranked constraint becomes active that successfully demands the presence of case on the object in atypical environments.

Against this background, we can ask whether there is a problem for cyclicity constraints posed by differential object marking as in (60). This is not the case because the phenomenon is strictly local. Thus, suppose that object case in (60) is assigned by v; therefore, the decision whether case is assigned or not must be made on the v′ cycle. At this stage, the properties of an object

DP within the VP are all accessible. Therefore, the decision can be made immediately (in accordance with the Cyclic Principle), and without a need to later go back (in accordance with the Strict Cycle Condition), resulting in Agree between v and the DP complement of V and thus bringing about differential case marking in (60b).

However, things are different with case splits that are not *local*, but *global*. Here the case-marking of one argument depends on properties of this argument with respect to one (or more) of the hierarchies in (59) *and* on properties of its co-argument; thus, the decision about differential case marking cannot be local but must be global (and this is why Silverstein (1986, 178-179) came up with the term "global case-marking"). Abstracting away from the similar phenomenon of direct vs. inverse marking in Algonquian, global case splits appear to be somewhat rare. A well-known example is the global split with object case marking in Yurok; cf. (61).

(61) a. keʔl [nek ki newoh-paʔ]
2.SG.NOM 1.SG.NOM FUT see-2>1SG
'You will see me.'

b. yoʔ [nek-*ac* ki newoh-peʔn]
3.SG.NOM 1.SG.OBJ FUT see-3SG>1SG
'He will see me.'

The split is determined by the person hierarchy, which is 1/2 $\succ$ 3, and it involves differential object marking: The internal argument of the verb bears accusative case if it is higher on the person hierarchy than the external argument, i.e., if both arguments are atypical. Global case splits as in (61) have been addressed by, i.a., Aissen (1999), de Hoop and Malchukov (2008), Béjar and Řezáč (2009), Keine (2010), Georgi (2012), Bárány (2017), and Bárány and Sheehan (2021). However, as noted by Georgi (2012), given a derivational, bottom-up approach, most analyses of the phenomenon emerge as counter-cyclic upon closer inspection: When the derivation has reached the v′ stage, with v the head that may assign accusative case to an object, the decision cannot yet be taken because the subject is not yet part of the structure; cf. (62a). Subsequently, the subject is merged in Specv; cf. (62b). And it is only at this point that the decision can be made, leading to accusative case assignment to the object by v if the subject is third person and the object is first person, as in (61b); cf. (62c). This final step is counter-cyclic; under

present assumptions, it violates the Strict Cycle Condition (and perhaps also the Cyclic Principle, depending on the exact formulation of the conditions for accusative case assignment).

(62) a. $[_{v'}$ v ... $DP_{[1]}$]
b. $[_{vP}$ $DP_{[3]}$ $[_{v'}$ v ... $DP_{[1]}$]]
c. $[_{vP}$ $DP_{[3]}$ $[_{v'}$ v ... $DP_{[1]}$-acc]]

From the perspective of cyclicity, the problem with a derivation of global case splits along the lines of (62) is that it is unclear how the head that assigns the case features to the external or internal argument can know about the remaining argument's properties before this latter argument is actually present. According to the derivational branching strategy, it does not, but a preliminary decision is taken nonetheless. The analysis developed in Georgi (2012) is exactly of this type. In what follows, I sketch a somewhat simple-minded reconstruction of the gist of Georgi's approach that focusses on the cyclicity and derivational branching issues and leaves out many intricacies (e.g., related to the nature of case and structure-building features, to the concept of maraudage that plays an important role in the analysis, and to the nature of the underlying optimization procedure).

Suppose first that there are two relevant constraints of the type proposed in Aissen (2003) that are active in the syntax of Yurok, viz., (63a) and (63b).[56] (63a) is a *violable* constraint; but (63b) is *inviolable* in well-formed outputs.

(63) a. A local (first or second) person object must be case-marked.
b. A local (first or second) person object must be case-marked if the subject is third person.

(63a) is the constraint that locally, within v′, triggers the repair; this repair must be tentative because what it really wants to preclude is a violation of (63b), which cannot yet be detected at this point. There is an economy constraint counter-acting (63a), which may thus block the repair. If the two constraints are tied, optionality of case-marking arises with first or second person objects; but there will never be case-marking of third person objects (there is no trigger,

[56] In Aissen's approach, these constraints are generated via harmonic alignment of prominence scales with grammatical functions (cf. Prince and Smolensky (2004)) and local conjunction with a markedness constraint requiring case on DPs (cf. Smolensky (2006)) to yield (63a), and via local conjunction of the resulting constraints to yield (63b).

and the operation is therefore always precluded by economy). However, and this is Georgi's (2012) core idea, if case-marking by v takes place, triggered by (63a), the v head is changed in such a way that it can only subsequently combine with a third person subject, and not with a first or second person subject. Thus, the subcategorization feature that v has for the subject is now something like [•$D_{[3]}$•], rather than [•D•].[57] Consequently, if a local person object is present and v assigns case to it, the subject that is merged subsequently can only be third person. Alternatively, if v has not assigned object case on the v′ cycle, its subcategorization feature for the external argument is not affected (it is still [•D•]), and it can freely combine with a first or second person subject. In principle, v can now also still be merged with a third person subject; but in this case, (63b) is violated, which (by assumption, since this constraint is classified as inviolable) leads to illformedness.

Thus, under a derivational branching approach, the initial, tentative repair (viz., case-marking of the object) survives in exactly the (inverse) environment where it is required; absence of repair prevails otherwise (with third person objects or first/second person subjects); and the account is fully compatible with the Strict Cycle Condition (and the Cyclic Principle).

3.5. Epenthesis in Icelandic

Finally, as a fourth case study I would like to briefly, and speculatively, extend the derivational branching approach to a cyclicity issue arising in morphology/phonology interactions. The background assumption is that morphology and phonology are cyclically interspersed (i.e., governed by the concept of *cyclicity*$_k$, in the terminology introduced in section 1). More specifically, the derivation starts with a morphological root domain, next applies phonological operations that belong to this domain, then adds morphological exponents, which establishes a new cyclic domain, then carries out phonological operations that apply in this cyclic domain, and so on, until the final morphological cycle has been reached, and the final phonological

[57] Georgi (2012) accounts for this by invoking a concept of feature *maraudage*: To accomodate additional case-marking of the object, features that are required to subcategorize for a local subject are used up on v. This presupposes that person features and case features or ontologically of the same type, at least in Yurok and other languages exhibiting global case splits. Alternatively, one might want to view this change of the subcategorization properties of v as a *weakening* of the *strength* of v. This presupposes an approach like Gradient Harmonic Grammar (cf. Smolensky and Goldrick (2016)), where strength is a primitive property of lexical items.

operations have applied. In approaches that envisage such an interaction of morphology and phonology (but not in strictly representational approaches like Standard Parallel Optimality Theory as devised in Prince and Smolensky (2004)), the Cyclic Principle in (9) is widely adopted.[58]

Based on Kiparsky (1985), Gleim (2022) discusses the case of vowel epenthesis in Icelandic, which takes place so as to break up a consonant cluster in the coda. As shown in (64), epenthesis inside roots is bled by word-level re-syllabification. This looks like a counter-cyclic interaction: Epenthesis takes place on the root cycle (cf. (64a)), but subsequent attachment of the definite determiner on a later cycle that adds inflectional exponents leads to counter-cyclic suppression of epenthesis on the root cycle (cf. (64b)).

(64) a. livr → livyr 'liver'
b. livr-in → livrin 'liver-DEF.FEM.NOM'

As Gleim notes, the phenomenon is complicated by the fact that gender plays a role. The stem *livr* in (64) is feminine, but things are different with masculine noun stems. As shown in (65), in this case seemingly counter-cyclic suppression of epenthesis does not show up when a definite article is added to the stem – epenthesis applies across the board.

(65) a. hamstr → hamstyr 'hamster'
b. hamstr-in → hamstyrin 'hamster-DEF.MASC.NOM'

Gleim's (2022) solution to this problem for cyclicity is as follows. First, the analysis envisages three cyclic domains beyond the root for words in Icelandic: (a) the stem level; (b) the word level; and (c) the phrase level. Second, there is evidence that epenthesis applies between the word level and the phrase level. Third, the definite article exponent does not belong to the

[58]The case is different with the Strict Cycle Condition. Based on the interaction of vowel deletion and Schwa epenthesis ('Sonorant Cluster') in Klamath, Kean (1974) argued that a version of the Strict Cycle Condition that looks exactly like the one in (13) is active in phonology, in addition to the Cyclic Principle. Following this, Mascaró (1976) and Kiparsky (1982*b*, 1985) advance formulations of the Strict Cycle Condition for phonology that are major deviations from the original Chomskyan concept, and designed to cover additional kinds of phenomena (in particular, derived environment effects are now accounted for by a stipulation to this effect that is part of a modified Strict Cycle Condition). Nowadays, the Strict Cycle Condition does not seem to be generally adopted anymore in phonology. See Gleim (2023) for extensive discussion.

stem level; it follows exponents that belong to the stem level. Fourth, the definite article exponent can neither uniformly be added early, at the word level (because then epenthesis would be blocked throughout, i.a., also in (65b), due to resyllabification and breaking up of the consonant cluster in the coda), nor uniformly be added late, at the phrase level (because then epenthesis would occur in all forms, including the one in (64b)). Fifth and finally, Gleim's conclusion is that the definite article exponent is merged *before* epenthesis with feminine (and neuter) nouns (i.e., at the word level), and *after* epenthesis with masculine nouns (i.e., at the phrase level).

This analysis seems to work well, and is in accordance with the Cyclic Principle. However, it may give rise to a potential problem: The definite article exponents that are added are identical with feminine and masculine nouns in nominative singular environments – a minor difference in orthography (that has been adjusted in the above examples) notwithstanding, it is the same *in* in (64) and (65); and the inflected forms of the article in other environments may differ, but they clearly share a common core. However, if one takes the hypothesis seriously that the definite article exponent attaches at the word level with feminine noun stems, and at the phrase level with masculine noun stems, the conclusion suggests itself that the analysis must envisage *two* separate definite article exponents in the mental lexicon of Icelandic speakers; this, in turn, means that a likely case of systematic syncretism remains unaccounted for.

For this reason, it might be worth pursuing the question of what a direct transfer of the derivational branching approaches presented for apparently counter-cyclic syntactic phenomena in the previous three subsections could look like in the case at hand. In what follows, I sketch a possible line of approach.

At an early stage of the derivation, i.e., before a definite article exponent is present, the feminine noun stem *livr* can choose to either carry out epenthesis or not, based on the outcome of an optimization procedure (with a counter-acting faithfulness constraint prohibiting epenthesis, and the two constraints tied). Importantly, epenthesis can be suppressed here (in the hope that this may ultimately pay off) even though the context for this operation to apply is present. This produces derivational branching. (In contrast, the masculine noun *hamstr* always carries out epenthesis; there is no optionality involved here). Thus, non-application of epenthesis with feminine *livr* is locally unmotivated, just like locally unmotivated index copying (without prior Agree-Link), locally

unmotivated resumption, and locally unmotivated case-marking in the earlier reanalyses of seemingly counter-cyclic phenomena in syntax.

So, at this point two continuations need to be considered. In the first one, epenthesis has not applied. If a definite article exponent is added, this leads to syllabification, and everything is fine; cf. (66a). If, on the other hand, no exponent is added and the form stays the same, ungrammaticality arises; cf. (66b). This can be modeled by assuming that the markedness constraint requiring epenthesis is stronger (i.e., higher-ranked) at the word (or phrase) level than it is at the root level. The constraint violated by (66b) can be assumed to be inviolable in an optimal output (e.g., by assuming that it outranks the constraint blocking the null parse).

(66) a. livr → livr → livr-in
b. *livr → livr → livr-Ø

Alternatively, epenthesis does take place with *livr* on the root cycle. Suppose that the special nature of this vowel (giving rise to derivational branching) is indicated by a diacritic: *livy*$^{+}$*r*. In this case, if there is no subsequent attachment of a definite article exponent, well-formedness can be derived; cf. (67b). However, if the article exponent is added, as in (67a), the diacritic on the epenthetic vowel ensures that a high-ranked (in effect, again, inviolable) constraint against unmotivated epenthetic vowels (i.e., vowels accompanied by $^{+}$) is violated in the final output, and the null parse wins again.

(67) a. *livr → livy^{+}r → livy^{+}r-in
b. livr → livy^{+}r → livy^{+}r-Ø

To sum up: Such an analysis would certainly not be entirely unproblematic because it would require an otherwise unjustified diacritic; but it would respect the Cyclic Principle (and the Strict Cycle Condition, if it exists in the phonological component), and it would be straightforwardly compatible with the (morphologically motivated) assumption that there is only one definite article exponent in Icelandic for masculine and feminine environments.

References

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2012): *Phases. An Essay on Cyclicity in Syntax*. Vol. 543 of *Linguistische Arbeiten*, De Gruyter, Berlin.

Abels, Klaus and Kristine Bentzen (2011): Are Movement Paths Punctuated or Uniform?. *In:* A. Alexiadou, T. Kiss and G. Müller, eds, *Local Modelling of Non-Local Dependencies in Syntax*. Linguistische Arbeiten, De Gruyter, Berlin.

Adger, David, Susana Béjar and David Harbour (2003): 'Directionality of Allomorphy: A reply to Carstairs-McCarthy', *Transactions of the Philological Society* **101**(1), 109–115.

Aissen, Judith (1999): 'Markedness and Subject Choice in Optimality Theory', *Natural Language and Linguistic Theory* **17**, 673–711.

Aissen, Judith (2002): Bidirectional Optimization and the Problem of Recoverability in Head Marking Languages. Ms., University of California, Santa Cruz.

Aissen, Judith (2003): 'Differential Object Marking: Iconicity vs. Economy', *Natural Language and Linguistic Theory* **21**, 435–483.

Andermann, Felicitas (2023): Cyclicity in Morphological Movement: The Case of Potawatomi Inverse Marking. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Arregi, Karlos and Andrew Nevins (2012): *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Springer, Heidelberg.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller and Philipp Weisser (2015): 'Ergatives Move Too Early. On an Instance of Opacity in Syntax', *Syntax* **18**, 343–387.

Baker, Mark (2008): *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.

Bárány, András (2017): *Person, Case, and Agreement: The Morphosyntax of Inverse Agreement and Global Case Splits*. Oxford University Press, Oxford.

Bárány, András and Jenneke van der Wal (2021): 'We Don't Agree (Only) Upward', *Linguistic Inquiry* pp. 1–21. Early Access Corrected Proof.

Bárány, András and Michelle Sheehan (2021): Challenges for Dependent Case. Ms., Universiät Bielefeld and Anglia Ruskin University.

Barbiers, Sjef (2002): Remnant Stranding and the Theory of Movement. In: *Dimensions of Movement*. Benjamins, Amsterdam, pp. 47–67.

Barnickel, Katja (2014): Opazität in AcI-Konstruktionen. Master's thesis, Universität Leipzig.

Barss, Andrew (1986): Chains and Anaphoric Dependence. Ph.d. thesis, MIT, Cambridge, Mass.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**, 35–73.

Belletti, Adriana and Luigi Rizzi (1988): 'Psych-Verbs and Θ-Theory', *Natural Language and Linguistic Theory* **6**, 291–352.

Bhatt, Rajesh and Stefan Keine (2019): Secondary Strong Crossover in Hindi and the Typology of Movement. *In:* M. Baird and J. Pesetsky, eds, *Proceedings of NELS*. Vol. 49, GSLA, Amherst, pp. 125–134.

Bjorkman, Bronwyn and Hedde Zeijlstra (2014): Upward Agree is Superior. Ms., University of Toronto and Universität Göttingen.

Bobaljik, Jonathan (2000): The Ins and Outs of Contextual Allomorphy. *In:* K. Grohmann and C. Struijke, eds, *University of Maryland Working Papers in Linguistics*. Vol. 10, University of Maryland, College Park, pp. 35–71.

Bobaljik, Jonathan and Samuel Brown (1997): 'Interarboreal Operations: Head Movement and the Extension Requirement', *Linguistic Inquiry* **28**, 345–356.

Börjesson, Kristin and Gereon Müller (2020): Long Distance Agreement and Locality: A Reprojection Approach. *In:* P. Smith, J. Mursell and K. Hartmann, eds, *Agree to Agree: Agreement in the Minimalist Programme*. Open Generative Syntax, Language Science Press, pp. 225–265.

Bošković, Željko (2007): 'Agree, Phases, and Intervention Effects', *Linguistic Analysis* **33**, 54–96.

Bošković, Željko and Howard Lasnik (1999): 'How Strict is the Cycle?', *Linguistic Inquiry* **20**, 691–703.

Branigan, Phil (2013): Cyclicity and the Approach the Probe Principle. Ms., Memorial University of Newfoundland.

Browning, Marguerite (1991): 'Bounding Conditions on Representations', *Linguistic Inquiry* **22**, 541–562.

Bruening, Benjamin (2014): 'Precede-and-Command Revisited', *Language* **90**, 342–388.

Bruening, Benjamin and Eman Al Khalaf (2019): 'No Argument-Adjunct Asymmetry in Reconstruction for Binding Condition C', *Journal of Linguistics* **55**, 247–276.

Büring, Daniel (2005): *Binding Theory*. Cambridge University Press, Cambridge.

Butt, Miriam and Tracy King (2004): 'The status of case', *Clause Structure in South Asian Languages* pp. 153–198.

Caballero, Gabriela and Alice Harris (2012): A Working Typology of Multiple Exponence. *In:* F. Kiefer et al., eds, *Current Issues in Morphological Theory: (Ir)Regularity, Analogy, and Frequency*. Benjamins, Amsterdam, pp. 163–188.

Caballero, Gabriela and Sharon Inkelas (2013): 'Word Construction: Tracing an Optimal Path Through the Lexicon', *Morphology* **23**, 103–143.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.

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 (1975): *The Logical Structure of Linguistic Theory*. Plenum Press, New York.

Chomsky, Noam (1977): On Wh-Movement. *In:* P. Culicover, T. Wasow and A. Akmajian, eds, *Formal Syntax*. Academic Press, New York, pp. 71–132.

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

Chomsky, Noam (1986): *Barriers*. MIT Press, Cambridge, Mass.

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

Chomsky, Noam (2001): Derivation by Phase. *In:* M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2004): Beyond Explanatory Adequacy. *In:* A. Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures*. Vol. 3, Oxford University Press, Oxford, pp. 104–131.

Chomsky, Noam (2007): Approaching UG from Below. *In:* U. Sauerland and H.-M. Gärtner, eds, *Interfaces + Recursion = Language?*. Mouton de Gruyter, Berlin, pp. 1–31.

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, C. Otero and M. L. Zubizarreta, eds, *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.

Chomsky, Noam (2013): 'Problems of Projection', *Lingua* **130**, 33–49.

Chomsky, Noam (2019): Lectures. UCLA.

Chomsky, Noam, Ángel Gallego and Dennis Ott (2019): 'Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges', *Catalan Journal of Linguistics* (Special Issue), 229–261.

Chomsky, Noam and Morris Halle (1968): *The Sound Pattern of English*. MIT Press, Cambridge, Mass.

Chung, Sandra (1994): 'Wh-Agreement and 'Referentiality' in Chamorro', *Linguistic Inquiry* **25**, 1–44.

Chung, Sandra (1998): *The Design of Agreement: Evidence from Chamorro*. Chicago University Press, Chicago.

Cinque, Guglielmo (1990): *Types of A-bar Dependencies*. MIT Press, Cambridge, Mass.

Clements, George, James McCloskey, Joan Maling and Annie Zaenen (1983): 'String-Vacuous Rule Application', *Linguistic Inquiry* **14**, 1–17.

Cole, Peter and Gabriella Hermon (2000): Partial Wh-Movement: Evidence from Malay. *In:* U. Lutz, G. Müller and A. von Stechow, eds, *Wh-Scope Marking*. Benjamins, Amsterdam, pp. 101–130.

Collins, Chris (1993): Topics in Ewe Syntax. PhD thesis, MIT, Cambridge, Mass.

Collins, Chris (1994): 'Economy of Derivation and the Generalized Proper Binding Condition', *Linguistic Inquiry* **25**, 45–61.

Collins, Chris (1997): *Local Economy*. MIT Press, Cambridge, Mass.

Collins, Chris and Edward Stabler (2016): 'A Formalization of Minimalist Syntax', *Syntax* **19**(1), 43–78.

de Hoop, Helen and Andrej Malchukov (2008): 'Case-Marking Strategies', *Linguistic Inquiry* **39**(4), 565–587.

Doliana, Aaron (2013): The Super-Strong Person-Case Constraint: Scarcity of Resources by Scale-Driven Impoverishment. In: *Rule Interaction in Grammar*. Vol. 90 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig, pp. 177–202.

Embick, David (1998): Voice Systems and the Syntax/Morphology Interface. *In:* H. Harley, ed., *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*. Vol. 32, MITWPL, Cambridge, Mass., pp. 41–72.

Embick, David (2010): *Localism vs. Globalism in Morphology and Phonology*. MIT Press, Cambridge, Mass.

Embick, David and Rolf Noyer (2001): 'Movement Operations after Syntax', *Linguistic Inquiry* **32**, 555–595.

Epstein, Sam, Erich Groat, Ruriko Kawashima and Hisatsugu Kitahara (1998): *A Derivational Approach to Syntactic Relations*. Oxford University Press, Oxford and New York.

Fanselow, Gisbert (1983): Zur Behandlung der Anaphora in der Generativen Grammatik. Ein Überblick. Technical Report 84, Sonderforschungsbereich 99, Universität Konstanz.

Fanselow, Gisbert (1990): On the Nature of Proper Government and Syntactic Barriers. *In:* W. Abraham, W. Kosmeijer and E. Reuland, eds, *Issues in Germanic Syntax*. Mouton de Gruyter, Berlin, pp. 33–48.

Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.

Fanselow, Gisbert (2003): Münchhausen-Style Head Movement and the Analysis of Verb-Second. *In:* A. Mahajan, ed., *Head Movement and Syntactic Theory*. Vol. 3 of *Syntax at Sunset*, UCLA & Universität Potsdam Working Papers in Linguistics, Los Angeles & Potsdam, pp. 40–76.

Fanselow, Gisbert and Damir Ćavar (2001): Remarks on the Economy of Pronunciation. *In:* G. Müller and W. Sternefeld, eds, *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 107–150.

Featherston, Sam and Wolfgang Sternefeld (2003): The Interaction of Factors in Judgements of Reflexive Structures: Data from Object Coreference in German. *In:* L. Gunkel, G. Müller and G. Zifonun, eds, *Arbeiten zur Reflexivierung*. Niemeyer, Tübingen, pp. 25–50.

Ferguson, Scott and Erich Groat (1994): Defining 'Shortest Move'. Ms., Harvard University.

Fischer, Silke (2004): Towards an Optimal Theory of Reflexivization. PhD thesis, Universität Tübingen.

Fox, Danny (2000): *Economy and Semantic Interpretation*. MIT Press, Cambridge, Mass.

Frampton, John and Sam Gutmann (2002): Crash-Proof Syntax. *In:* S. D. Epstein and T. D. Seely, eds, *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 90–105.

Freidin, Robert (1992): *Foundations of Generative Syntax*. MIT Press, Cambridge, Mass.

Freidin, Robert (1994): Generative Grammar: Principles and Parameters Framework. *In:* R. Asher and J. Simpson, eds, *The Encyclopedia of Language and Linguistics, vol. III*. Pergamon Press, Oxford, pp. 1370–1385.

Freidin, Robert (1999): Cyclicity and Minimalism. *In:* S. D. Epstein and N. Hornstein, eds, *Working Minimalism*. MIT Press, Cambridge, Mass, pp. 95–126.

Frey, Werner (1993): *Syntaktische Bedingungen für die Interpretation*. Akademieverlag, Berlin.

Fritzsche, Rosa (2020): Ordering Discontinuous Φ-Feature Agree: Verbal -s in North Eastern English. Ms., Universität Leipzig. To appear in *Journal of Comparative Germanic Linguistics*.

Fritzsche, Rosa (2023): Anti-Local Agree and Cyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Gallego, Ángel (2020): Strong and Weak 'Strict Cyclicity' in Phase Theory. Ms., Universitat Autònoma de Barcelona.

Georgi, Doreen (2012): A Local Derivation of Global Case Splits. *In:* A. Alexiadou, T. Kiss and G. Müller, eds, *Local Modelling of Non-Local Dependencies in Syntax*. Linguistische Arbeiten, De Gruyter, Berlin, pp. 306–336.

Georgi, Doreen (2014): Opaque Interactions of Merge and Agree. PhD thesis, Universität Leipzig.

Georgi, Doreen (2017): 'Patterns of Movement Reflexes as the Result of the Order of Merge and Agree', *Linguistic Inquiry* **48**(4), 585–626.

Georgi, Doreen and Gereon Müller (2010): 'Noun Phrase Structure by Reprojection', *Syntax* **13**(1), 1–36.

Georgi, Doreen, Martin Salzmann and Marta Wierzba (2019): Condition A Reconstruction in German A′-Movement. An Empirical Investigation. In: *Proceedings of NELS 49*. GLSA, Amherst.

Gleim, Daniel (2022): Cyclic Locality in Phonology. Ms., Universität Leipzig.

Gleim, Daniel (2023): The Empirical Scope of the Strict Cycle Condition in Phonology. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Gleim, Daniel, Gereon Müller, Mariia Privizentseva and Sören Tebay (2022): 'Re-

flexes of Exponent Movement in Inflectional Morphology: A Study in Harmonic Serialism', *Natural Language and Linguistic Theory* **41**, 103–158.

Gong, Zhiyu Mia (2022): Case in Wholesale Late Merger: Evidence from Mongolian Scrambling. Ms., University of California, Santa Cruz.

Graf, Thomas (2013): Local and Transderivational Constraints in Syntax and Semantics. PhD thesis, UCLA.

Grewendorf, Günther (1983): 'Reflexivierung in deutschen AcI-Konstruktionen – Kein transformationsgrammatisches Dilemma mehr', *Groninger Arbeiten zur Germanistischen Linguistik* **23**, 120–196.

Grewendorf, Günther (1988): *Aspekte der deutschen Syntax*. Narr.

Grewendorf, Günther (1989): *Ergativity in German*. Foris, Dordrecht.

Grimshaw, Jane (1997): 'Projection, Heads, and Optimality', *Linguistic Inquiry* **28**, 373–422.

Grofulović, Jelena and Gereon Müller (2023): Cyclicity and Extended Exponence. Ms., Universität Leipzig.

Gunkel, Lutz (2003): *Infinitheit, Passiv und Kausativkonstruktionen im Deutschen*. Stauffenburg Verlag, Tübingen.

Hale, Ken (1972): A New Perspective on American Indian Linguistics. *In:* A. Ortiz, ed., *New Perspectives on the Pueblos*. University of New Mexico Press, Albuquerque, pp. 87–103.

Halle, Morris and Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. *In:* K. Hale and S. J. Keyser, eds, *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.

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

Heck, Fabian (2022): Syntactic Repair and Cyclic Optimization. Ms., Universität Leipzig.

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

Heck, Fabian and Gereon Müller (2003): 'Derivational Optimization of Wh-Movement', *Linguistic Analysis* **33**, 97–148.

Heck, Fabian and Gereon Müller (2007): Extremely Local Optimization. Proceedings of WECOL 2006. California State University, Fresno.

Heck, Fabian and Gereon Müller (2013): Extremely Local Optimization. *In:* H. Broekhuis and R. Vogel, eds, *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 135–166.

Heim, Irene (1989): Survey of Formal Semantics. Ms., MIT, Cambridge, Mass.

Heim, Irene and Angelika Kratzer (1998): *Semantics in Generative Grammar*. Blackwell, Oxford.

Hicks, Glyn (2009): *The Derivation of Anaphoric Relations*. Benjamins, Amsterdam.

Himmelreich, Anke (2017): Case Matching Effects in Free Relatives and Parasitic Gaps – A Study on the Properties of Agree. PhD thesis, Universität Leipzig.

Huang, Cheng-Teh James (1982): Logical Relations in Chinese and the Theory of Grammar. PhD thesis, MIT, Cambridge, Mass.

Huang, Cheng-Teh James (1993): 'Reconstruction and the Structure of VP', *Linguistic Inquiry* **24**, 103–138.

Huber, Walter and Werner Kummer (1974): *Transformationelle Syntax des Deutschen I*. UTB: Fink, München.

Jacobson, Pauline and Paul Neubauer (1974): Extraposition Rules and the Cycle. In: *Berkeley Studies in Syntax and Semantics*. Vol. 1:viii, pp. 1–99.

Kalin, Laura and Philipp Weisser (2021): Minimalism, Morphology, and the Lexicon: Then and Now. Ms., Princeton University and Universität Leipzig.

Katamba, Francis (1993): *Morphology*. Palgrave, New York and London.

Kean, Mary-Louise (1974): 'The Strict Cycle in Phonology', *Linguistic Inquiry* **5**, 179–203.

Keine, Stefan (2007): Reanalysing Hindi Split Ergativity as a Morphological Phenomenon. *In:* J. Trommer and A. Opitz, eds, *1 2 Many*. Vol. 85 of *Linguistische Arbeitsberichte*, Universität Leipzig, pp. 73–127.

Keine, Stefan (2010): *Case and Agreement from Fringe to Core. Impoverishment Effects on Agree*. Linguistische Arbeiten, Mouton de Gruyter, Berlin.

Keine, Stefan (2016): Probes and Their Horizons. PhD thesis, University of Massachusetts, Amherst.

Keine, Stefan and Gereon Müller (2015): Differential Argument Encoding by Impoverishment. *In:* I. Bornkessel-Schlesewsky, A. Malchukov and M. Richards, eds, *Scales and Hierarchies*. Trends in Linguistics, Mouton de Gruyter, Berlin, pp. 75–130.

Kiparsky, Paul (1982*a*): From Cyclic Phonology to Lexical Phonology. *In:* H. van der Hulst and N. Smith, eds, *The Structure of Phonological Representations*. Vol. 1, Foris, Dordrecht, pp. 131–175.

Kiparsky, Paul (1982*b*): Lexical Morphology and Phonology. *In:* I.-S. Yang, ed., *Linguistics in the Morning Calm*. Hanshin Publishing Company, Seoul, pp. 3–91.

Kiparsky, Paul (1985): 'Some Consequences of Lexical Phonology', *Phonology Yearbook* **2**, 85–138.

Kiss, Tibor (2001): Anaphora and Exemptness. A Comparative Treatment of Anaphoric Binding in German and English. *In:* D. Flickinger and A. Kathol, eds, *Proceedings of the 7th International Conference on HPSG*. CSLI Publications, Stanford University, pp. 182–197.

Kitahara, Hisatsugu (1997): *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, Mass.

Kobele, Gregory (2023): Cyclicity in Minimalist Syntax. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Korsah, Sampson and Andrew Murphy (2020): 'Tonal Reflexes of Movement in Asante Twi', *Natural Language and Linguistic Theory* **38**, 827–885.

Kuno, Susumu (1972): 'Pronominalization, Reflexivization, and Direct Discourse', *Linguistic Inquiry* **3**, 161–195.

Kuno, Susumu (1987): *Functional Syntax*. Chicago University Press, Chicago.

Lahne, Antje (2009): Where There is Fire There is Smoke. Local Modelling of Successive-Cyclic Movement. PhD thesis, Universität Leipzig.

Lakoff, George (1970): 'Global Rules', *Language* **46**, 627–639.

Lasnik, Howard (1976): 'Remarks on Coreference', *Linguistic Analysis* **2**.

Lebeaux, David (1988): Language Acquisition and the Form of the Grammar. PhD thesis, University of Massachusetts, Amherst.

Lohndal, Terje (2010): 'Freezing Effects and Objects', *Journal of Linguistics* **47**, 163–199.

Longenbaugh, Nicholas (2019): On Expletives and the Agreement-Movement Correlation. PhD thesis, MIT, Cambridge, Mass.

Mahajan, Anoop (1990): The A/A-bar Distinction and Movement Theory. PhD thesis, MIT, Cambridge, Mass.

Mascaró, Joan (1976): Catalan Phonology and the Phonological Cycle. PhD thesis, MIT, Cambridge, Mass.

Matushansky, Ora (2005): 'Going Through a Phase', *MITWPL* **49**, 157–181. Perspectives on Phases.

Matushansky, Ora (2006): 'Head-Movement in Linguistic Theory', *Linguistic Inquiry* **37**, 69–109.

McCawley, James (1984): Exploitation of the Cyclic Principle as a Research Strategy in Syntax. *In:* W. de Geest and Y. Putseys, eds, *Sentential Complementation*. Foris, Dordrecht, pp. 165–183.

McCawley, James (1998): *The Syntactic Phenomena of English*. University of Chicago Press, Chicago.

McCloskey, James (1979): *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish*. Reidel, Dordrecht.

McCloskey, James (2002): Resumptives, Successive Cyclicity, and the Locality of Operations. *In:* S. D. Epstein and T. D. Seely, eds, *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 184–226.

Müller, Gereon (1995): *A-bar Syntax*. Mouton de Gruyter, Berlin.

Müller, Gereon (1999*a*): 'Imperfect Checking', *The Linguistic Review* **16**, 359–404.

Müller, Gereon (1999*b*): 'Optimality, Markedness, and Word Order in German', *Linguistics* **37**, 777–818.

Müller, Gereon (2001): Order Preservation, Parallel Movement, and the Emergence of the Unmarked. *In:* G. Legendre, J. Grimshaw and S. Vikner, eds, *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 279–313.

Müller, Gereon (2007): Extended Exponence by Enrichment. Argument Encoding in German, Archi, and Timucua. *In:* T. Scheffler, J. Tauberer, A. Eilam and L. Mayol, eds, *Proceedings of the 30th Annual Penn Linguistics Colloquium*. Vol. 13.1 of *Penn Working Papers in Linguistics*, University of Pennsylvania, Philadelphia, pp. 253–266.

Müller, Gereon (2011): *Constraints on Displacement. A Phase-Based Approach*. Vol. 7 of *Language Faculty and Beyond*, Benjamins, Amsterdam.

Müller, Gereon (2014): *Syntactic Buffers*. Vol. 91, Linguistische Arbeits Berichte, Universität Leipzig.

Müller, Gereon (2015): Optimality-Theoretic Syntax. *In:* T. Kiss and A. Alexiadou, eds, *Syntax. An International Handbook*. Vol. 2, De Gruyter, Berlin, pp. 875–936.

Müller, Gereon (2020): *Inflectional Morphology in Harmonic Serialism*. Advances in Optimality Theory, Equinox, Sheffield. http://home.uni-leipzig.de/muellerg/harm/imhs.pdf.

Müller, Gereon (2021): Constraints on Grammatical Dependencies. *In:* N. Allot, T. Lohndal and G. Rey, eds, *A Companion to Chomsky*. Wiley Blackwell, pp. 190–209.

Müller, Gereon (2022): *Structure Removal in German Syntax*. Book ms., Universität Leipzig.

Müller, Gereon and Wolfgang Sternefeld (2001): The Rise of Competition in Syntax: A Synopsis. *In:* G. Müller and W. Sternefeld, eds, *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 1–68.

Murphy, Andrew and Zorica Puškar (2018): 'Closest Conjunct Agreement is an Illusion', *Natural Language and Linguistic Theory* **36**, 1207–1261.

Murugesan, Gurujegan (2019): Predicting the Anaphor Agreement Effect and Its Violations. PhD thesis, Universität Leipzig.

Murugesan, Gurujegan (2022): 'Deriving the Anaphor-Agreement Effects and the Violations of It', *Syntax* **25**, 39–83.

Newman, Elise (2021): The (In)Distinction between Wh-Movement and C-Selection. PhD thesis, MIT, Cambridge, Mass.

Nissenbaum, Jon (2000): Investigations of Covert Phrase Movement. PhD thesis, MIT, Cambridge, Mass.

Noyer, Rolf (1997): *Features, Positions and Affixes in Autonomous Morphological Structure*. Garland Publishing, New York.

Perlmutter, David and Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.

Pesetsky, David (1979): Russian Morphology and Lexical Theory. Ms., MIT.

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

Pesetsky, David (1998): Some Optimality Principles of Sentence Pronunciation. *In:* P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky, eds, *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 337–383.

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, Mass., pp. 355–426.

Pesetsky, David and Esther Torrego (2006): Probes, Goals and Syntactic Categories. Ms., MIT.

Preminger, Omer (2014): *Agreement and Its Failures*. MIT Press, Cambridge, Mass.

Preminger, Omer (2018): How to Tell a Syntactic Phenomenon When You See It. Ms., University of Maryland.

Prince, Alan and Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell, Oxford.

Privizentseva, Mariia (2022*a*): Inverse Case Attraction in Moksha Mordvin Relative Clauses and the Nature of Merge. PhD thesis, Universität Leipzig.

Privizentseva, Mariia (2022*b*): Semantic Agreement in Russian: Gender, Declension, and Morphological Ineffability. Ms., Universität Leipzig. To appear in *Natural Language and Linguistic Theory*.

Privizentseva, Mariia (2023): Late Merge and Cyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Pullum, Geoffrey (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York.

Pullum, Geoffrey (1992): The Origins of the Cyclic Principle. In: *Proceedings of the Regional Meeting of the Chicago Linguistic Society*. Vol. 28.2, pp. 209–236.

Reinhart, Tanya (1976): The Syntactic Domain of Anaphora. PhD thesis, MIT, Cambridge, Massachusetts.

Reinhart, Tanya (1983): *Anaphora and Semantic Interpretation*. Croom Helm, London.

Reis, Marga (1976): 'Reflexivierung in deutschen A.c.I.-Konstruktionen. Ein transformationsgrammatisches Dilemma', *Papiere zur Linguistik* **9**, 5–82.

Reuland, Eric (2001): 'Primitives of Binding', *Linguistic Inquiry* **32**, 439–492.

Reuland, Eric (2011): *Anaphora and Language Design*. MIT Press, Cambridge, Mass.

Richards, Marc (2007): 'On Feature Inheritance', *Linguistic Inquiry* **38**, 563–572.

Richards, Norvin (1999): Featural Cyclicity and the Ordering of Multiple Specifiers. *In:* S. D. Epstein and N. Hornstein, eds, *Working Minimalism*. MIT Press, Cambridge, Mass, pp. 126–158.

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

Riemsdijk, Henk van and Edwin Williams (1981): 'NP-Structure', *The Linguistic Review* **1**, 171–217.

Rizzi, Luigi (1982): *Issues in Italian Syntax*. Foris, Dordrecht.

Rizzi, Luigi (1990): 'On the Anaphor Agreement Effect', *Rivista di Linguistica* **2**, 27–42.

Roberts, Ian (1997): *Comparative Syntax*. Arnold, London.

Roberts, Ian (2021): One Cycle to Rule Them All. Ms., University of Cambridge.

Ross, John (2012): A Preliminary, but Fattened, List of Transformations. Ms., University of North Texas.

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

Schneider-Zioga, Patricia (2005): Wh-Agreement as an Intervention Effect. Ms., University of California, Irvine.

Schwarzer, Marie-Luise (2022): Determiner Sharing in German. An Argument for Movement-Based Approaches to Ellipsis. PhD thesis, Universität Leipzig.

Sells, Peter (1984): Syntax and Semantics of Resumptive Pronouns. PhD thesis, University of Massachusetts, Amherst.

Šereikaitė, Milena (2021): 'Active Existential in Lithuanian: Remarks on Burzio's Generalization', *Linguistic Inquiry* **52**, 747–789.

Silverstein, Michael (1986): Hierarchy of Features and Ergativity. *In:* H. v. Riemsdijk and P. Muysken, eds, *Features and Projections*. Foris, Dordrecht, pp. 163–232.

Smolensky, Paul (2006): Harmonic Completeness, Local Constraint Conjunction, and Feature Domain Markedness. *In:* P. Smolensky and G. Legendre, eds, *The Harmonic Mind*. Vol. II, MIT Press, Cambridge, Mass., chapter 14, pp. 27–160.

Smolensky, Paul and Matthew Goldrick (2016): Gradient Symbolic Representations in Grammar: The Case of French Liaison. ROA 1286.

Speas, Margaret (1990): *Phrase Structure in Natural Language*. Kluwer, Dordrecht.

Stabler, Edward (2013): 'Two Models of Minimalist, Incremental Syntactic Analysis', *Topics in Cognitive Science* **5**, 611–633.

Stiebels, Barbara (2002): *Typologie des Argumentlinkings: Ökonomie und Expressivität*. Akademie Verlag, Berlin.

Stiebels, Barbara (2015): Multiple Exponence: Typology, Triggering Factors, and Theoretical Modelling. Ms., Universität Leipzig.

Streffer, Helene (2023): Incorporation as a Repair Mechanism and Cyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Stump, Gregory (2001): *Inflectional Morphology*. Cambridge University Press, Cambridge.

Svenonius, Peter (2004): On the Edge. *In:* D. Adger, C. de Cat and G. Tsoulas, eds, *Peripheries. Syntactic Edges and their Effects*. Kluwer, Dordrecht, pp. 261–287.

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

Trommer, Jochen (1999): Morphology Consuming Syntax' Resources. In: *Proceedings of the ESSLI Workshop on Resource Logics and Minimalist Grammars*. University of Nijmegen, pp. 37–55.

Trommer, Jochen (2020): Overall Description of the Research Unit and the Coordination Proposal. Ms., Universität Leipzig. (Grant proposal for a DFG Research Unit).

Uriagereka, Juan (1999): Multiple Spell-Out. *In:* S. Epstein and N. Hornstein, eds, *Working Minimalism*. MIT Press, Cambridge, Mass., pp. 251–282.

van Urk, Coppe (2015): A Uniform Syntax for Phrasal Movement: A Dinka Bor Case Study. PhD thesis, MIT, Cambridge, Mass.

Vikner, Sten (1989): 'Object Shift and Double Objects in Danish', *Working Papers in Scandinavian Syntax* **44**, 119–141.

Wasow, Tom (1972): Anaphoric Relations in English. PhD thesis, MIT, Cambridge, Mass.

Watanabe, Akira (1995): Conceptual Basis of Cyclicity. *In:* R. Pensalfini and H. Ura, eds, *Papers on Minimalist Syntax*. Vol. 27, MITWPL, Cambridge, Mass., pp. 261–291.

Webelhuth, Gert (1992): *Principles and Parameters of Syntactic Saturation*. Oxford University Press, Oxford.

Wexler, Ken and Peter Culicover (1980): *Formal Principles of Language Acquisition*. MIT Press, Cambridge, Mass.

Wunderlich, Dieter (1997): A Minimalist Model of Inflectional Morphology. *In:* C. Wilder, H.-M. Gärtner and M. Bierwisch, eds, *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 267–298.

Zyman, Erik (2018): On the Driving Force for Syntactic Movement. PhD thesis, University of California, Santa Cruz.