# LINGUISTISCHE ARBEITS BERICHTE

---

# Cyclicity

Edited by

Mariia Privizentseva, Felicitas Andermann & Gereon Müller

95

Institut für Linguistik
Universität Leipzig

**Herausgeber**: Institut für Linguistik
Universität Leipzig
Beethovenstr. 15
D-04107 Leipzig
www.uni-leipzig.de/~asw

**ISSN 0138-4597**

**Druck**: MERKUR DRUCK
Salomonstr. 20
D-04103 Leipzig

**Vertrieb**: Sabine Tatzelt
Institut für Linguistik
Universität Leipzig
Beethovenstr. 15
D-04107 Leipzig

Tel.: +341 97 37 610
Fax.: +341 97 37 609
E-mail: infl@uni-leipzig.de

# Contents

## Part I: Conceptualizing Cyclicity

# Preface

The seeds of this volume were planted at the first retreat of DFG Research Unit *Cyclic Optimization* (FOR 5175) that took place on May 12-13, 2022 at Wilhelm Ostwald Park in Großbothen; in its present form, it originates in the subsequent workshop *(Strict) Cyclicity* at Leipzig University on June 15, 2022. The goals of this workshop were to identify and compare different conceptions of cyclicity in phonology, morphology, and syntax, to discuss known evidence against cyclicity, and to talk about the role cyclicity plays in individual analyses of linguistic phenomena. These are also the goals of the current volume.

The volume has three parts. Part I, *Conceptualizing Cyclicity*, considers existing approaches to cyclicity in grammar, their empirical coverage and necessity. Part II, *Exploring Evidence*, shows new analyses of non-trivial linguistic effects enabled by cyclicity, but also re-analyses of some well-known patterns that eliminate cyclicity. Finally, part III, *Removing Obstacles*, investigates counter-cyclic analyses and patterns and demonstrates how they can be reconciled with cyclicity.

# Part I
# Conceptualizing Cyclicity

# 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 $\mathrm{T}_1$, $\mathrm{T}_2$ 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*$_k$, $\mathrm{T}_1$ involves *structure-building* whereas $\mathrm{T}_2$ 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, $\mathrm{T}_1$ is the system of lexical morphology, and $\mathrm{T}_2$ 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 $\mathrm{T}_1$ involves applications of Merge in a certain local domain (e.g., the *phase*; cf. Chomsky (2001)), and $\mathrm{T}_2$ 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 $\mathrm{T}_1$ and $\mathrm{T}_2$ 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 $\alpha$ ... [$_{\beta}$ ... [$_{\gamma}$ ... $\delta$ ... ] ... ] ..., movement of $\delta$ to $\alpha$ cannot apply if $\beta$ and $\gamma$ 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. [$_{\text{DP}_1}$ Which book ] do [$_{\text{TP}_2}$ you think [$_{\text{CP}}$ $t_1'$ [$_{\text{C}}$ (that) ] [$_{\text{TP}_4}$ John read $t_1$ ]]] ?
b. ?*[$_{\text{DP}_1}$ Which book ] do [$_{\text{TP}_2}$ you wonder [$_{\text{CP}}$ [$_{\text{PP}_3}$ to whom ] C [$_{\text{TP}_4}$ 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) $[_{\mathrm{CP}_0} \ldots [_{\mathrm{CP}_1} \mathrm{XP}(\leftarrow O_1)\ [_{\mathrm{C}'}\ \mathrm{C}(\leftarrow O_2)\ [_{\mathrm{TP}}\ \mathrm{YP}(\leftarrow O_3)\ [_{\mathrm{T}'}\ \mathrm{T}(\leftarrow O_4)\ [_{\mathrm{vP}}$
$\mathrm{WP}(\leftarrow O_5)\ [_{\mathrm{v}'}\ \mathrm{v}(\leftarrow O_6)\ [_{\mathrm{VP}}\ \mathrm{ZP}(\leftarrow O_7)\ [_{\mathrm{V}'}\ \mathrm{V}(\leftarrow O_8)\ \mathrm{CP}_2\ ]]]]]]]]$

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 $\mathrm{D}_{x-1}$ 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) ?*[$_{\mathrm{DP}_1}$ Which book ] do [$_{\mathrm{TP}_2}$ you wonder [$_{\mathrm{CP}}$ [$_{\mathrm{PP}_3}$ to whom ] C [$_{\mathrm{TP}_4}$ 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 $C_{[\bullet\mathrm{wh}\bullet]}$ 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*:

[$_{\mathrm{DP}_2}$ Which claim [$_{\mathrm{CP}}$ 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) [$_{\mathrm{DP}_2}$ Every [$_{\mathrm{NP}}$ argument [$_{\mathrm{CP}}$ 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 $\alpha$ really means that Merge of $\alpha$ 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. $[_{\text{CP}_1}$ $[_{\text{DP}_1}$ Some students ] try $[_{\text{CP}_2}$ ~~$[_{\text{DP}_1}$ some students ]~~ to be in the lecture hall ]]

b. $[_{\text{CP}_1}$ 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) *[$_{\mathrm{CP}_1}$ [$_{\mathrm{DP}}$ Some students ] try [$_{\mathrm{CP}_2}$ there to be in the lecture hall ]]

Importantly, a derivation giving rise to (25) is not excluded by the Cyclic Principle: On the initial $\mathrm{CP}_2$ 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 $\mathrm{CP}_1$ 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 $\mathrm{CP}_1$ 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 $\mathrm{CP}_1$ has been affected; but subsequent application of There-Insertion exclusively affects the embedded cyclic domain $\mathrm{CP}_2$, 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. [$_{\text{CP}}$ [$_{\text{C}}$ that ] John wanted what$_1$ ]
b. Mary say [$_{\text{CP}}$ [$_{\text{C}}$ that ] John wanted what$_1$ ]
c. [$_{\text{CP}}$ [C$_{[\bullet \text{w}h\bullet]}$ did ] Mary say [$_{\text{CP}}$ [$_{\text{C}}$ that ] John wanted what$_1$ ]
d. [$_{\text{CP}}$ [C$_{[\bullet \text{w}h\bullet]}$ did ] Mary say [$_{\text{CP}}$ [C$_{[\bullet \text{w}h\bullet]}$ that ] John wanted what$_1$ ]
e. [$_{\text{CP}}$ [C$_{[\bullet \text{w}h\bullet]}$ did ] Mary say [$_{\text{CP}}$ what$_1$ [C$_{[\bullet \text{w}h\bullet]}$ that ] John wanted t$_1$ ]
f. [$_{\text{CP}}$ what$_1$ [C$_{[\bullet \text{w}h\bullet]}$ did ] Mary say [$_{\text{CP}}$ t$_1'$ [$_{\text{C}}$ 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 $\phi$ 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 $\phi$ and tense features from C, T carries out agreement with the subject, via an Agree operation that values the $\phi$ 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:□*],[*$\pi$:□*], [Tns:PAST]}
b. [$_{\text{CP}}$ C:{[*#:□*],[*Gen:□*],[*$\pi$:□*], [Tns:PAST]} [$_{\text{TP}}$ T [$_{\text{vP}}$ [$_{\text{DP}}$ D:{[#:pl],[Gen:fem],[$\pi$:2],[*case:□*]} ⇒
c. [$_{\text{CP}}$ C [$_{\text{TP}}$ T:{[*#:□*],[*Gen:□*],[*$\pi$:□*], [Tns:PAST]} [$_{\text{vP}}$ [$_{\text{DP}}$ D:{[#:pl],[Gen:fem],[$\pi$:2],[*case:□*]} ⇒
d. [$_{\text{CP}}$ C [$_{\text{TP}}$ T:{[#:pl],[Gen:fem],[$\pi$:2], [Tns:PAST]} [$_{\text{vP}}$ [$_{\text{DP}}$ D:{[#:pl],[Gen:fem],[$\pi$: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 $\phi$ 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 $\phi$ 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, $\pi$ 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) [$_{\text{TP}}$ [$_{\text{DP}_1}$ John ] T [$_{\text{vP}}$ t$_1$ wondered [$_{\text{CP}}$ [$_{\text{DP}_3}$ which picture of himself$_{1,2}$]
C$_{[+wh]}$ [$_{\text{TP}}$ [$_{\text{DP}_2}$ Bill ] T [$_{\text{vP}}$ 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) [$_{\text{TP}}$ [$_{\text{DP}_1}$ John ] T [$_{\text{vP}}$ $t_1$ wondered [$_{\text{CP}}$ whether [$_{\text{TP}}$ [$_{\text{DP}_2}$ Bill ] T [$_{\text{vP}}$ $t_2$ saw [$_{\text{DP}_3}$ 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) [$_{\text{CP}}$ [$_{\text{DP}_3}$ Which picture of himself$_{1,2}$] [$_{\text{C}_{[+wh]}}$ does ] [$_{\text{TP}}$ [$_{\text{DP}_1}$ John ] T [$_{\text{vP}}$ $t_1$ think [$_{\text{CP}}$ $t'_3$ [$_{\text{C}_{[-wh]}}$ that ] [$_{\text{TP}}$ [$_{\text{DP}_2}$ Bill ] T [$_{\text{vP}}$ $t_2$ 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) [$_{\text{CP}}$ [$_{\text{DP}_3}$ 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 $\text{DP}_3$, 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 $\text{DP}_3$ 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. $[_{\text{DP}_3}$ which picture of himself$_{*\square*}$]
b. $[_{\text{vP}}$ $[_{\text{DP}_2}$ Bill ] saw $[_{\text{DP}_3}$ which picture of himself$_2$]]
c. $[_{\text{vP}}$ $[_{\text{DP}_2}$ Bill ] saw $[_{\text{DP}_3}$ which picture of himself$_\square$]]
d. $[_{\text{vP}}$ $[_{\text{DP}_1}$ John ] wondered $[_{\text{CP}}$ $[_{\text{DP}_3}$ which picture of himself$_1$] $[_{\text{C}'}$ C$_{[+wh]}$ $[_{\text{TP}}$ T $[_{\text{vP}}$ $[_{\text{DP}_2}$ 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 [$_{\text{vP}}$ [$_{\text{DP}_1}$ der König ] [$_{\text{v}'}$ [$_{\text{VP}}$ [$_{\text{DP}_2}$ den Sklaven ] [$_{\text{V}'}$ sich$_{1/2}$
that the king$_{nom}$ the slave$_{acc}$ REFL
im Spiegel zeigt ]] v ]]
in the mirror shows

b. dass [$_{\text{vP}}$ [$_{\text{DP}_1}$ die Gastgeber ] [$_{\text{v}'}$ [$_{\text{VP}}$ [$_{\text{DP}_2}$ die Gäste ] [$_{\text{V}'}$
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*:
[$_{\text{vP}}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]
b. *Object shift*:
[$_{\text{vP}}$ DP$_2$ [$_{\text{v}'}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V t$_2$ ]]]]
c. *Merge of T*:
[$_{\text{TP}}$ T [$_{\text{vP}}$ DP$_2$ [$_{\text{v}'}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V t$_2$ ]]]]]
d. *EPP-movement of the subject*:
*[$_{\text{TP}}$ DP$_1$ [$_{\text{T}'}$ T [$_{\text{vP}}$ DP$_2$ [$_{\text{v}'}$ t$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ 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*:
[$_{\text{vP}}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]
b. *Merge of T*:
[$_{\text{TP}}$ T [$_{\text{vP}}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]]
c. *EPP-movement of the subject*:
[$_{\text{TP}}$ DP$_1$ [$_{\text{T}'}$ T [$_{\text{vP}}$ t$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]]]
d. *Object shift*:
*[$_{\text{TP}}$ DP$_1$ [$_{\text{T}'}$ T [$_{\text{vP}}$ DP$_2$ [$_{\text{v}'}$ t$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ 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*:
[$_{\text{vP}}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]
b. *Merge of T*:
[$_{\text{TP}}$ T [$_{\text{vP}}$ DP$_1$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V DP$_2$ ]]]]
c. *First step of EPP-movement – DP$_1$ to workspace*:
[$_{\text{TP}}$ T [$_{\text{vP}}$ v [$_{\text{VP}}$ V DP$_2$ ]]] $\boxed{\text{DP}_1}$
d. *First step of v-to-T movement – T to workspace*:
[$_{\text{vP}}$ v [$_{\text{VP}}$ V DP$_2$ ]] $\boxed{\text{DP}_1}$, $\boxed{\text{T}}$
e. *First step of object shift – DP$_2$ to workspace*:
[$_{\text{vP}}$ v [$_{\text{VP}}$ V ]] $\boxed{\text{DP}_1}$, $\boxed{\text{T}}$, $\boxed{\text{DP}_2}$

---

[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*:
[$_{\text{vP}}$ [$_{\text{v}'}$ DP$_2$ [$_{\text{v}'}$ v [$_{\text{VP}}$ V ]]]] $\boxed{\text{DP}_1}$, $\boxed{\text{T}}$,
g. *Second step of v-to-T movement – v to T in workspace*:
[$_{\text{vP}}$ [$_{\text{v}'}$ DP$_2$ [$_{\text{v}'}$ [$_{\text{VP}}$ V ]]]] $\boxed{\text{DP}_1}$, $\boxed{\text{T-v}}$,
h. *Third step of v-to-T movement – T-v is remerged*:
[$_{\text{TP}}$ [$_{\text{T}}$ T v ] [$_{\text{vP}}$ [$_{\text{v}'}$ DP$_2$ [$_{\text{v}'}$ [$_{\text{VP}}$ V ]]]]] $\boxed{\text{DP}_1}$
i. *Second step of EPP-movement – DP$_1$ to SpecT*:
[$_{\text{TP}}$ DP$_1$ [$_{\text{T}'}$ [$_{\text{T}}$ T v ] [$_{\text{vP}}$ [$_{\text{v}'}$ DP$_2$ [$_{\text{v}'}$ [$_{\text{VP}}$ 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 $\alpha$ and $\beta$ can be inserted in either order), then to X (cf. (42c), with the exponent $\gamma$ inserted), and finally to Y (cf. (42d), where the vocabulary item $\delta$ is inserted).

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

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., [$\mathrm{f}_1$]) are a subset of the morpho-syntactic features borne by /b/ (e.g., [$\mathrm{f}_1$,$\mathrm{f}_2$]). 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/$\leftrightarrow$[$\mathrm{f}_1$] and /b/$\leftrightarrow$[$\mathrm{f}_1$,$\mathrm{f}_2$] 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*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$
b. *Root lexicalization*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$
c. *Feature copying on X cycle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1],[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$
d. *Vocabulary insertion on X cycle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}_{[\text{f}_1]}}\ /\text{a}/\ ]]\ \text{Y}_{[\text{f}_2]}\ ]$
e. *Vocabulary insertion on Y cycle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}}\ /\text{a}/\ ]]\ [_{\text{Y}}\ /\text{b}/\ ]]$

---

[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*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1]}\ ]$
b. *Root lexicalization*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1]}\ ]$

c. *Feature copying on Y cycle* **Cyclic Principle*
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_1]}\ ]$

d. *Vocabulary insertion into X on Y cycle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}}\ /\text{b}/\ ]]\ \text{Y}_{[\text{f}_1]}\ ]$

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

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 $[\text{f}_1]$ feature in question is in Y), but the derivation could have inserted /b/ into the smaller cyclic domain X first (which would then discharge $[\text{f}_1]$ 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 $[\text{f}_1]$ 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 *mʊn* 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 $\alpha$. Standardly, the concept of repair is modelled in such a way that the repair operation intrinsically violates some constraint $\beta$. This normally suffices to preclude application of the operation, except for specific environments where otherwise constraint $\alpha$ would have to be violated; in this case, as a last resort, $\beta$ can be minimally violated by the repair so as to satisfy $\alpha$. This thus presupposes a general violability of constraint $\beta$ in favour of a compliance with $\alpha$. 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 $\phi$-features.[42] The

---

[42] This conclusion suggests itself on the basis of very simple data, such as sentences exhibiting agreement with respect to $\phi$-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. [$_{\text{DP}}$ Sich$_{*2,1}$ (selbst) ] denkt Maria$_2$ [$_{\text{CP}}$ 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. ??[$_{\text{DP}_2}$ Die Studenten ] finden [$_{\text{CP}}$ [$_{\text{DP}}$ einander$_{*2,1}$ ] sollten [$_{\text{DP}_1}$
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*:
$[_{\text{vP}}\ \text{DP}_1\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ \text{V}\ [_{\text{DP}}\ \text{REFL}_{[*\square*]}\ ]]]]$
b. *Agree-Link*:
$[_{\text{vP}}\ \boxed{\text{DP}_1}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ \text{V}\ [_{\text{DP}}\ \boxed{\text{REFL}_{[\square]}}\ ]]]]$
c. *Agree-Copy*:
$[_{\text{vP}}\ \boxed{\text{DP}_1}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ \text{V}\ [_{\text{DP}}\ \boxed{\text{REFL}_{[1]}}\ ]]]]$

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*:
$[_{\text{DP}}$ D ... $[_{\text{DP}}$ REFL$_{[*\square*]}$ ]]
b. *Failure of Agree-Link*:
$[_{\text{DP}}$ D ... $[_{\text{DP}}$ REFL$_{[*\square*]}$ ]]
c. *Agree-Copy as a repair*:
$[_{\text{DP}}$ D ... $[_{\text{DP}}$ 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) [$_{\text{CP}_0}$ Maria$_3$ erzählt [$_{\text{CP}_1}$ [$_{\text{DP}_2}$ 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) [$_{\text{CP}_0}$ Maria$_3$ erzählt [$_{\text{CP}_1}$ [$_{\text{DP}_2}$ 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 $\mathrm{XP}_1$ 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 $\mathrm{XP}_n$ 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. $[_{\mathrm{vP}}\ \mathrm{XP}_{1\,[\mathrm{rel}]}\ [_{\mathrm{v}'}\ ...\ [_{\mathrm{VP}}\ -\ \mathrm{V}\ ]\ \mathrm{v}\ ]]$
b. $[_{\mathrm{vP}}\ \mathrm{XP}_{1\,[\mathrm{rel}]:[\bullet 1 \bullet]}\ [_{\mathrm{v}'}\ ...\ [_{\mathrm{VP}}\ \mathrm{XP}'_{1}\ \mathrm{V}\ ]\ \mathrm{v}\ ]]$

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 $\pi$ (cf. (57b)); and finally, the phase $\pi$P can now be left, in accordance with the PIC (cf. (57c)).

(57)
a. $[_{\pi\text{P}}\ [_{\text{YP}}\ \text{XP}_{1[\text{rel}]:[\bullet 1 \bullet]}\ [_{\pi'}\ \pi\ ...\ \text{XP}'_1\ ...\ ]]]$
b. $[_{\pi\text{P}}\ \text{XP}_{1[\text{rel}]}\ [_{\pi'}\ [_{\text{YP}}\ [_{\pi'}\ \pi\ ...\ \text{XP}'_1\ ...\ ]]]]$
c. $\text{XP}_{1[\text{rel}]}\ ...\ [_{\pi\text{P}}\ \text{t}_1\ [_{\pi'}\ [_{\text{YP}}\ [_{\pi'}\ \pi\ ...\ \text{XP}'_1\ ...\ ]]]]$

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 Spec$\pi$ is impossible (cf. (58b)); hence, the island YP remains strict at later steps because any extraction from YP (and $\pi$P) will now violate the PIC.

(58)
a. $[_{\pi\text{P}}\ [_{\text{YP}}\ \text{XP}_{1[\text{rel}]}\ [_{\pi'}\ \pi\ ...\ ]]]$
b. $*[_{\pi\text{P}}\ \text{XP}_{1[\text{rel}]}\ [_{\pi'}\ [_{\text{YP}}\ \text{t}_1\ [_{\pi'}\ \pi\ ...\ ]]]]$
c. $*\text{XP}_{1[\text{rel}]}\ ...\ [_{\pi\text{P}}\ [_{\text{YP}}\ \text{t}_1\ [_{\pi'}\ \pi\ ...\ ]]]$

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, $\pi$ 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): ‘Reflexes 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.

# Cyclicity in Minimalist Syntax

Gregory M. Kobele

## Abstract

What is the relationship between the extension condition and the cycle in e.g. phonology? I explore the analytical landscape and conclude that we must distinguish derivational and derived structure.

## 1. Introduction

The cycle, and references to cyclicity, appears throughout linguistics. Many terms in linguistics do double duty both as a name for a mechanism or analysis and as a label for a set of facts that that mechanism is intended to explain. As noted by Freidin (1978, 1999), cyclicity was introduced in transformational grammars to restrict the power of the model. Certain kinds of ungrammatical constructions (superraising, relativized minimality) could be blocked by means of this mechanism. By a *façon de parler* any mechanism which blocks these constructions could be termed cyclic. One might be forgiven, however, for wondering whether this situation is the reflection of a deep truth about language, as opposed to a historical accident. If this were the case, we might expect that the phenomena associated with cyclicity could and should be given a uniform formal treatment. In this short paper I will investigate whether cyclicity as currently implemented in minimalist syntax can be unified with cyclicity as implemented in other domains, and if so, how. My answer will be that cyclicity is exclusively about the mapping from one structure to another, and thus that there is no meaningful unification of *syntactic cyclicity* with *interface cyclicity*. However, I will argue that the popular conception of syntactic cyclicity is more properly viewed as an interface condition.

---

*Cyclicity*, 73–82
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

## 2. Cyclicity in Syntax and Elsewhere

Cyclicity is used in the domain of (morpho-)phonology to describe the interaction between modules: phonological rules apply to the intermediate outputs of bottom-up morphological structure building. As an equation:

$$\pi([\mathsf{m}\ \text{-}\mathsf{AFF}]) = \textsc{Phon}(\pi(\mathsf{m}), \text{-}\mathsf{AFF})$$

Here, $\pi(\cdot)$ is the (phonological) interpretation of a morphological structure. The operation $\textsc{Phon}(\cdot,\cdot)$ is responsible for attaching the right form of the affix to the right spot in $\pi(\mathsf{m})$, and applying the relevant phonological rules to the result. This allows for both phonological rules to depend on morphological structure, as well as limitations on this dependence to be stated.

Looking for an analogue of this in syntax, we would expect cyclicity to be a property of the interfaces: semantic terms (for example) should be computed incrementally during the derivation. However, Chomsky (1995: chap.3) asserts that requiring that all structure building target the root of the tree (with the implicit assumption that it target *only* the root of the tree) is the proper way to bring the concept of (strict) cyclicity into modern syntactic theory. Demanding that syntactic structure building target the root—the Extension Condition (EC)—seems a completely different kind of thing from allowing rules from one domain to depend on rules from another domain.

When cyclicity was first applied to syntactic theory, the formal model of syntax was very different. By *Aspects*, Chomsky (1965) had adopted the idea of a context-free base component. A grammar consisted of a context-free base, together with a (sometimes ordered) set of transformational rules. As a given transformational rule could apply in principle to any number of parts of a given input, some strategy was necessary to adjudicate between possibilities. Simple and natural rules seemed to require a *cyclic* (or *inside-out*) mode of scheduling: rules applied to a subtree before applying to anything containing it. This statement could perhaps be interpreted to allow a rule to apply to a subtree by actually applying to a proper subpart of this subtree. A *strict* reading of this statement would require that the rule apply to the entire subtree in question. An obvious question was which subtrees rules should apply to. A very simple proposal (see McCawley 1988: Chapter 6) is that rules should apply at *every* subtree. McCawley (1988) observes that this allows for rule ordering to be abandoned: rules apply in a (strict) cyclic manner whenever their structural descriptions are met. *This* sounds a lot more like what we began with: if you apply as soon as you can, you will (of course) be targeting the root of the current tree! A more influential proposal has it that rules are applied only at subtrees with certain syntactic properties (i.e. having a particular syntactic category). They then needn't apply to the entire subtree, but rather to any part of it that properly includes the previous subtree with that property. This in turn sounds very similar to ideas about *feature inheritance* (Chomsky 2008), which necessitates mild abrogations of the EC. It would seem that the identification of cyclicity in minimalist syntax with the EC (or relaxations of it) is faithful to the original conception of cyclicity in transformational grammar.

But what of the *other* notion of cyclicity, that which seemed to point towards the interfaces? How do these notions of cyclicity relate to one another, aside from onomastically? We will see that the original notion of cyclicity in transformational grammar, properly understood, was in fact the same as the interface notion. However, this raises issues with the glib identification of cyclicity in minimalist syntax with the EC.

## 3. Cyclicity in Transformational Grammar

The popular conception of transformational grammar is that a derivation proceeded by first obtaining a base structure (a tree), and then applying transformations to it in a cyclic manner. Transformations could in principle interact with (i.e. feed, bleed, etc) one another, but these could not retroactively affect the base structure—if, working bottom up, transformations change what was once a VP into an NP, this doesn't change the fact that this NP née VP was combined with an NP on its left to make an S in the base component. In other words, *syntactic selection* was purely a matter of the context-free base component, whereas transformations manipulated the God-given base structure.

Cyclic rule application over a tree is naturally stated in terms of interleaving the tree construction process with the rule application process. For this purpose, it is more useful to view context-free grammar productions in a bottom-up way; thus the production $S \rightarrow NP\ VP$ is viewed *not* as rewriting an $S$ into $NP$ and $VP$, but rather, as combining a pre-existing $NP$ and $VP$ together to obtain an $S$. Because the bottom-up interpretation of a production rule is less familiar, I will use a different notation so as to remind us that productions are to be so interpreted: $S \vdash NP\ VP$. Thinking of rules in a bottom-up way, we now actually *have* a particular $NP$ (say $u$) and a particular $VP$ (say $v$) that we are combining to produce an $S$. This $S$ is constructed out of the $NP$ $u$ and the $VP$ $v$ by introducing a new node labeled $S$ as the parent of these two (in that order): $[_S\ u\ v]$. This information is implicit in the production rule notation, but we can make it explicit in the bottom-up notation by writing the derived object in parentheses after the category name: $X(w)$, which can be read as '$w$ is an object of category $X$'. Our rule now looks as follows: $S([_S\ u\ v]) \vdash NP(u)\ VP(v)$.

As we want to treat categorial selection differently from the objects we construct (selection cannot be changed by transformational rules, the object we build can), it is useful to have a notation of this sort that distinguishes them. Curry (1961) calls the aspect of grammar dealing with categorial selection *tectogrammar*, and the aspect dealing with the objects constructed *phenogrammar*. This same distinction is made in Abstract Categorial Grammars (de Groote 2001: (ACGs)), where it is easier to see how this notion applies in general to *interfaces*: the tectogrammatical structure is the input to an interface, and the *phenogrammatical* structure is its output. The ACG perspective views the rule $S([_S\ u\ v]) \vdash NP(u)\ VP(v)$ as an operator $\rho$ of type $NP \to VP \to S$, which is interpreted as a function $[\![\rho]\!] = \lambda u, v. [_S\ u\ v]$ of type $\mathsf{tree} \to \mathsf{tree} \to \mathsf{tree}$. A well-typed term $M$ of atomic type can be thought of as representing a derivation of the tree $[\![M]\!]$, where $[\![M(N)]\!] = [\![M]\!]([\![N]\!])$.

Let us now write $\mathsf{tr}$ for the process of applying an ordered sequence of transformational rules to a structure. Then, as we surely want to apply transformations at $S$ nodes, we can add this information to our base rule as follows: $S(\mathsf{tr}([_S\ u\ v])) \vdash NP(u)\ VP(v)$. In other words, if we have an NP $u$ and a VP $v$, we can construct an S from them by applying a round of transformational rules to the object $[_S\ u\ v]$. Note that the property of being an S (say) is different from being an object whose root is labeled with the symbol $S$. Being an S means that you are something that can be used by a rule that has S on its right hand side. What is important here is that we see that the transformational component (the symbol $\mathsf{tr}$) lies squarely in the phenogrammatical component. This shows that the transformational notion of cyclicity is in fact the same as the interface notion of cyclicity familiar in the morphophonological world: it's just that the 'interface' here is the map between the base structure and the surface structure.

## 4. Cyclicity in Minimalism

One of the changes that occured on the road from transformational grammar to minimalism was the syntactification of transformations. In other words, transformations were subsumed by the operations of the base—in particular by the single binary operation MERGE.[1] The effects of MERGE are commonly thought of in terms of adding a new node which is an immediate parent of both of its arguments: $[\![\text{MERGE}(\alpha,\beta)]\!] = [[\![\alpha]\!]\ [\![\beta]\!]]$.[2]

In terms of the derivation, MERGE as defined above applies directly to its two arguments. This builds in the extension condition, at least derivationally (i.e. syntactically). As we wish to see how extension relates to cyclicity, we would like to avoid building extension into the system. In linguistic terms, we incorporate *timing* into the definition of MERGE. This requires us to permit MERGE to apply to subparts of expressions. In order to make this precise, MERGE must take *four* arguments—$\text{MERGE}(A,a,B,b)$—where two arguments (lower case $a$ and $b$) are the two terms which will be merged together, and the other two arguments (upper case $A$ and $B$) indicate at what point in the derivation this is to happen. We intend $a$ to be a subterm of $A$, and $b$ to be a subterm of $B$.[3] This is easiest to visualize via a picture, as shown in figure 1. In the figure, the left and right subtrees are the $A$ and $B$ arguments respectively. The dotted subtrees of each are the $a$ and $b$

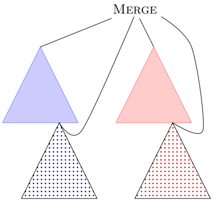


Figure 1: $\text{MERGE}(A,a,B,b)$

---

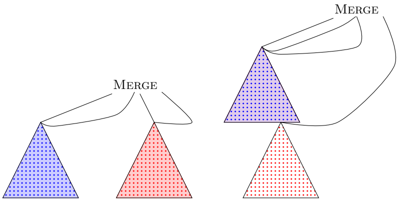
[1] This is not entirely accurate. The effects of transformations have been distributed across components of the grammar. Some effects of transformations have been moved to the interfaces, for example copy deletion/trace conversion.

[2] This is canonically written in set notation: $[\![\text{MERGE}(\alpha,\beta)]\!] = \{[\![\alpha]\!],[\![\beta]\!]\}$. This is of course just another notation for unordered trees where all daughters of a node are distinct.

[3] We can enforce this by changing the types of $a$ and $b$ from terms to pointers to nodes: $\text{MERGE} : \mathsf{forall}\ (A : \mathsf{Term}),\ \mathsf{Node}(A) \rightarrow \mathsf{forall}\ (B : \mathsf{Term}), \mathsf{Node}(B) \rightarrow \mathsf{Term}$.

arguments respectively. Intuitively, we want to understand $\textsc{Merge}(A,a,B,b)$ as saying “merge objects $a$ and $b$ together, but retroactively, after embedding them in $A$ and $B$ respectively”.

We will say that MERGE is countercyclic in $a$ if $A \neq a$, and countercyclic in $b$ if $B \neq b$.[4] Being countercyclic in this sense means the same thing as violating the extension condition (in a particular argument). If $A = a$ and $B = b$ then we have normal (cyclic) external merge. Internal merge obtains if $A = a$ and $A = B$ but $B \neq b$. That is, internal merge is countercyclic in $b$. Note that in both cases there are exactly two distinct arguments: in (cyclic) external merge these are the $A = a$ and $B = b$ arguments, and in internal merge these are the $A = a = B$ and the $b$ arguments. Thus in both cases, merge can be treated as a binary operation. These cases of merge are depicted in figure 2.



(a) external (b) internal

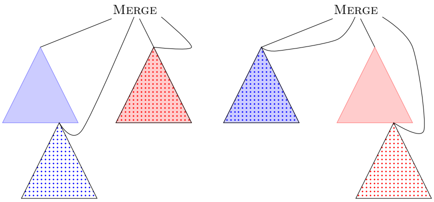
Figure 2: Canonical merge

Now let $A \neq B$. If $A \neq a$ but $B = b$ then we have countercyclic merge of $b$ inside of $A$, what Citko (2005) calls *parallel merge* and van Riemsdijk (2006) calls *grafting*. In the reverse situation ($A = a$ but $B \neq b$), we have what Nunes (2001) calls *sideward movement*. This is depicted in figure 3.

Crucially, allowing any countercyclicity in MERGE (including the countercyclicity of internal merge) means that the derivations are no longer proper trees, but rather graphs (multiple dominance structures). If there is a finite upper bound on the number of possible targets of reentrant arcs in any given structure, as is enforced by Ed Stabler’s so-called **SMC** constraint (Stabler

---

[4] This terminology is unfortunate, as it sounds like it has something to do with the notion of cyclicity under discussion. Whether it does is the subject of this paper. This terminology reflects current linguistic practice, which presupposes the connection.



(a) parallel (b) sidewards

Figure 3: Non-canonical merge

1997) in the case of internal merge, then these graphs can be encoded as trees.

### 4.1. The Extension Condition

As in transformational grammar, we have in minimalism two levels of structure: 1. the derivation, and 2. the structure so derived.We wish to ask whether the extension condition holds at which levels.

Here we are confronted with the fact that the EC is stated for trees, rather than for multi-dominance structures. By their very nature, all derivational operations satisfy the no tampering condition (NTC)—requiring that the inputs to an operation be preserved in the output—and this is often thought of as being stricter than the EC. On the other hand, we might wish to require that structure building *exclusively* target the root. This would then rule out as violating the EC any sort of reentrancy—in our terms, any merge step countercyclic in any of its arguments. As alluded to at the end of the previous section, reentrancy needn't be explicitly represented (and thus, can be formally eliminated) if the targets of reentrancy are uniquely determinable. As an example, many constraints on movement proposed in the minimalist program have a 'superlative' flavor—*Shortest Move*, *Attract Closest*, *Minimal Link* (though consider in this context the notion of *equidistance*)—which suggest that the identity of the mover might be uniquely recoverable just from the information that at a particular point in the derivation a movement took place. More generally, sidewards movement and parallel merge can be made compatible with this restricted derivational EC, so long as the sidewards mover and the target of parallel merger respectively can be reconstructed from the derivational stage at which merge applies.

The main question of interest is thus whether the EC holds of derived structure, which, as we have seen in the case of transformational grammar, can be thought of as the phenogrammar (i.e. in terms of interfaces). Viewing MERGE, as described above, as adding a new node to a graph, which immediately dominates its arguments, the objects it derives satisfy the EC just in case two of the following statements are true: $A = a$, $B = b$ and $A = B$.[5] In other words, this allows for exactly the structure building effects of cyclic merge and move, as the EC was designed to do. In contrast to our initial situation, now that derivation and derived structures have been distinguished, we see that this involves the interleaving of structure building and interpretation.

The difficulty with derived structure is that it is of necessity somewhat ephemeral—its entire *raison d'être* is to serve as the input to some other process, such as linearization. As we change our respresentation of derived structure, so too changes what might count as conforming to the EC. As a concrete example, consider the PF-interface, which we suppose is responsible solely for linearizing the terminals in our derived structure. As shown by Michaelis (2001) and Harkema (2001) (and intuited by Brosziewski (2003)), the mapping from derivation to string in Stabler's Minimalist Grammar formalism can be achieved by maintaining a tuple of strings without the need for derived structure. The EC no longer straightforwardly applicable to such a representation. However, if we understand the intent of the EC to be that of limiting changes to already constructed structures (along the lines of the NTC), then the generalized EC holds of a tuple of strings if only string concatenation is used to combine the components of tuples with one another (as opposed to substitution, or infixation). Splitting a single derived object up into parts (i.e. a tuple) allows for operations conforming to the generalized EC to apply which would have fallen afoul of the EC on the original derived object. For example, tucking-in movement (Richards 1999) satisfies the generalized EC so long as the sequence of specifiers of a head is split off from

---

[5] There are four possibilities, of which the following two have not yet been discussed: 1. $A = B = b \neq a$, and 2. $A = a = B = b$.The first case is the mirror image of internal merge, where the second to-be-merged argument of the MERGE operation properly contains the first. One could potentially think of this in linguistic terms as reprojection. The second case could be thought of as self-merge.

this head and its complement—this makes the innermost specifier position directly accessible without modifying an already built structure.

## 5. Conclusion

Defining cyclicity uniformly as interleaving structure building with interpretation accounts for what was called the transformational cycle in transformational grammar. The extension condition in minimalism can also be viewed in this manner, with the EC regulating the mapping from derivation to derived structure. However, it is extremely sensitive to representational choices, and thus appears *ad hoc*.

The EC can be applied to syntax proper (the derivation), where it requires that the targets of operations be uniquely determinable from their point of application. However, this does not have anything to do with cyclicity as a formal mechanism.

An important aspect of cyclicity as a mechanism is that the objects being constructed are only semipermeable to subsequent manipulation. The sensitivity of the EC to representational choices can be understood in this light—instead of condemning the EC as representation dependent, we use the kind of manipulation which is possible for already constructed derived object to infer a representation which allows (just) those to take place under the EC.

## References

Brosziewski, Ulf (2003): *Syntactic Derivations: A Nontransformational View*. Linguistische Arbeiten, Max Niemeyer Verlag, Tübingen.

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

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

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

Citko, Barbara (2005): ‘On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge’, *Linguistic Inquiry* **36**(4), 475–496.

Curry, Haskell B. (1961): Some Logical Aspects of Grammatical Structure. *In:* R. O. Jakobson, ed., *Structure of Language and its Mathematical Aspects*. Vol. 12 of

*Symposia on Applied Mathematics*, American Mathematical Society, Providence, pp. 56–68.

de Groote, Philippe (2001): Towards Abstract Categorial Grammars. In: *Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter, Proceedings of the Conference*. pp. 148–155.

Epstein, Samuel David and Norbert Hornstein, eds (1999): *Working Minimalism*. Number 32 *in* 'Current Studies in Linguistics', MIT Press, Cambridge, Massachusetts.

Freidin, Robert (1978): 'Cyclicity and the Theory of Grammar', *Linguistic Inquiry* **9**(4), 519–549.

Freidin, Robert (1999): Cyclicity and Minimalism. *In* Epstein and Hornstein (1999), chapter 5, pp. 95–126.

Harkema, Henk (2001): Parsing Minimalist Languages. PhD thesis, University of California, Los Angeles.

McCawley, James D. (1988): *The Syntactic Phenomena of English*. Vol. 1, The University of Chicago Press.

Michaelis, Jens (2001): On Formal Properties of Minimalist Grammars. PhD thesis, Universität Potsdam.

Nunes, Jairo (2001): 'Sideward Movement', *Linguistic Inquiry* **32**(2), 303–344.

Richards, Norvin (1999): Featural Cyclicity and the Ordering of Multiple Specifiers. *In* Epstein and Hornstein (1999), chapter 6, pp. 127–158.

Stabler, Edward P. (1997): Derivational minimalism. *In:* C. Retoré, ed., *Logical Aspects of Computational Linguistics*. Vol. 1328 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, pp. 68–95.

van Riemsdijk, Henk (2006): Grafts follow from Merge. *In:* M. Frascarelli, ed., *Phases of Interpretation*. Vol. 91 of *Studies in Generative Grammar*, Mouton de Gruyter, pp. 17–44.

# The Empirical Scope of the Strict Cycle Condition in Phonology

Daniel Gleim[*]

## Abstract

This paper examines four influential proposals to introduce the Strict Cycle Condition (SCC) from syntax to phonology, namely Kean (1974), Mascaró (1976), Kiparsky (1982), and Kiparsky (1985); and compares the empirical predictions each version makes. As has been noted previously (e.g. Kiparsky 1993, Rubach 2003), the two patterns that have been accounted for by the SCC, cyclic counterfeeding and derived environment effects, are both problematic for the SCC: cyclic counterfeeding might not exist, and derived environment effects are not general enough to be handled with such a rigid tool. One set of data that can be accounted for with (some versions of) the SCC remains: sandhi that are restricted to word boundaries. These can, however, also be accounted for by different, representational means. Still, there is a difference in predictions: representational analyses of these sandhi either predict a feeding relationship with other phrasal processes or make no predictions; the SCC predicts a counterfeeding relationship.

## 1. Introduction

The Strict Cycle Condition (SCC) was first adopted to phonology by Kean (1974), and, in its definitions by Mascaró (1976) and Kiparsky (1982, 1985) remained a staple of phonological theory during the seventies and eighties until it was abandoned, most prominently by Kiparsky (1993).

Historically, the SCC played an important role in accounting for derived environment effects, and to such an extent that it has often been confounded as solely a tool to account for those; however, the original empirical argument for the SCC is cyclic counterfeeding.

The precise empirical scope and predictions of the SCC in phonology have varied depending on the precise definition of the SCC itself and related

---

[*] I want to thank Gereon Müller, without whose input this paper would not exist, and the participants of the mini-workshop on the Strict Cycle Condition 2022 at Leipzig University for their valuable feedback.

*Cyclicity*, 83–104
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

concepts, importantly ‘cyclic rules’. The table in (1) gives an overview of the four implementations of the SCC discussed here and the empirical predictions they make.[1]

(1)

| | Cyclic cntrfeeding | DEE | Rules bleed SCC |
|---|---|---|---|
| Kean 1974 | Yes | No | No |
| Mascaró 1976 | Yes | Yes | Yes |
| Kiparsky 1982 | Yes | Partial | Yes |
| Kiparsky 1985 | Only within stems | Yes | Yes |

Deriving cyclic counterfeeding[2] is the core property of all versions of the SCC. A rule S that applies on a cycle i cannot, due to the SCC, feed a rule R that could apply on a cycle j.

(2)

| ABC | | Input cycle i | |
|---|---|---|---|
| — | Rule R | A → E __ D | *counterfed due to rule ordering* |
| ADC | Rule S | B → D __ C | |
| $[[ADC]_i Z]_j$ | | Input cycle j | |
| — | Rule R | A → E __ D | *blocked by SCC* |
| — | Rule S | B → D __ C | |
| ADCZ | | | |

A second property of the SCC, introduced by Mascaró (1976), is that it derives derived environment effects (DEE): the failure of phonological processes to apply within roots. The mechanism that conjoins the derivation of DEEs and the SCC may differ in the exact implementation; this is the reason why Kiparsky (1982) enforces DEEs only partially. The last crucial difference of implementation in the SCC lies in whether a phonological rule R on a cycle j, which changes some material belonging to cycle i, can feed a rule S that only uses material contained in i (3).

---

[1] DEE = Derived environment effects.

[2] Cyclic counterbleeding, on the other hand, follows directly from cyclic rule application and is thus not reliant on the SCC.

(3)

| | | | |
|---|---|---|---|
| AB | | Input cycle i | |
| — | Rule R | B → D __Z | |
| — | Rule S | A → E __ D | |
| $[[AB]_iZ]_j$ | | Input cycle j | |
| $[[AD]_iZ]_j$ | Rule R | B → D __Z | |
| ?? | Rule S | A → E __ D | *blocked by the SCC?* |

The core empirical questions regarding the SCC and its versions are thus the following: 1) Does cyclic counterfeeding exist, and if so, is it restricted to stems? The cases of cyclic counterfeeding that were important for the discussion have been reanalysed without cyclic counterfeeding or can be reanalysed along the same lines (e.g. Kiparsky 2000, 2015, Rubach 2003, Bermúdez-Otero 2006, 2011, 2018). These reanalyses, however, make a prediction for phrasal phonology that differs from versions of the SCC that extend to phrasal phonology: namely, whether general sandhi rules can or cannot feed sandhi rules restricted to word boundaries. 2) Do DEEs exist, and if so, are they obligatory? Here the answer seems to be well substantiated: DEEs do exist but are in no way obligatory. A mechanism like a DEE-inclusive SCC that enforces DEEs is thus undergenerating. 3) Are there rules applying in an inner cycle fed by a rule applying in an outer cycle? This question seems settled as well – such patterns do exist. I am not aware of a counterfeeding case, as would be predicted by Kean. This constitutes a crucial difference between the SCC in phonology and in morphosyntax, where the absence of precisely such a feeding interaction is a core argument for the SCC, compare e.g. Perlmutter and Soames (1979) or, in this volume, Müller (2023).

This paper is structured as follows. First, I discuss the four proposals by Kean, Mascaró, and Kiparsky in historical order. This discussion should suffice to answer the empirical question 3). In section 3, I will briefly deviate and discuss some conceptual problems regarding the SCC in phonology before, in section 4, coming back to the two remaining empirical questions. In this paper I sketch re-analyses without the SCC for Mascaró's Catalan and Kiparsky's Spanish case studies, but nor for Kean's Klamath case study since the original analysis with the SCC is actually not successful.[3]

---

[3] See Müller (2023) and Trommer (2023), both this volume, for reanalyses.

## 2. Versions of the SCC

### 2.1. Kean

The first work to introduce the SCC into phonology was Kean (1974). Her definition is the closest to the one of Chomsky (1973) for the SCC in morphosyntactic structure building and reprinted in (4).

(4) On any cycle A no cyclic rule may apply to material within a previous cycle B without making crucial use of material uniquely in A. (Kean 1974: 179)

There are two important differences between Kean's SCC and later iterations: (a) it does not derive derived environment effects, and (b) it is blind to phonological changes.Neither property is explicitly mentioned, which is unsurprising since her version precedes the others. The first property is crucial for her core case study of Klamath; a monomorphemic form such as /dewy/ must undergo the cyclic rule 'Sonorant Cluster' (6), in a first cycle, because otherwise it would not be able to undergo vowel deletion (7)[4] in a second cycle, as seen in the derivation in (8).

(5) /de-dewy/ → dedwiː 'shoot a bow and arrow'

(6) Sonorant cluster
Ø → ə / C__ [+son]{C,#}

(7) Vowel deletion
a. V → Ø / prfx [C__.CV
b. V → Ø / prfx [__

(8)

| | | |
|---|---|---|
| [dewy] | | |
| — | Vowel deletion | |
| [dewəy] | Sonorant Cluster | *not blocked by SCC!* |
| [de[dewəy] | | |
| [de[dwəy]] | Vowel Deletion | |
| — | Sonorant Cluster | |
| [de[dwiː]] | Postcyclic rules | |
| dedwiː | | |

---

[4] Rules are slightly simplified with respect to Kean.

The second difference follows directly from the definition of the SCC. Every change of a string PQ in an inner cycle is blocked unless its direct context is outside that cycle. Let us assume the structure in (9) and the rules in (10) in the given order.

(9) $[[PQ]_i Z]_j$

(10) a. Q → Y /__Z
b. P → X /__Y

On cycle i, there is no context for any of the rules, so none can apply. On cycle j, the first rule can and does apply, changing Q into Y, but this cannot feed the second rule since both the new Y and P are properly contained within cycle i.

Kean assumes that this property is crucial for her analysis. Her principal argument comes from the underapplication of Sonorant Cluster in Consonant-Sonorant-Sonorant-Consonant sequences, where the rule applies only once. Consider the underlying form in (11).

(11) [nt’iw-[otn-[el’g-a]]] → nt’iwtəllga ‘falls against something’

Here, vowel deletion on cycle j feeds Sonorant Cluster twice. However, Sonorant Cluster (SC) applies only once.[5]

(12)

| [otn[el’ga]] | | |
|---|---|---|
| [otn[l’ga]] | VD | |
| [otən[l’ga]] | SC | *applies only once* |

What needs to be blocked is the iteration of Sonorant Cluster on one cycle j, not the application of Sonorant Cluster on cycle k. Kean’s formalisation of the SCC is not capable of doing so. The entire context for Sonorant Cluster is there in cycle j, so it is not an instance of a phonological rule triggered

---

[5] Müller (2023; this volume) tries to repair Kean’s analysis by referring to morphological affiliation; Sonorant Cluster must insert the schwa inside a morpheme, that is, non-adjacent to a bracket. Therefore, SC is possible in the first cluster, but not in the second; the schwa would be (left or right) adjacent to the bracket. Bracket Erasure applies after SC, counterfeeding it. On the next cycle then, the SCC blocks SC. This works for the data at hand, but it fails to account for data such as (i), where we do find the epenthetic vowel directly at the morpheme boundary, both between the root and the suffix and the root and the prefix.

(i) /has-way’asg’-ys/ → hasəwy’asg’əys ‘loincloth’ in a later cycle that feeds a rule in an embedded cycle, as Kean alleges. It is also not an instance of counterfeeding that needs to be preserved since vowel deletion is needed to feed Sonorant Cluster. Kean’s analysis works if we amend the Sonorant Cluster rule with a diacritic which states that this rule applies non-iteratively only to the leftmost cluster. In this way, epenthesis to the second cluster on cycle i is blocked by the diacritic and on cycle j by the SCC.

(13)

| | | |
|---|---|---|
| [otn[el’ga]] | | |
| [otn[l’ga]] | VD | |
| [otən[l’ga]] | | *non-iterative version* |
| [nt’iw[otən[l’ga]]] | | |
| [nt’iw[tən[l’ga]] | VD | |
| — | SC | *blocked by SCC* |
| [nt’iw[təl[lga]]] | Postcyclic rules | |
| nt’iwtəllga | | |

However, forcing a non-iterativity requirement on a non-spreading rule is highly unusual or even unheard of.[6] Such a mechanism predicts typologically unwanted patterns galore, e.g. final devoicing only in the first voiceless coda that is new in each cycle.

(14) a. /kad/ → kat
b. /kadmad/ → katmad

The second case study in Kean, Welsh main stress, is not an undergeneration argument against theories without the SCC, but an argument of parsimony: if the SCC is accepted, it allows for a simpler and more elegant derivation of the Welsh facts.

Summarised, Kean gives an empirical argument against the SCC as a tool for derived environment effects, but she gives no convincing argument for the SCC itself since the Klamath case does not stand further scrutiny.

---

[6] Trommer (2023; this volume) derives the non-iterativity as an epiphenomenon by referring to some sort of antifaithfulness formalised in containment theory: epenthesis is blocked in positions where an underlying vowel is deleted. This derives the data at hand but fails to account for cases such as (i), where the loci of deletion and insertion accidentally align.

(i) bah-el’g’-a → bahəlg’a *bahlg’a ‘dries up’

### 2.2. Mascaró

Mascaró's definition of the SCC builds on Kean's and both restricts and extends its application.

(15) Strict Cycle Condition (Mascaró 1976: 9)
A cyclic rule R applies proverly [*sic*] on cycle j if either a, b or c is met:
a. R makes specific use of information uniquely in cycle j. That is, it refers specifically to some A in [$_j$XAY[$_{j\text{-}1}$ ... ]Z] or [$_j$Z [$_{j\text{-}1}$...]XAY].
b. R makes specific use of information within different constituents of the previous cycle which cannot be referred to simultaneously until cycle j. R refers thus to some A, B in [$_j$ X[$_{j\text{-}1}$ ...A... ] Y [$_{j\text{-}1}$ ...B...]Z].
c. R makes specific use of information assigned on cycle j by a rule applying before R.

He claims two crucial differences with respect to Kean, encapsulated in his clauses b. and c. Clause c. is indeed a crucial difference: Kean's SCC explicitly excludes phonological rules from being able to circumvent the SCC. There is, however, no data in Kean (1974) that necessarily needs the SCC to apply under such circumstances. On the other hand, it is a necessary weakening of the SCC for Mascaró given the data he analyses. In Catalan, vowel reduction changes unstressed non-high vowels to either *ə* or *u*. This change is fed by (lexical) de-stressing. The de-stressing rule itself does not violate the SCC according to Kean. The vowel reduction rule, however, does: It does not refer to any element outside of the inner circle. Under Mascaró's definition, its application is not blocked by the SCC because of the change from [+stress] to [–stress], information that was not available in the previous cycle.

(16) [[[trióm f]ál]ízm] → triumfəlízmə'triumphalism'

(17)

| | | |
|---|---|---|
| [[trióm f]ál] | | |
| [[triomf]ál] | De-stressing | |
| [[triumf]ál] | V reduction | *not blocked by SCC* |
| [[[triumf]ál]ízm] | | |
| [[[triumf]al]ízm] | De-stressing | |
| [[[triumf]əl]ízm] | V reduction | *not blocked by SCC* |
| [[[triumf]əl]ízmə] | Other rules | |

The second difference that Mascaró claims to have from Kean is assuming that the SCC does not hold over two previous cycles i which were not accessible together before a cycle j.

While such a structure is not explicitly discussed in Kean's definition, it becomes obvious from her discussion of Welsh stress that she does not intend her SCC to block such cases. In Welsh compounds and certain phrases, stress is shifted to the penult if the second member is monosyllabic. (18) shows the mapping that Kean assumes for a cycle j.

(18) [[cánhwyll]$_i$[brén]$_i$]$_j$ → [[cànhwýll]$_i$[brèn]$_i$]$_j$ 'candle-stick'

Both members have undergone previous cycles and have been assigned main stress. If Kean intended her SCC to work the way Mascaró seems to suppose she does, the stress shift would not be derivable because all information is in (separate) previous cycles i. One can thus conclude that Mascaró's clause b. is not so much a modification of Kean's SCC, but rather a more precise formulation.

The most compelling evidence for the SCC that Mascaró gives comes from the interaction of vowel reduction, mentioned above, and gliding. In Catalan, an unstressed high vowel is realised as an off-glide after another vowel. However, gliding is not fed by vowel reduction due to de-stressing *inside words.* Across word boundaries, however, gliding does apply to reduced vowels.

(19) a. /álʒebra-ik/ → əlʒəbrájk 'algebraic'
b. /raím-ɛ́t/ → rəimɛ́t *rəjmɛ́t 'raisin'
c. /prudúírá ɔ́ksidásjó/ → pruduiráwksidasjó 'produces oxidation'

(20)

| | | |
|---|---|---|
| [[pruduí]rá] | | |
| — | Gliding | *counterfed due to rule ordering* |
| [[prudui]rá] | De-stressing | |
| — | V reduction | |
| [[pruduirá][uksidəsjó]] | | |
| [[pruduirá][wksidəsjó]] | Gliding | *applies across word boundary, but blocked inside: *ui→uj* |

There is another important difference between Mascaró and Kean: Mascaró's version does account for derived environment effects. This is not explicitly stated in his definition, and this is the reason why authors such as Scheer (2010: 168) credit Kiparsky with the introduction of an SCC that aims to derive derived environment effects – unlike Kiparsky himself, who cites Mascaró as the originator. He does that by assuming that a so-called 0th cycle, which contains only the root, has two properties: no phonological rules apply, so the root is mapped onto its underlying form, and the 0th cycle suffices to count as a cycle, so the SCC will protect its output at later cycles (Mascaró 1976: 13).

### 2.3. Kiparsky 1982 and 1985

Kiparsky, arguably the author whose conception of the SCC was most influential in phonology, had actually proposed various versions of the constraint before abandoning it in Kiparsky (1993) and subsequent work. I will focus on Kiparsky (1982) and Kiparsky (1985).

Kiparsky (1982) redefines Mascaró's SCC so that it formally includes the derivation of derived environment effects.

(21) Strict Cycle Condition (Kiparsky 1982: 41)
a. Cyclic rules apply only to derived representations.
b. A representation Φ is *derived* w.r.t. rule R in cycle j iff Φ meets the structural analysis of R by virtue of a combination of morphemes introduced in cycle j or the application of a rule in cycle j.

In 1982, he derives the SCC from another principle, the Elsewhere Condition, and the assumption that for every entry in the lexicon, there is a rule that returns its underlying representation. This very specific rule blocks all other less specific rules that could manipulate the underlying representation. This system, unlike Mascaró's, cannot employ vacuous rules (see section 4 for a discussion of vacuous rules) in order to circumvent the SCC. Kiparsky claims that the data follow nonetheless, though it is not clear how. Another difference is that in Kiparsky (1982), phonology applies to roots – it is just blocked by the Elsewhere Condition in most cases. Syllable structure and stress , however, may be assigned on the root-only cycle so long as they do not conflict with pre-specified stress and syllabification and, in return, may feed the application of cyclical rules. This is exemplified with the Spanish example in (22). Final de-palatalisation turns a palatal sonorant into a coronal sonorant at the end of a word, as seen in (22a). It is bled by certain affixes, namely derivation and verbal inflection (22b), but counterbled by others, namely nominal inflection such as the plural (22c).

(22) a. /desdeɲ/ → desden ‘disdain’
b. /desdeɲ-es/ → desdeɲes ‘you disdain.SUBJ’
c. /desdeɲ-es/ → desdenes ‘disdains’

Kiparsky 1982 assumes that syllabification happens on the first cycle, which is only *desdeɲ* in the case of the plural but *desdeɲ+a* in the case of the verb form.[7]

(23)

| | |
|---|---|
| [desdeɲ] | |
| [des.deɲ.] | Syllabification |
| [des.den.] | De-palatalisation |
| [[des.den.]es] | |
| [[des.de.n]es. | Syllabification |
| — | De-palatalisation |
| des.de.nes | |

(24)

| | |
|---|---|
| [desdeɲ+a] | |
| [des.de.ɲa.] | Syllabification |
| — | De-palatalisation |
| [[des.de.ɲa.]es] | |
| [[des.de.ɲ.]es.] | Vowel deletion |
| [[des.de.ɲ]es.] | Syllabification |
| — | De-palatalisation |
| des.de.ɲes | |

Data of this type are problematic under Mascaró’s approach, where we would expect the SCC to block de-palatalisation until a boundary is merged or a postcyclic rule applies, in any case yielding the output *desdeɲes* for both the noun and the verb.

In *Some Consequences of Lexical Phonology*, Kiparsky (1985) divorces the SCC from the Elsewhere Condition again, creating an SCC that is more similar to Mascaró’s version, cf. (25).

---

[7] The thematic vowel -a is deleted preceding the subjunctive -e.

(25) Strict Cycle Condition (Kiparsky 1985: 89)
If W is derived from a lexical entry W’, where W’ is nondistinct from XPAQY and distinct from XPBQY, then a rule A→B / XP __ QY cannot apply to W until the word level.

This version rescinds the assumption from 1982 that stress or syllabification may feed rule application in monomorphemic domains – there are both empirical arguments of undergeneration and overgeneration against it. Its major innovation is a theory about which rules are cyclic and obey the SCC and which rules are non-cyclic and do not obey the SCC, via the introduction of a second lexical level: the word level. All stem level rules are cyclic whereas all word-level and all postlexical rules are non-cyclic. This connects cyclic effects and obeyance of the SCC with independent aspects of stem vs. word-level morphophonology. A similar proposal with minor differences was made by Booij and Rubach (1987). In order to analyse the aforementioned Spanish data under these new assumptions, de-palatalisation must be considered a word-level rule, the plural *-es* needs to be a word-level affix, and de-palatalisation must precede re-syllabification on the word level.

(26)

| | | |
|---|---|---|
| [desdeɲ] | | Stem level |
| [des.deɲ.] | Syllabification | |
| — | De-palatalisation | *Blocked by SCC* |
| [[des.deɲ.]es] | | Word level |
| [[des.den.]es] | De-palatalisation | *SCC does not hold* |
| [[des.de.n]es. | Syllabification | |
| des.de.nes | | |

It follows that the empirical scope of Kiparsky’s (1985) SCC is much smaller than Mascaró’s (1976); the SCC is expected to hold only in stem-level phonology and does not affect word- or phrase-level phonologies. Mascaró’s Gliding across word boundaries, for example, falls outside the scope of this SCC.

## 3. Conceptual Issues with the SCC

### 3.1. Modularity

The major conceptual issue with the SCC is that which part of the phonological or morphological representation triggers the SCC remains ill-defined. If we assume that the morphological brackets are an object which the phonological computation refers to, we run into one conceptual and one empirical problem. First, such a system would be grossly non-modular since the morphological bracketing is non-phonological information; under the assumption of modularity, phonological computation should not have access to this type of information. This is a problem that will accompany most or even all potential implementations of the SCC. The empirical problem lies with Mascaró's clause c., which exempts from the SCC a process applying in a smaller cyclic domain i on a cycle j, iff it is fed by a process that applies on cycle j. It is not immediately clear how the mechanism that checks for SCC violations could be circumvented if all it sees are brackets. Consider the structure and the rules in (27). According to Mascaró, rule R must be able to feed rule S. However, if the SCC is (informally) defined as 'On cycle j, do not apply a rule if focus and context are uniquely between the brackets labelled j-1', rule S will necessarily be blocked as well.

(27) a. $[[AB]_i Z]_j$
b. Rule R: B → D / __ Z
c. Rule S: A → E /__ D

An alternative to brackets would be some sort of index or feature attached to some part of the phonological representation. Let us assume this cycle index is attached to features, and every feature-change deletes the index. This would help us to avoid Mascaró's problem, but the conceptual issue remains: these indices are not genuine phonological material but rather a way for phonology to track morphological structure building and thus violate modularity. If, on the other hand, we assume them to be true phonological elements, we would expect them to be lexicalised (Chung 1983, Bermúdez-Otero 2012, Scheer 2020) into underlying representations. Such features predict two things: 1) affixes or words that idiosyncratically block processes because they are marked with the index and b) lexically specific non-derived environment effects. The latter might seem like a desirable result, but since the applicability of the process is tied exclusively to the underlying element, it is potentially still problematic. Take the Catalan root /kaɔs/, for example. This root undergoes vowel reduction but not gliding ([kaus] not *[kaws]), showing a derived environment effect for one process but not for the other. An approach along these lines, that is a modularity compatible SCC with indices, has recently been proposed under the name Harmonic Layer Theory (Trommer 2019, Zimmermann and Trommer 2021, 2022).

### 3.2. Bracket Erasure

Pesetsky (1979) noted that the SCC is at odds with the Chomsky-Hallean notion of Bracket Erasure, a mechanism employed to enforce the blindness of phonology to morphological structure.

(28) Chomsky-Hallean Bracket Erasure (Pesetsky (1979) based on Chomsky and Halle (1968: 15))
Given the nested constituents
$[\ .\ .\ .\ [\ .\ .\ .\ ]_{n-1}\ .\ .\ .\ ]_{n}$
the first rule of cycle j is: Erase brackets j-1.

Under this definition of Bracket Erasure and the conception that the SCC is enforced by reference to brackets, the SCC cannot hold because it has no brackets to operate on.

In order to reconcile Bracket Erasure and the SCC, Pesetsky redefines the former in (29).

(29) Pesetskian Bracket erasure (Pesetsky 1979)
Given the nested constituents
$[\ .\ .\ .\ [\ .\ .\ .\ ]_{n-1}\ .\ .\ .\ ]_{n}$
the last rule of cycle j is: Erase brackets j-1.

However, as Scheer (2010: 144ff) notes, the bracketing itself gives the power to restrict many processes to new cycles. In this way, cyclic counterfeeding can be accounted for without the SCC. If brackets and bracket erasure are combined with the SCC, the job of the SCC decreases: it is not needed for cyclic counterfeeding, and since Mascaró's clause c. is necessary, it is not useful for blocking a rule fed by another rule. Its remaining purpose is consequently to derive DEEs, which is, as will be discussed in the next section, also undesirable.

## 4. Predictions of the SCC

### 4.1. Derived Environment Effects

As discussed above, the first version of the SCC in phonology, Kean's transferral, does not make any statement about derived environment effects. On the contrary, in her analysis of Klamath, it is necessary to apply a process to a non-derived root prior to further affixation, compare the derivation in (8).

Mascaró modifies Kean's definition by adding two clauses, neither of which addresses monomorphemic domains. However, he assumes that phonology does not apply to bare roots, which he calls the 0th cycle. In the first cycle, which minimally contains two morphemes, the output of the 0th cycle (which, given that no phonology has applied, is identical to the input of the 0th cycle) is protected by the SCC. Thus, for Mascaró (1976), derived environment effects fall out from the combination of the SCC and the 0th cycle assumption. This effect is put to use for two processes in his analysis of Catalan: idiosyncratic failure of vowel reduction and underlying stressed tense mid vowels. The first concerns monomorphemic words which appear with unstressed full vowels [a, e, o], which normally reduce to [ə] or [u] depending on their roundness.

(30) a. bóston *bóstun 'Boston'
b. kólerə *kólərə 'cholera'

In the 0th cycle, phonology does not apply: ergo, there is no vowel reduction and /bóston/ is mapped to $[\text{bóston}]_i$. In any subsequent cycle, the SCC protects the unstressed /o/.[8] The second process regards stressed tense ó and é. Mascaró assumes that there is a rule that turns mid vowels lax if stressed.

(31) $[-\text{high}, -\text{low}, +\text{stress}] \rightarrow [-\text{tense}]$

This is argued for because if pre-stressing affixes shift the stress to otherwise tense vowels, they surface as mid.

---

[8] This derivation is, however, problematic for the root /kaɔs/, which has an underlying unstressed mid vowel that surfaces reduced: [kaus].

(32) a. /númər/ → númər ‘number’
b. /númər-ik/ → numɛ́rik ‘numeric’
c. /séntr/ → sentrə ‘centre’
d. /séntr-ik/→ sɛ́ntrik ‘centric’

Mascaró assumes that the rule in (31) can apply to the root séntr in /séntr-ik/ because of the (vacuous) stress shifting induced by the pre-stressing affix. If the underlying stress is replaced by a stress assigned by a rule, the laxing rule is free to apply. In forms like ‘séntrə’ on the other hand, the underlying stress and underlying vowel quality have been fixed by the 0th cycle, so under the absence of new information (even if vacuous), laxing is blocked from applying by the SCC.[9]

Kiparsky (1982, 1985) formally integrates derived environment effects into his definition of the SCC. Since then, the discussion of the SCC and derived environment effects has become more and more overlapping, whereas the other clauses of the SCC have found less discussion.[10] While the evidence for the existence of derived environment effects is overwhelming, the conjecture that no phonology applies to monomorphemic domains prior to concatenation must be considered obsolete (Kiparsky 1993). For this reason, Kiparsky abandoned the SCC in 1993. However, most of the evidence against the SCC came from the untenability of a generalised derived environment effect. The evidence against the Keanian core of the SCC with the relevant modifications by Mascaró seems to be much weaker.

---

[9] This analysis is, however, not compatible with the entire data given in Mascaró (1976). If this analysis were correct, all affixes containing a mid stressed vowel should undergo laxing as they are not protected by the SCC – the affix has not undergone a 0th cycle. This is, however, not the case. We find both affixes with tense vowels such as -és (IMPF.SUBJ) and lax vowels such as -ɛ́m (1PL).

[10] There are, of course, exceptions. Kaye (1992) adopts Kean’s definition of the SCC and explicitly excludes DEEs from its scope. It is somewhat unclear how it fits conceptually in his framework of ‘minimalist phonology’ because he forbids processes from having access to the derivational history. It is also unclear which data the SCC should derive or exclude under his wider assumptions: cyclic counterfeeding cannot be derived, because there is no counterfeeding on a given cycle in the first place. It could still block a process from applying if its context is created by a proces in an outer cycle. Just like Kean, Kaye does not offer an example for such an interaction.

### 4.2. Cyclic Counterfeeding

The SCC as defined in Mascaró, without the assumption that derives derived-environment effects, makes predictions only in systems that have either intra- or intercyclic counterfeeding opacity. If the system allows for intracyclic opacity, the SCC effect is created in the following fashion: A rule R and a rule S apply cyclically. On the first cycle, rule R cannot apply because its context is not met. Rule S applies and creates the context for rule R, counterfeeding it. On the second cycle, rule R is blocked by the SCC from applying in the context previously created by rule S. This is Mascaró's analysis for the de-stressing – gliding interaction in Catalan. De-stressing counterfeeds gliding on a cycle i, the SCC blocks gliding on a cycle j; compare the derivation in (20).

However, according to Cabré and Prieto (2004), the characterisation of the data by Mascaró is not entirely accurate: The affixes they discuss never undergo gliding. Mascaró and Cabré and Prieto (2004) do not discuss the same affixes. So if the data of both are correct, we have to conclude that a) gliding across affix boundaries is morpheme specific and b) these morpheme classes are not obviously stratally organised (33).

(33) a. /korne-u/ → kornew 'cultivated land'
b. /korne-u/ → korne.u 'I cultivate'

If we adopt a reasonable analysis with two different underlying shapes for gliding vs. hiatus formation, the intra-stratally opaque aspect of gliding in Catalan disappears.

(34) a. /korne-w/ → kornew 'cultivated land'
b. /korne-u/ → korne.u 'I cultivate'

Thus, in a form like *rəimɛ́t* thus, there is no blocking of gliding, because there is no gliding in the first place.

The interstratal counterfeeding aspects, however, remain. Across word boundaries, there is gliding of unstressed high vowels. This gliding only applies at word boundaries; other vowel-high vowel sequences are not affected.

If the claim of the absence of intrastratal opacity (e.g. Kiparsky 2000, 2015, Bermúdez-Otero 2018) is correct, the empirical domain of the SCC is restricted drastically. Kiparsky's (1985) version loses all meaning since it was restricted to the stem level, which is one stratum. But the scope of the more general versions is also reduced: the SCC's purpose would basically be to block sandhi that apply across word boundaries from applying inside words. Of course, other rules of phrasal phonology, such as flapping in English, do not care about word boundaries. This has always been acknowledged, and frameworks that employ a version of the SCC have two ways of deriving them. One is that the process in question is declared postcyclic and, as such, outside the scope of the SCC. Any attempt to categorise rules as either cyclic or postcyclic by some general mechanisms has failed (e.g. Kiparsky 1993, Scheer 2010). Another option is explored in Mascaró (1976). He employs fully vacuous rules that take an element X and return an identical element X, with the difference being that the new X is new and, thus, does not fall under the SCC. As mentioned above, the unstressed *o* of *bóston* does not undergo vowel reduction. In an affixed form, however, it reduces. Mascaró (1976) assumes a vacuous rule that turns an unstressed vowel into an unstressed vowel preceding a stressed vowel, enabling reduction to apply.

(35) a. bóston 'Boston'
b. bustuɲá 'Bostonian'

An approach that uses such means makes very few predictions, of course, so I will not further discuss this alternative.

The SCC approach to sandhi is surely not the only one. One can just as easily refer to the word boundary itself in the rule in order to make it only apply across boundaries. The rule in (36) would only affect /produirá##uksidasjó/ but not /rəimɛ́t/ simply because the latter does not meet its context.

(36) [+high +syll] → [–syll] / V##__

The similarity between these types of sandhi and derived-environment effects is, of course, striking; the process can apply in an environment derived in syntax, but not in a morphologically derived environment. Accordingly, approaches to blocking in non-derived environments that do not rely on the SCC can also be transferred to these cases. Take for example the underspecification approach, developed in Kiparsky (1993) and developed further in Rasin (2016). Here, a segment at a morpheme edge is underspecified for some certain feature, whereas it is fully specified morpheme-internally. Under concatenation, the underspecified segment may undergo a process that is blocked for the fully specified one. Transferred to phrasal phonology, this means that a word-edge segment has been underspecified in a previous cycle of phonology. Returning to the Catalan example ‘pruduiráwksidajó’, this entails that /ɔ́ksidásjó/ has become *Uksidasjó* by some rule like (37) by the time it enters phrasal phonology.

(37) u → U #__

This underspecified U is then either mapped to the glide *w* if it precedes a vowel or to *u* elsewhere (38).

(38) a. U → w /V__
b. U → u

If we now consider the derivation for the phrase in (39), it becomes obvious why these rules lead to gliding of the *u* but not of the *i*: the latter has never become underspecified and, thus, does not meet the rule’s description.

(39)

| | |
|---|---|
| [ɔ́ksidásjó] | |
| [ɔksidasjó] | De-stressing |
| [uksidəsjó] | V reduction |
| [Uksidəsjó] | Initial underspecification |
| [[pruduirá][Uksidəsjó]] | |
| [[pruduirá][wksidəsjó]] | U/I-Gliding |
| pruduiráwksidəsjó | |

Very similar data from Ecuadorian Spanish and Catalan – voicing of intervocalic s across words, but not inside – has been analysed by Bermúdez-Otero (2006, 2011) along exactly these lines. Another representational alternative for Catalan gliding could refer to the prosodic structure: the gliding could be triggered by a constraint that forces prosodic word initial unstressed syllables to have an onset, but not word-internal ones. These approaches shift complexity from the computation to the representation. However, they do not introduce representational mechanisms that have not been introduced for independent reasons such as underspecification or prosodic structure.

The predictions of the SCC and representational alternatives are however divergent with respect to process interactions in phrasal phonology. Generally, in SCC approaches, rules that do not obey the SCC, such as English flapping, are taken to apply after cyclic rules and appropriately called ‘postcyclic’. Sandhi rules that are sensitive to word boundaries – i.e. that obey the SCC –

are cyclic and must thus precede non-cyclic rules. Therefore, they must be counterfed or counterbled by postcyclic sandhi rules that are insensitive to word boundaries.

Depending on the mechanism that is employed to account for opacity, representational approaches would either make no predictions or predict the exact opposite, namely only transparent interactions in phrasal phonology. Imagine a structure like the one in (40) and the two phrasal processes in (41).

(40) $[[ABC]_i[DEF]_i]_j$

(41) a. G → D /__ D *across a word boundary*
b. C → G /__ B *irrespective of morphosyntactic structure*

With the SCC, process b. must counterfeed process a., because it is postcyclic, and process a. is cyclic (since it obeys the SCC). If we employ representational means to derive the restriction on process a. together with the adoption of a framework that does not allow for opacity on the phrase level, such as Stratal OT, process b. must feed process a. If opacity on the phrase level is allowed, both a feeding and a counterfeeding order of processes a. and b. is derivable.

## 5. Summary

The Strict Cycle Condition in phonology is a tool that has served to account for and predict the existence of two patterns: cyclic counterfeeding and derived environment effects. Due to its inviolability, at least in the cyclic part of phonological grammar, it also excludes the opposite patterns: feeding of a rule R by a rule S from a previous cycle and cyclic rules applying to underived structures. The SCC, therefore, seems to be both too strong and too weak: While DEEs exist, they are not general, and on the other hand, the evidence for cyclic counterfeeding seems to be increasingly scarce.

## References

Bermúdez-Otero, Ricardo (2006): 'Phonological Domains and Opacity Effects: A New Look at Voicing and Continuancy in Catalan', *Handout, Workshop on Approaches to Phonological Opacity, GLOW XXIX, Barcelona, 5 April 2006* .

Bermúdez-Otero, Ricardo (2011): Cyclicity. *In:* M. van Oostendorp, C. J. Ewen, E. V. Hume and K. Rice, eds, *The Blackwell Companion to Phonology*. Wiley Blackwell.

Bermúdez-Otero, Ricardo (2012): The architecture of grammar and the division of labour in exponence. *In:* J. Trommer, ed., *The morphology and phonology of exponence (Oxford Studies in Theoretical Linguistics 41)*. Oxford University Press, Oxford, pp. 8–83.

Bermúdez-Otero, Ricardo (2018): Stratal Phonology. *In:* S. Hannahs and A. R. K. Bosch, eds, *The Routledge Handbook of Phonological Theory*. Routledge, Abingdon, pp. 100–134.

Booij, Geert and Jerzy Rubach (1987): 'Postcyclic versus postlexical rules in Lexical Phonology', *Linguistic inquiry* **18**(1), 1–44.

Cabré, Teresa and Pilar Prieto (2004): 'Prosodic and analogical effects in lexical glide formation in Catalan', *Probus* **16**, 113–150.

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 and Morris Halle (1968): *The Sound Pattern of English*. Harper & row, New York, Evanston and London.

Chung, Sandra (1983): 'Tansderivational constraints in Chamorro phonology', *Lingua* **59**(1), 35–66.

Kaye, Jonathan (1992): On the interaction of theories of Lexical Phonology and theories of phonological phenomena. In: *Phonologica 1988. Proceedings of the 6th International Phonology Meeting*. pp. 141–155.

Kean, Mary-Louise (1974): 'The strict cycle in phonology', *Linguistic Inquiry* **5**(2), 179–203.

Kiparsky, Paul (1982): 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* **2**(1), 85–138.

Kiparsky, Paul (1993): Blocking in nonderived environments. *In:* S. Hargus and E. Kaisse, eds, *Studies in lexical phonology*. pp. 277–313.

Kiparsky, Paul (2000): 'Opacity and cyclicity', *The Linguistic Review* **17**, 351–267.

Kiparsky, Paul (2015): Stratal OT: A synopsis and FAQs. *In:* Y. E. Hsiao and L.-H. Wee, eds, *Capturing Phonological Shades Within and Across Languages*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 2–44.

Mascaró, Joan (1976): Catalan phonology and the phonological cycle. PhD thesis, Massachusetts Institute of Technology.

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

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’, *MIT Manuscript* .

Rasin, Ezer (2016): ‘Morpheme structure constraints and blocking in nonderived environments’, *MIT Manuscript* .

Rubach, Jerzy (2003): ‘Polish palatalization in derivational optimality theory’, *Lingua* **113**(3), 197–237.

Scheer, Tobias (2010): *A guide to morphosyntax-phonology interface theories*. De Gruyter Mouton, Berlin / New York.

Scheer, Tobias (2020): ‘On the lexical character of intermodular communication’, *Radical 1* .

Trommer, Jochen (2019): ‘A Harmonic Layer Account of Levantine Arabic Syncope’, *Talk at the OCP XVI, Verona, on January 17th, 2019* .

Trommer, Jochen (2023): The Strict Cycle Condition in Stratal OT. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig.

Zimmermann, Eva and Jochen Trommer (2021): ‘Getting stronger or weaker at every stratum: A new approach to tonal morphophonology’, *Poster at the Manchester Phonology Meeting on May 26th, 2021* .

Zimmermann, Eva and Jochen Trommer (2022): ‘Getting weaker across layers: The tonal phonology of Shona without stratal re-ranking’, *Talk at the University of Warsaw, linguistics colloquium, June 17th 2022* .

# Cycles in Syntax, Morphology, and Phonology

Paula Fenger

## Abstract

This paper is a brief investigation into determining what is necessary to investigate whether domains between modules come from the same source or not. Opposing views have been put forward, which will briefly be discussed. In order to come closer to an answer, it is important to investigate different modules in a single language, as well as conducting cross-linguistic work. This paper discusses various ways this can be investigated.

## 1. Introduction

Domains have been around in various forms in the different modules of grammar. For example, in the syntax domains have been defined in terms of bounding nodes, barriers, or most recently, phases. In the morphology domains are defined by assuming #/+ boundaries, levels, strata, or, just as in syntax, through phases. Even though linguistic theory has made extensive use of some form of a domain to account for the presence or absence of cyclic rule applications, it remains an open question whether domains in the different modules come from the same source.

There is a growing body of work that connects syntactic phases to morphological and phonological domains. This type of work looks extensively at mismatches between either the syntax and the phonology or the syntax and the morphology, leading to different conclusions (Marvin 2003, Adger 2007, Ishihara 2007, Cheng and Downing 2007, 2016, Pak 2008, Newell 2008, Kahnemuyipour 2009, Embick 2010, D'Alessandro and Scheer 2015, Sande et al. 2020, Harðarson 2022). One strand of research argues that it is better to have a single device in the grammar that delimits domains, while the other strand of research argues that there are too many mismatches between the different modules to maintain this.

It thus seems that a satisfying answer to the questions of whether and how domains across modules are related is difficult to find. Crucially, in

---

*Cyclicity*, 105–120
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

most cases the syntax-phonology interface is discussed separately from the syntax-morphology interface, and independent operations that are used in the syntactic literature to mask domains have generally not been considered when looking at domain mismatches. There is some research that does work on the syntax, morphological and phonological side in a single language, but not for the same phenomenon Bobaljik and Wurmbrand (2013), Harðarson (2022). However, recently is there work on the same phenomenon in a single language across syntax, morphology, and phonology: Bogomolets (2020), Fenger (2020), Georgieva et al. (2021), Georgieva and Borise (2022), Fenger and Weisser (2022).

The aim of this paper is to evaluate what is needed to investigate whether domains across modules are the same. Section 2 reviews the two types of research that have been done before. Then in section 3, the focus will be on what can be done to come closer to the question of whether domains are related, and it presents two preliminary case studies.

## 2. The State of the Art: Cycling through Previous Work

### 2.1. Starting from the Top

The first family of analyses that investigates if domains in the syntax have a direct effect on the phonology takes syntax as a starting point (cf. Selkirk 2011, Downing 2013, Cheng and Downing 2016, Bonet et al. 2019) . That is, they follow certain syntactic works who have argued for certain heads being the least controversial to be domain delimitors: *v*, C and D. The reasoning is that if these are domain-delimitors in the syntax, and the phonology is read off of the syntactic structure, one should find evidence of these three domains in the phonology as well . However, such a perfect mapping does not seem to be found when looking at various patterns, and the general conclusion is that there are too many mismatches to be accounted for by this type of direct inheritance of syntactic domains into the phonology.

To illustrate, consider penultimate vowel lengthening (PVL) in various Bantu languages (Kanerva 1990, Kenstowicz and Kisseberth 1990, Cheng and Downing 2007: a.o.). This phonological process is one in which there is lengthening of the penultimate vowel in a specific domain. This domain seems to be roughly similar to the syntax, but not completely. The illustrations below are from Chichewâ. In a simple mono-clausal sentence there is a single instance of PVL (1). When there are adverbials present, each of these are treated as a different domain from the clause, and as such there can be more instances of VL (2).

(1) físi a-na-dyá ḿ-k$\boxed{\text{áa}}$ngo
hyena 1.SUBJ-TAM-eat CL3-lion
'The hyena ate the lion' (Downing and Mtenje 2011: p.1968)

(2) Ti-ná-pírikitsa m-b$\boxed{\text{áa}}$lá [ ku-chókéra mu-m-s$\boxed{\text{ii}}$ka ] [
we-TAM-chase CL9-thief INF-leave LOC-CL3-market
ku-ítá ku-tchál$\boxed{\text{íi}}$tchi ]
INF-arrive LOC-church
'We chased the thief from the market to the church' (Downing and Mtenje 2011: p.1972)

The intuition, based on this type of data, is that the syntax plays a role in determining the domains for vowel lengthening, since adverbials are parsed separately from the main clause. However, the 'basic' delimitors, *v*, C, and D are not visible in the phonology. Thus, there is in fact a syntactic-phonology mismatch when looking at (1). The bracketed structure for (1) is given below, where there are two DPS, a vP and a CP. If syntax is what matters for the application of PVL, one should expect it on all three of the elements, contrary to fact. Only the last DP, lion, has a lengthened vowel.

(3) [$_{CP}$ [$_{DP}$ físi ] [$_{vP}$ a-na-dyá [$_{DP}$ ḿ-k$\boxed{\text{áa}}$ngo ]]]
hyena 1.SUBJ-TAM-eat CL3-lion

Although the process of PVL is sensitive to syntax, there is no perfect inheritance from syntax to phonology, since there is one instance of PVL but four syntactic domains. Because of these mismatches, researchers concluded that syntax is not directly mapped onto phonology.

Before coming to this conclusion, however, several issues need to be considered. Specifically, independent operations that can mask domains (Pak 2008, Harðarson 2022) in the syntax. One example of such an operation is movement inside DPs. That is, in Chicheŵa the word order in the DP is generally noun initial with modifiers to the right of the head noun (Mchombo 2004). To account for this word order pattern, N-to-D movement has been proposed (Downing and Mtenje 2011, Dehé and Samek-Lodovici 2009).

Coupled with a theory of spell-out where it has been argued that the only the complement of the phase head gets sent to the interfaces (Chomsky 2008), Harðarson (2022) accounts for the fact that there is no penultimate vowel lengthening in (1) for the DPs. If there is movement of the noun, it moves to the edge of the DP and therefore is not part of the spell-out cycle of the DP. On the next cycle, the DP is part of the domain together with the verb, and therefore there is only a single instance of PVL.

It is therefore crucial to consider independent operations that can mask syntactic domains, before considering whether or not the phonology is sensitive to the syntax directly, or an intermediate step is needed. Of course, one should not account for the data above by assuming N-to-D movement without actual evidence for this movement operation.

## 2.2. Starting from the Bottom

The other group of approaches takes under-application of phonological processes as a starting point. (Marvin 2003, Newell 2008, Newell and Piggott 2014, D'Alessandro and Scheer 2015, Creemers et al. 2018, Sande et al. 2020: a.o.). The guiding idea behind these approaches is that it is not desirable to have different operations that divides up pieces of grammar in different modules. Since there is some understanding of what this device is in the syntax, i.e., currently phases, it means that phases play a role everywhere in grammar, also in case of phonological mismatches.

An example of such a phonological mismatch is given for stress in Turkish. Turkish is considered an agglutinating language, and verbs generally expresses tense, mood and aspect morphology as suffixes. Moreover, stress assignment is generally an indicator of wordhood, and falls at the edge of a 'word' (Kornfilt 1997). Thus, stress can fall on any verbal morpheme, (4), even when there is an additional suffix before it, such as the causative, (5).

(4) a. kal-ˈ**du**
stay-PST
'stayed'

b. koş-ˈ**tur**
run-CAUS
'make run'

c. kal-ˈ**ıyor**
stay-PROG
's/he is staying'

(5) a. bit-**ir**-ˈ**iyor**
finish-CAUS-PROG
's/he is finishing'

b. koş-**tur**-ˈ**du**
run-CAUS-PST
'*x* made *y* run'

However, not every combination leads to stress on the final morpheme. Stress does not fall on the past when there is an aspectual morpheme, (6).

(6) a. gid-ˈ**ıyor-du**
stay-PROG-PST
'was staying'

b. gid-ˈ**ecek**-i-**di**-m
stay-FUT-COP-PST-1SG
'I will have gone '

Even though the past is the second morpheme from the stem in both (5b) and (6), the morpheme that the past attaches to matters. Specifically it means that (6) constitutes a mismatch, in that it seems that the past is suffixed to the verb stem, but stress seems to fall in the middle of the verb word.[1] Newell (2008) analyzes this mismatch as follows. She argues that the phonology is interpreted from the syntax, and that phases play a role in delimiting when stress is assigned. Her analysis is presented in (7): she assumes that the copula is the phase head *v*, and that the aspectual morpheme is below this phase head. The tense morpheme is above the phase head.

(7) [[[[ gid- $_{\sqrt{}}$] $\boxed{\text{ˈecek}}$ $_{asp}$] -i $\boxed{\text{v}}$] -di-m $_{T}$]
go FUT COP PST-1.SG

She argues thus that, in line with Chomsky (2001) that *v* is a phase head, and this head triggers spell-out of its complement, i.e., the heads below it which include the verb stem and the aspectual marker. Stress is assigned at this point as well. This analysis is different from several phonological analyses that treat markers such as the aspectual marker as special, in that they have a diacritic marking them for stress, and has the advantage that the stress assignment aligns with what is considered to be a phase in many syntactic works, namely *v*.

However, there is no independent research showing that the syntactic structure that is proposed based on the phonological phenomena is in fact the correct structure. For example, it is generally not common that viewpoint aspect is below the domain defining head *v*. Second, one counterargument that is being made against criticism of these type of approaches is that syntactic correlates of phase heads like *v, n, a* are not available, as they occur inside of

---

[1] One indication that the past is part of the same phonological word is based on vowel harmony: when there is a single vowel harmony domain, there is a single phonological word.

words, and as such they should have a different status than syntactic phase heads. Even though this might be true for cases where for example stress-shifting and stress-neutral affixes have been reanalyzed from level-1/level-2 affixes to below or above the phase head (Marvin 2003, Lowenstamm 2015, Creemers et al. 2018), this argument does not go through for a case like Turkish, since this is directly reflecting the syntactic clausal structure.[2]

Moreover, even though in the analysis by Newell it is assumed that syntactic and phonological phases align, similar types of research of other phonology-morphology mismatches have been arguing that other heads can be a phase head, since syntactic analyses haven't settled on which heads are phases (Sande et al. 2020). This could be true, but in order to investigate this, independent syntactic evidence is needed.

To summarize, although it is possible to analyze most phonological mismatches as coming from the syntax, the syntactic analyses proposed for these mismatches seem to not be independently corroborated, and lack advantages over the existing morpho-phonological analyses.

## 3. What to Do Next

The above sections very briefly illustrated the two lines of research, and what steps have been made to investigate domains. However, it also showed that there is a gap in what should be researched in order to look at mismatches more carefully. The next steps are laid out in this section.

First of all, independent operations that can mask domains are not taken into account. For example, in the syntactic and morphological literature on boundaries the following have played a role for variable domains: variation in (syntactic) movement is said to have played a role in domains extending (den Dikken 2007, Gallego and Uriagereka 2007), the question if domains are cross-linguistically the same (Bošković 2014), and if all (morpho-)phonological properties are sensitive to syntax or not. This means that, depending on the phenomenon and the language in question, the question of whether the interfaces require an different mapping procedure might look different. That is, it is important to compare minimal pairs by taking into account these different

---

[2] Fenger (2020) has a similar analysis to that of Newell, but argues that the phase head can be aspect above *v*, following Harwood (2014).

variables, which might lead to syntax-phonology or syntax-morphology mismatches.

The second property that should be considered is what diagnostics are used to count as a domain, and whether this can vary with the independent operations cross-linguistically. For example, Harwood (2014) shows that it is important to look if different tests cluster together, and that a richer aspectual structure shows in more detail where boundaries in English are.

Relatedly, and most importantly, in order to determine whether or not syntactic domains play a role in the phonology, it is important to do in-depth research in a single language for a single phenomenon across all modules. Generally, when the interfaces are investigated, only a single module (or two) are researched, leaving open whether or not the other modules align or not.

In the remainer of this paper I show preliminary results for two case studies that differ minimally. Namely, I consider verbal morphology in Japanese and Sinhala. These languages are both head final, agglutinating languages, and therefore have a rich testing ground for domains in the verbal domain. They overlap to a large extent, but a single syntactic operation (syntactic head movement to T) in one language, but not in the other, leads to different results in the domains and the mismatches.

## 3.1. Japanese

Japanese has been discussed extensively in the syntactic literature, and several works have argued for a *vP* phase. Interestingly, this boundary is also visible by looking at word-internal pitch accent (Fenger 2020). Syntactically, Tense in Japanese seems to be excluded from various syntactic processes. For example, fronting of the verb (+object), as a form of VP fronting, excludes Tense; and Tense cannot be elided (Funakoshi 2020).

(8) a. $[_{\text{TP}}$ aogaeru-o **tabe-ta**-sae ] Kaonashi-ga $t_{\text{TP}}$
Aogaeru-ACC eat-PST-even No.Face-NOM
'No Face even ate Aogaeru'

b. $[_{\text{VP}}$ aogaeru-o **tabe**-sae ] Kaonashi-ga $t_{\text{VP}}$ si-**ta**
Aogaeru-ACC eat-even No.Face-NOM do-PST
'No Face even ate Aogaeru' based on Funakoshi (2020)

Crucially, Aspect has generally been excluded from these debates, but interestingly it patterns with the root for VP fronting and ellipsis (to the exclusion of Tense) (Fenger 2020).

(9) ?[$_{\text{ASPP}}$ aogaeru-o **tabe-te**-sae ] Kaonashi-ga t$_{\text{ASPP}}$ **i-ta**
Aogaeru-ACC eat-ASP-even No.Face-NOM be-PST
'No Face was even eating Aogaeru' (Fenger 2020)

Interestingly, (Harwood 2014) argues that there is a split between (progressive) aspect and tense for English. This thus means that Japanese behaves syntactically similar to English with regard to domains.

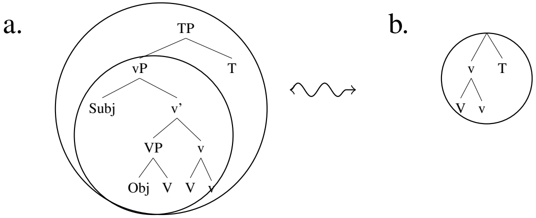
Turning to morphology and phonology, a similar split between the root, voice, aspect on the one hand, and tense is visible in auxiliary patterns and pitch accent (Fenger 2020). That is, T and Asp can never occur on the same verb, and an auxiliary is needed to host the tense morpheme. Wordhood tests, including conjunction, and putting material 'inside' words, reveal the same pattern. The difference between a causative form, and a form with aspect shows this difference: The tense morpheme is included in the pitch pattern in (10), but is excluded when the progressive is present. The (a.) examples provide the phonological breakdown, (b.) the morphological breakdown.

(10) a. ( L H H H H L )
he. da. ta. ra. se. ru
b. [ hedatar -sase -ru ]
be.distant -CAUS -PRES
'To make it distant'

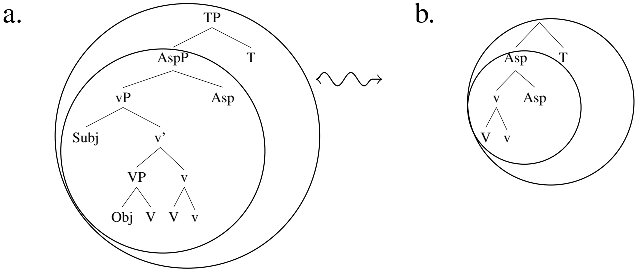
(11) a. (( L H H L ) L )
he. da. ta. te. ru
b. [ hedatar te ] -ru
be.distant -PROG -PRES
'It is being distant '

The exception to this pattern is (10): in these cases the root and the tense morpheme can form a single morphological and phonological unit. However, as shown above, syntactically Tense and the root are not part of the same domain. Put differently, there is in this case a mismatch between syntax and the morphology: the syntax shows two domains but the morphology only a single domain. However, in most other cases the syntax and the morphology align. there are always two domains in the syntax, which can include aspectual information, but never tense. There is a mismatch in the verbal domain when looking at verbal tenses, i.e. when the aspectual morpheme is missing. Thus, there is a morpho-syntax mismatch in very restricted environments and the morphological cycle can sometimes be bigger than the syntactic cycle. This can be represented as follows (12-13), where there are always two domains in the syntax (a.), and it can include aspect. However, when the heads are being mapped to phonological domains, there is a procedure in place that can delete the boundary in case there is no overt aspectual material present (see Fenger (2020) for details, following Embick (2010)).

(12) Multiple domains in syntax, single domain in morphology



(13) Multiple domains in syntax, mulitple in morphology



Thus, there is a mismatch in the morphology only, in a very restricted environment. This would not have been clear when only looking at a single TMA morpheme, i.e., simple tenses in the syntax, morphology, and phonology. Nor would it have been clear when only looking at only one module.

### 3.2. Sinhala

Sinhala (Indo-Aryan) is minimally different from Japanese in that it syntactically shows differences between simple tenses and complex tenses.[3] That is, even though on the surface both Japanese and Sinhala form a simple tense form (without aspect) synthetically, in Sinhala, unlike Japanese, these forms are a single unit in the syntax as well. Crucially, there is evidence that syntactically V+T also form a single domain in the language.

---

[3] I report here a small part of a larger project (Fenger and Weisser 2022, 2023).

Morpho-phonologically there is a distinction between umlaut triggers in the verbal domain. Sinhala has a rich inventory of verbal affixes, and among them are those that trigger fronting of the vowel on the stem. For example, affixes such as the causative (/wa/), the non-past (/nə/), or the indicative (/waa/) do not trigger fronting of the vowel, (14). However, the past (GEM/u) or the perfect (/laa/) do trigger fronting of the stem, (15-16).[4]

(14) adə-wə-nə-wa
pull-CAUS-NP-IND
'causes to pull'

(15) æd-d-a
pull-PST-IND
'pulled' (Past)

(16) ædə-la
pull-PERF
'pulled' (Perf)

Other umlaut triggers are the passive, the progressive, and the informal imperative. Among the umlaut-triggering morphemes there is a split. Even though they can all trigger fronting of the vowel when the trigger is adjacent to the target, they differ when morphemes intervene. Certain triggers, such as the past and the passive, can trigger vowel fronting across the causative, whereas triggers such as the perfect, the progressive, and the informal imperative can not. This is shown for the past in (17), and the perfect in (18).

(17) æd-də-u-wa
pull-CAUS-PST-IND
'made X pull'

(18) adə-wə-la
pull-CAUS-PERF
'have made X pull'

The behaviour of these different morphemes could be analyzed by assuming different domains for the different triggers. Since the past and the passive always trigger fronting, they seem to end up in the same morpho-phonological domain as the verb stem. Other morphemes, such as the perfect, seem to be variable: they are generally outside of the domain of the stem, but when adjacent they 'count' as if they are in the same domain. The question is whether these domains are only word-internal, or if they come from the syntax.

It turns out that the same split, for the same group of morphemes, can be observed for the choice of clausal negation. To see this, first consider the following sentence with an embedded and matrix clause that both are negated:

---

[4] There are various phonological processes that mask the underlying form. Short vowels, such as in the causative, are often reduced to schwa, and long vowels, such as in the indicative, are shortened. For clarity only the surface forms of the morphemes are given in the examples.

(19) [oyaa bat **no**-ka-nə-wa nisaa], oyaa-ʈə sindu-ak ki-u-e
2SG rice NEG-eat-NP-IND because 2SG-DAT song-DET say-PST-F
**nææ**.
NEG
'Because you don't eat rice, I didn't sing you a song.' (Slomanson 2008)

The embedded sentence has a prefixal negation /no-/; the matrix clause has a negation particle /nææ/. This thus means that there is a split for the choice of negation depending on being in an lower or higher domain. Crucially the simple tense form in the matrix clause cannot be negated with the prefix, (20a). However, the perfect can be, (20b).

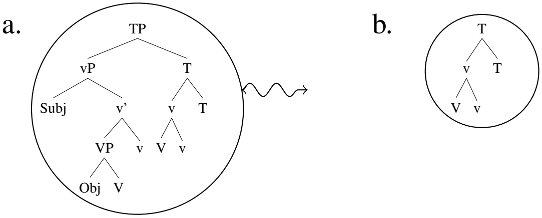
(20) a. *Mamə **no**-giy-a.
1SG NEG-go.PST-IND
'I didn't go'

b. Mamə **no**-anɖə-la
1SG NEG-cry-PERF
'I haven't cried'

Other morphemes that can take the prefix-negation in a matrix clause are the progressive and the informal imperative. This means that those morphemes that can only trigger umlaut when adjacent to the verb stem, correlate with taking clausal negation for embedded domains. The other umlaut triggers, which are always part of the same domain as the verb stem, are those that cannot take prefixal negation. This means that for the same group of morphemes both the syntax and the morpho-phonology make reference to the same domains.

One way to analyze this, is to say that for simple tenses there is movement of the verb to T. Since it moves to this position, it also carries along the phase head, and as such it extends the domain of the internal phase (den Dikken 2007: a.o.). Since there is only a single domain in this case, the only negation available is the matrix negation. Under the assumption that single morpho-syntactic domains are mapped onto single morpho-phonological domains, the umlaut trigger that is the passive or the past will always be in the same domain as the verb stem. This correspondence between domains is shown in (21).

(21) Single domains



This is then different from simple tenses in Japanese, where there is no extension in the syntax. The overt aspectual cases are the same in both languages. In this case there is no head movement to a higher position above the phase, and thus there is no domain extension of the vP phase. This means that syntactically there are two clause-internal domains, and the negation can be expressed with /no-/. The Perfect also remains outside of this first domain, under the assumption that material is interpreted at the interfaces cyclically.[5]

This means that in Sinhala there is more often than in Japanese a direct correspondence between the syntax and the morpho-phonology. The difference between the two languages stems from an independent syntactic mechanism that is present in one but not the other language.

### 3.3. Implications

Both Japanese and Sinhala have simple tense forms in the morphology, which are derived differently. Since syntactic domains can vary, possibly through the presence or absence of verb movement, differences in the morphology and phonology are expected. Thus, these two languages differ across one syntactic operation, leading to differences in the other modules as well. Moreover, both languages have the same type of syntax-morphology mismatch, for different features (simple tenses or the perfect). These mismatches are restricted, and can be derived through morphological extension.

Crucially, only looking at one form might give the impression that there is just random variation between the two languages, but turns out to be systematic when considering all modules, and independent points of language variation.

---

[5] For space reasons the pattern where the perfect can trigger umlaut is not discussed, but see Fenger and Weisser (2022). In essence this requires the same type of analysis as the simple tenses in Japanese.

## 4. Conclusion

This paper is a brief investigation into determining what is necessary to investigate whether domains between modules come from the same source or not. Before coming to the conclusion that there is no direct mapping between the modules, it is necessary to investigate various independent processes in the languages that can mask the output of one or the other module, such as head movement in the syntax. Crucially, in order to investigate domains, it is important to take a cross-modular and cross-linguistic approach.

## References

Adger, David (2007): 'Stress and phasal syntax', *Linguistic Analysis* **33**, 238–266.

Bobaljik, Jonathan David and Susi Wurmbrand (2013): Suspension Across Domains. *In:* O. Matushansky and A. Marantz, eds, *Distributed Morphology Today: Morphemes for Morris Halle*. MIT Press.

Bogomolets, Ksenia (2020): Lexical accent in languages with complex morphology. PhD thesis, University of Connecticut.

Bonet, Eulàlia, Lisa Lai-Shen Cheng, Laura J. Downing and Joan Mascaró (2019): '(In)direct Reference in the Phonology-Syntax Interface under Phase Theory: A Response to "Modular PIC"', *Linguistic Inquiry* **50**, 751–777.

Bošković, Željko (2014): 'Now I'm a phase, now I'm not a phase', *LI* **45**(1), 27–89.

Cheng, Lisa Lai-Shen and Laura J. Downing (2007): 'The Prosody and Syntax of Zulu Relative Clauses', *SOAS Working Papers in Linguistics* **15**, 51–63.

Cheng, Lisa Lai-Shen and Laura J Downing (2016): 'Phasal Syntax = Cyclic Phonology?', *Syntax* **19**(2), 156–191.

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

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, D. Michaels, C. P. Otero and M. L. Zubizaretta, eds, *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*. MIT Press, Cambridge, MA, pp. 133–166.

Creemers, Ava, Jan Don and Paula Fenger (2018): 'Some affixes are roots and some are heads', *Natural Language and Linguistic Theory* **36**(1), 45–84.

D'Alessandro, Roberta and Tobias Scheer (2015): 'Modular PIC', *LI* **46**(4), 593–624.

Dehé, Nicole and Vieri Samek-Lodovici (2009): 'On the prosody and syntax of DPs: Evidence from Italian noun adjective sequences', *Natural Language and Linguistic Theory* **27**, 45–75.

den Dikken, Marcel (2007): 'Phase Extension: Contours of a theory of the role of head movement in phrasal extraction', *Theoretical Morphology* **33**(1), 1–41.

Downing, Laura J. (2013): Issues in the Phonology-Syntax Interface in African Languages. *In:* O. O. Orie and K. W. Sanders, eds, *Selected Proceedings of the 43rd Annual Conference on African Linguistics*. Sommerville, MA: Cascadilla, pp. 26–38.

Downing, Laura J. and Al Mtenje (2011): ‘Prosodic phrasing of Chichewa relative clauses’, *Journal of African Languages and Linguistics* **32**, 65–112.

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

Fenger, Paula (2020): Words within Words: The internal syntax of verbs. PhD thesis, UConn.

Fenger, Paula and Philipp Weisser (2022): ‘Matching domains: The syntax, morphology, and phonology of the verb in Sinhala’, Ms. Universität Leipzig.

Fenger, Paula and Philipp Weisser (2023): What umlaut tells us about the underlying morphological structure of verbs in Sinhala. In: *Proceedings of (Formal) Approaches to South Asian Languages 12*. p. 105–123.

Funakoshi, Kenshi (2020): ‘Verb-raising and VP-fronting in Japanese’, *The Linguistic Review* **37**(1), 117–146.

Gallego, Ángel J. and Juan Uriagereka (2007): Subextraction from subjects: A Phase Theory account. *In:* J. Camacho, N. Flores-Ferrán, L. Sánches, V. Deprez and M. J. Cabrera, eds, *Romance Linguistics 2006: Selected papers from the 26th Linguistic Symposium on Romance Languages (LSRL)*. John Benjamins, Amsterdam, pp. 155–168.

Georgieva, Ekaterina and Lena Borise (2022): ‘Towards a DM account of verbal morphophonology in Udmurt’, Paper presented at IMM 20.

Georgieva, Ekaterina, Martin Salzmann and Philipp Weisser (2021): ‘Negative verb clusters in Mari and Udmurt and why they require postsyntactic top-down word-formation’, *Natural Language and Linguistic Theory* **39**(2), 457–503.

Harðarson, Gísli Rúnar (2022): Converging syntactic and phonological domains. In: *NELS proceedings*. NELS, pp. 1–14.

Harwood, William (2014): ‘Rise of the Auxiliaries: a case for auxiliary raising vs. affix lowering’, *The Linguistic Review* **31**(2), 295–362.

Ishihara, Shinichiro (2007): ‘Major phrase, focus intonation, multiple spell-out (MaO, FI, MSO)’, *The Linguistic Review* **24**, 137–167.

Kahnemuyipour, Arsalan (2009): *The Syntax of Sentential Stress*. Oxford University Press.

Kanerva, Jonni M. (1990): *Focus and Phrasing in Chichewȃ phonology*. Garland Pub.

Kenstowicz, Michael and Charles W. Kisseberth (1990): Cizigula tonology: The word and beyond. *In:* S. Inkelas and D. Zec, eds, *The Phonology-Syntax Connection*. CSLI, Stanford, pp. 133–166.

Kornfilt, Jaklin (1997): *Turkish*. Descriptive Grammars. Routledge.

Lowenstamm, Jean (2015): Derivational affixes as roots: Phasal spell-out meets English Stress Shift. *In:* A. Alexiadou, H. Borer and F. Schäfer, eds, *The syntax of roots and the roots of syntax*. Oxford University Press, Oxford, pp. 230–259.

Marvin, Tatjana (2003): Topics in the Stress and Syntax of Words. PhD thesis, Massachusetts Institute of Technology.

Mchombo, Sam (2004): *The syntax of Chichewa*. Cambridge University Press.

Newell, Heather (2008): Aspects of the morphology and phonology of phases. PhD thesis, McGill.

Newell, Heather and Glynne Piggott (2014): 'Interactions at the syntax-phonology interface: Evidence from Ojibwe', *Lingua* **150**, 332–362.

Pak, Marjorie (2008): The postsyntactic derivation and its phonological reflexes. PhD thesis, University of Pennsylvania.

Sande, Hannah, Peter Jenks and Sharon Inkelas (2020): 'Cophonologies by Ph(r)ase', *Natural Language and Linguistic Theory* **38**, 1211–1261.

Selkirk, Elisabeth (2011): The Syntax-Phonology Interface. *In:* J. Goldsmith, J. Riggle and A. Yu, eds, *The Handbook of Phonological Theory*. people.umass.edu, Oxford.

# Part II

# Exploring Evidence

# Cyclicity in Morphological Movement: The Case of Potawatomi Inverse Marking

Felicitas Andermann*

## Abstract

In this paper I show that analysing Potawatomi inverse marking in Harmonic Serialism (Müller (2020)), a derivational version of Optimality Theory, as a reflex of morphological movement obliterates the need for assuming two Voice heads in the syntax, nominative-accusative and absolutive-ergative alignment at the same time, or one exponent encoding both arguments. In addition to inverse marking, morphological movement and movement-related repair operations can derive participant reduction, i.e. the unexpected absence of certain exponents whenever they realise the less salient argument. My analysis crucially relies on the STRICT CYCLE CONDITION (SCC) as well as on the CYCLIC PRINCIPLE as assumed for Merge and movement operations in Transformational Grammar (Chomsky (1957)) (and as proposed for syntax and morphology by Bobaljik (2000: 3)): There are two cycles, one for Merge operations and a second one for movement, and Merge and movement are subject to the same cyclic domains, i.e. exponents must move in the same order in which they are merged. While the CYCLIC PRINCIPLE makes predictions for movement, the SCC makes predictions for deletion and insertion. Moreover, a branching derivation (see Müller (2023), this volume) of exponence-driven insertion strengthens the SCC in comparison to a non-branching derivation.

## 1. Introduction

In Potawatomi (Algonquian, North America) transitive animate (TA, see section 2) verbs (see Hockett (1948) or Stump (2001)), a direct (DIR) marker /a/ occurs when the subject is a speech act participant (SAP, i.e. 1st or 2nd person) and the object has 3rd person features, as in (1a), or when both subject and object are 3rd person but the object is marked as obviative (**less salient**), as in (1c). In the reverse cases, where a 3rd person subject acts on an SAP object,

---

[*] I would like to thank Sophia Meyer, Gereon Müller, Mariia Privizentseva and Adrian Stegovec as well as the other members of the research unit on Cyclic Optimization and the participants of the Morphology/Syntax colloquium for their helpful comments. All errors are my own.

*Cyclicity*, 123–158
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

as in (1b), or an obviative 3rd person subject acts on a proximate 3rd person object, as in (1d), instead of /a/, an inverse (INV) marker /UkO/ occurs in the inflected forms that are otherwise identical to the direct forms in (1a) and (1c).

(1) a. k-wapm-**a**-wa-k
‘you (pl.) see them’
b. k-wapm-**UkO**-wa-k
‘they see you(pl.)’
c. w-wapm-**a**-wa-n
‘they see him (obv.)’
d. w-wapm-**UkO**-wa-n
‘he (obv.) sees them’

In short, the direct marker occurs when the subject outranks the object on a person hierarchy as in (2), and the inverse marker occurs when the object outranks the subject.

(2) 1/2 > 3 > 3OBV

I derive this pattern as follows: An exponent realising the **less salient**[1] argument on a salience hierarchy 2/1 > 3 > OBV morphologically moves to the right edge of the word and leaves a copy in the base position which is overtly realised by the direct or inverse marker, where the direct marker /a/ is a copy of object movement and the inverse marker /UkO/ is a copy of subject movement.

There are numerous approaches to direct / inverse marking in numerous morphological theories, which differ regarding which grammatical category direct and inverse markers encode: They have been analysed as portmanteau markers encoding case and transitivity (Halle and Marantz (1993)), as case and person (Branigan and MacKenzie (2002), Henze and Zimmermann (2011), Bruening (2017)), as case, person, and animacy/salience (Wunderlich (1997), Stiebels (2002), Trommer (2001, 2006), as person markers that in reality must be assumed to have case diacritics (Despić et al. (2019), Steele (1995)), as instances of differential case marking where either ergative or accusative is assigned (Déchaine (1999), Kushnir (2015)). In Oxford (2018, 2022), the inverse marker is analysed as an elsewhere marker and the direct marker as a 3rd person object marker, in Stump (2001) direct and inverse markers realise a major reference feature that is assigned to either the subject or the object or no argument, and in Anderson (1992) the inverse marker is analysed as a reflex of modification of the morphosyntactic node into which exponents are inserted. The trade-offs of these approaches are that the DIR and INV markers

---

[1] Throughout the paper, exponents realising the **less salient** argument are **boldfaced** while exponents realising the *more salient* argument are *slanted.*

have to encode features of both subject and object (in the analyses of DIR and INV as portmanteaux), two Voice heads or case alignment systems have to be assumed (as in Déchaine (1999), Kushnir (2015), and Oxford (2018, 2022)), or the analysis relies on specific features or a morphosyntactic operation that cannot be independently argued for (Stump (2001), Anderson (1992)).

In this paper, I present an analysis previously developed in Andermann (2022) in which the distribution of the direct and inverse marker is derived via morphological movement. First, all exponents realising the **less salient** argument are merged, then all exponents realising the *more salient* argument are merged, and finally, exponents realising the **less salient** argument move to the right edge of the word, leaving a copy in the base position which is overtly realised by the direct marker in the case of object movement, as schematised in (3a-b) and by the inverse marker in the case of subject movement, as schematised in (3c-d).

(3) a. stem-**obj**-*subj* ⇒ c. stem-**subj**-*obj* ⇒
b. stem-**DIR**-*subj*-**obj** d. stem-**INV**-*obj*-**subj**

Morphological theories differ as to whether they allow for such movement of exponents or even predict it. Most morphological theories such as Paradigm Function Morphology (Stump (2001)), Network Morphology (Brown and Hippisley (2012)), Minimalist Morphology (Wunderlich (1997)) and Information-Based Morphology (Crysmann and Bonami (2016)) have no possibility of deriving morphological movement. In Distributed Morphology (Halle and Marantz (1993)), exponent movement is possible but has to be derived via additional operations such as lowering (Embick and Noyer (2001)), local dislocation (Embick and Noyer (2001)), or metathesis (Arregi and Nevins (2012)).

In a derivational optimality-theoretic framework like Harmonic Serialism (Müller (2020)), on the other hand, movement follows without further ado from the interaction of exponent realisation (henceforth *Merge*) and alignment constraints. In each step of the derivation, only one operation (Merge, movement, or deletion of an exponent) may be carried out. Given a ranking MERGE CONDITION » L ⇐ Root » MAX(X) » MAX(Y) » X ⇒ R, where the MERGE CONDITION requires Merge of exponents and the ranking of MAX constraints determines the order in which the exponents X and Y are merged, MAX(X) is ranked highest of all MAX constraints, and an exponent X must be merged as a suffix due to a high-ranked constraint L ⇐ Root requiring the root to be aligned with the left edge of the word. Subsequently, another exponent Y must be merged, also as a suffix, in violation of the constraint X ⇒ R that requires X to be right-aligned. Merging Y as a prefix would violate L ⇐ Root, and not merging Y would violate MAX(Y), which is ranked higher than X ⇒ R. In the next step of the derivation, however, X ⇒ R can be satisfied by movement of X to the right edge of the word. Note that movement is only possible because the constraints are satisfied one after another. In Standard Parallel Optimality Theory (SPOT), X and Y would be realised simultaneously, with X at the right edge. The derivational nature of Harmonic Serialism is therefore crucial for the analysis developed below.

In addition to being derivational, a movement-based analysis of the Potawatomi TA Independent Order paradigm must also be cyclic. Following Müller (2020), I adopt two notions of cyclicity. Firstly, there are two morphological cycles, each followed by a phonological cycle; one morphological cycle is finished when all exponents have been merged, and a further morphological cycle is finished when all other operations (movement, deletion, etc.) have taken place such that the inflectional form cannot be further optimised. Given this assumption, it turns out that the Potawatomi paradigm is completely regular and well-behaved as far as (first-cycle) Merge operations are concerned, and complications such as the unexpected occurrence of direct and inverse markers as well as the unexpected absence of some exponents, as described in section 2, are due to movement and movement-related operations that take place in the second cycle.

Secondly, both Merge and movement operations obey the STRICT CYCLE CONDITION (SCC, based on Chomsky (1973)).

(4) STRICT CYCLE CONDITION
Within the current domain $\delta$, an operation may not target a position included within another domain $\varepsilon$ that is dominated by $\delta$.

Merge proceeds from the root outwards, as in the toy example in (5), where $E_1$, $E_2$, $E_3$, and $E_4$ are exponents. Consequently, the current cyclic domain is always the domain that comprises the left and right edge of the inflectional form.

(5) a. root-$E_1$ c. root-$E_1$-$E_2$-$E_3$
b. root-$E_1$-$E_2$ d. $E_4$-root-$E_1$-$E_2$-$E_3$

This means that exponents may only be merged at the left or right edge in the first cycle and move only to the left or right edge in the second cycle, as in (6a). For deletion operations, Müller (2020) assumes a weak version of the SCC, represented in (6b). According to this weak version, deletion may also target a position adjacent to the leftmost or rightmost one provided that it is a consequence of the Merge or movement operation that applied immediately before. In (5c), for instance, as a consequence of merging $E_3$, $E_2$ may be deleted but not $E_1$. For repair-driven exponent insertion, an even weaker version is tacitly assumed in Andermann (2022) and made explicit in (6c), namely that this insertion need not even apply at a position adjacent to the edgemost one but must be a direct consequence of an immediately preceding operation that, in turn, must have targeted the leftmost or rightmost position.

(6) a. *Merge* and *movement* may only target the left or the right edge.
b. *Deletion* must target a position adjacent to the left- or rightmost position and must be the consequence of an immediately preceding Merge or movement operation (that has targeted the left or right edge, as per (6a))
c. *Repair-driven insertion* may apply to any position but must be a direct consequence of an immediately preceding Merge or movement operation (that has targeted the left or right edge, as per (6a)).

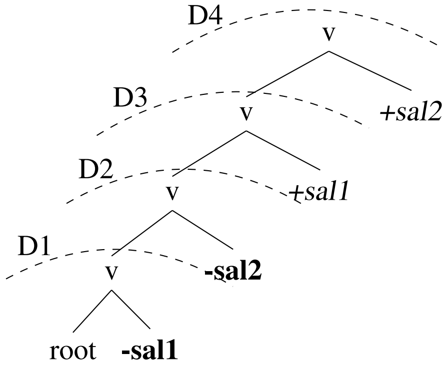
Moreover, as I show in section 4, for a successful derivation of the Potawatomi TA Independent Order paradigm, the CYCLIC PRINCIPLE as assumed for Transformational Grammar (formalised in (7), see e.g. Perlmutter and Soames (1979)) must hold for both first-cycle Merge and second-cycle movement operations in the same way such that both types of operations are subject to the same cyclic domains and each exponent constitutes a cyclic domain.

(7) CYCLIC PRINCIPLE
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.

In other words, all exponents must move in the same order in which they are merged, and an exponent may only move once the exponent previously merged has reached its final landing site.

For the case of Potawatomi, where first all exponents realising the less salient argument and then all exponents realising the more salient argument are merged, the cyclic domains are schematised in (8).

(8) *Cyclic domains in Merge and movement operations*



This resembles derivations in Transformational Grammar (Chomsky (1957), Perlmutter and Soames (1979)), where all basic phrase-structure building operations, which correspond to external Merge operations in both syntax and morphology, precede all transformations, including movement operations, and both types of operations proceed strictly bottom-up. Versions of the CYCLIC PRINCIPLE have been proposed for Minimalism under the name of the EARLINESS PRINCIPLE (see Pesetsky (1989), Pesetsky and Torrego (2001): A syntactic operation must apply as soon as its structural condition is met) and FEATURAL CYCLICITY (see Richards (2001), Preminger (2018): Active features that can trigger operations must do so as soon as possible). This notion of cyclicity has also been proposed by Bobaljik (2000: 3) for the morphology-syntax interface where a) morphology interprets syntax rather than feeding it, i.e. comes after syntax and b) morphology proceeds root-outwards. This means syntactic structure is interpreted via morphology in the same order in which it has been built, namely from the lowest, most embedded, to the highest domain.

This paper is structured as follows: section 2 is an overview of the person/number inflection paradigms of transitive animate verbs with which this paper is concerned. In section 3, I briefly discuss evidence for morphological movement and overt reflexes thereof and argue that these overt reflexes can be either full copies or minimally realised traces, as has been proposed by Pesetsky (1998), Hornstein (2000), and Bošković and Nunes (2007), among others, for overt (PF) reflexes of movement, and that in Potawatomi the inverse marker is a minimal trace rather than a full copy. Based on this reasoning, I illustrate my analysis of inverse marking as a minimal trace of exponent movement with two sample derivations, one of the direct form *n-wapm-**a**-k* ('I see them') and one of the inverse form *n-wapm-**UkO**-nan-k* ('they see us'). Section 5 concludes.

## 2. The Pattern

Potawatomi, like other Algonquian languages, has four types of verbs that differ by valency and animacy of their single or internal argument: Inanimate Intransitive (II); Animate Intransitive (AI); Transitive Inanimate (TI), where the object is inanimate; and Transitive Animate (TA), where the object is animate. Transitive Animate verbs have a direct paradigm where the subject outranks the object in the person hierarchy in (2), an inverse paradigm where the object outranks the subject, and a local paradigm where both arguments are speech act participants (SAP) and therefore ranked equally in the hierarchy. Furthermore, all verb types have different paradigms depending on whether they are used in main clauses (*independent order*) or subordinate clauses (*conjunct order*). This paper is concerned exclusively with TA verbs in the independent order.

Person/number inflection of Potawatomi TA verbs follows the template generally observed for Algonquian languages in the literature, as represented in (9). Inflectional forms consist of a prefix encoding person features of the *more salient* argument, the direct or inverse marker, also referred to as *theme sign* (Bloomfield (1946: 98-102)); a central ending (Goddard (1969: 38)) encoding person and number of the *more salient* argument; and a peripheral ending (Goddard (1969: 38)), which is either an obviative marker or realises person and number of the **less salient** argument.

(9) *Algonquian inflection template*

| Prefix | Stem | Theme sign | Central ending | Peripheral ending |
|---|---|---|---|---|
| PERS | | **DIR/INV** | PERS/NUM | **PERS/NUM//OBV** |

The direct and inverse paradigms of Potawatomi Transitive Animate verbs (adapted from Hockett (1948)) are represented in (10)-(11).

(10) *Independent Order Transitive Animate Direct*

| OBJ →<br>SUBJ ↓ | 3SG | 3PL | 3OBV |
|---|---|---|---|
| 1SG | n-wapm-**a**-∅ | n-wapm-**a**-k | n-wapm-**a**-n |
| 2SG | k-wapm-**a**-∅ | k-wapm-**a**-k | k-wapm-**a**-n |
| 3SG | | | w-wapm-**a**-n |
| 1PL.INCL | k-wapm-**a**-mUn | k-wapm-**a**-mUn | k-wapm-**a**-mUn |
| 1PL.EXCL | n-wapm-**a**-mUn | n-wapm-**a**-mUn | n-wapm-**a**-mUn |
| 2PL | k-wapm-**a**-wa | k-wapm-**a**-wa-k | k-wapm-**a**-wa-n |
| 3PL | | | w-wapm-**a**-wa-n |

(11) *Independent Order Transitive Animate Inverse*

| SUBJ →<br>OBJ ↓ | 3SG | 3PL | 3OBV |
|---|---|---|---|
| 1SG | n-wapm-**UkO**-∅ | n-wapm-**UkO**-k | |
| 2SG | k-wapm-**UkO**-∅ | k-wapm-**UkO**-k | |
| 3SG | | | w-wapm-**UkO**-n |
| 1PL.INCL | k-wapm-**UkO**-nan | k-wapm-**UkO**-nan-k | |
| 1PL.EXCL | n-wapm-**UkO**-nan | n-wapm-**UkO**-nan-k | |
| 2PL | k-wapm-**UkO**-wa | k-wapm-**UkO**-wa-k | |
| 3PL | | | w-wapm-**UkO**-wa-n |

In most cases, the direct and inverse forms differ only in the direct/inverse marker, with the exception of the 1PL ↔ 3 forms, where the marker /mUn/ appears when a 1PL subject acts on a 3rd person object, and /nan/ occurs instead of /mUn/ when a 3rd person object acts on a 1PL subject. This, however, is only the case in the present tense forms, while in the preterite, /mUn/ encodes both 1PL subjects and 1PL objects. This suggests that /mUn/ is a generic 1PL marker and /nan/ is a 1PL object marker whose occurrence is restricted to the present tense by some mechanism I will disregard here. Apart from /nan/, no other marker is specified for case. The person prefixes, /k-/ for 2nd person, /n-/ for 1st person, and /w-/ for 3rd person, appear in the direct as well as in the inverse paradigm, as does the central ending /wa/ that marks 2PL and 3PL arguments as well as the peripheral endings /-k/ encoding less salient 3PL arguments and /-n/ realising obviative arguments. I therefore assume the feature specifications in (12) for exponents, where person is decomposed into [±1 ±2 ±3], number into [±pl], and obviation into [±obv]; and the 1PL object marker /nan/ is additionally specified for a feature [+ob(ject)].

(12) *Feature specifications*

a. *Prefixes*
/$n_1$/ ↔ [+1],
/$k_1$/ ↔ [+2],
/w/ ↔ [+3],

b. *Central endings*
/mUn/ ↔ [+1 +pl],
/nan/ ↔ [+1 +pl +ob],
/wa/ ↔ [-1 +pl],

c. *Periph. endings*
∅ ↔ [+3 -pl]
/$k_2$/ ↔ [+3 +pl],
/$n_2$/ ↔ [+3 +obv]

Given these feature specifications, a further problem arises in addition to the distribution of the direct and inverse marker: one has to account for the fact that 1) the person prefix /w/ encoding 3rd person surfaces only once in 3 ↔ 3OBV and never in SAP ↔ 3rd person configurations, 2) in 1st person plural contexts only the 2nd person prefix /$k_1$/ appears but not the 1st person prefix /$n_1$/, 3) that the central ending /wa/ encoding 2PL or 3PL does not occur in 1 ↔ 3PL constellations and occurs only *once* in 2PL ↔ 3PL and 3PL ↔ 3PL contexts. This phenomenon, known as *participant reduction* (Trommer (2003)), follows without further ado from the interaction of alignment constraints and MAX constraints in an optimality-theoretic framework. The absence of /w/ in the 2 ← 3PL forms in (1), for instance, can be derived by a ranking L ⇐ +2 » L ⇐ +3 » MAX (+2) » MAX (+3): Both affixes /$k_1$/ ↔ +2 and /w/ ↔ +3 compete for the position at the left edge of the word. Deleting /w/ ↔ +3 yields the best constraint profile as it satisfies the highest-ranked constraint (+2 is now at the left edge) and does not violate the next-highest ranked constraints L ⇐ +3 and MAX (+2), but only the lowest-ranked constraint MAX (+3). Similarly, the fact that the exponents /$k_2$ and /$n_2$/ occur after the 2PL/3PL exponent /wa/ and the 1PL.OBJ marker /nan/ but are dropped after the generic 1PL exponent /mUn/ can be accounted for by assuming high-ranked right-alignment and MAX constraints referring to the exponent /mUn/.

Thus, under the assumption that there are two morphological cycles, one for the Merge operations and one for the movement operations, the Potawatomi person/number inflection paradigm can be derived in a simple and straightforward manner as far as the Merge operations are concerned. All affixes are merged neatly in a row. Their insertion follows from the basic mechanism of disjunctive blocking by compatibility and specificity (implemented in OT by MAX and DEP/IDENT constraints) without any impoverishment rules and without portmanteau agreement. Rather, both unexpected exponence in the form of the direct and inverse marker and unexpected non-exponence in the form of participant reduction arise only in the second morphological cycle where movement and movement-related operations take place as they are repair phenomena driven by the interaction of alignment and MAX constraints.

## 3. Movement-Related Copying

Evidence for repair-driven exponence triggered by morphological movement comes from Bantu languages. Hyman (2003) discusses cases of exponent copying in Chichewa (see (13)-(14)) resulting from conflict between the Causative-Applicative-Reciprocal-Passive (CARP) template and the Mirror Principle. The affix order in (13a), for instance, where the applicative suffix /il/ precedes the reciprocal suffix /an/, is grammatical under the compositional ([[Appl] Rec]) interpretation as well as the non-compositional ([[Rec]Appl]) interpretation, whereas the reverse affix order in (13b), which would mirror the composition [[Rec]Appl], is ungrammatical.

(13) a. mang-il-an
tie-APPL-REC
'tie for each other'
[[Appl] Rec]
'tie each other for/at'
[[Rec] Appl]
b. *mang-an-il
tie-REC-APPL

(14) a. mang-an-il-an
tie-REC-APPL-REC
'tie each other for/at'
[[Rec] Appl]
*'tie for each other'
[[Appl] Rec]
b. *mang-il-an-il
tie-APPL-REC-APPL

However, under the [[Rec]Appl] interpretation, the form in (14a), where the reciprocal affix both precedes the applicative affix and follows it, is grammatical, thus respecting both the Mirror Principle and the CARP template. On the other hand, doubling of the applicative suffix, as in (14b), is ungrammatical.

Moreover, for (14a), only the compositional interpretation ([[Rec]Appl]) is available. Hyman (2003: 256-257) therefore argues that copying of the reciprocal suffix in (14c) is an instance of repair, and Gleim et al. (2023: 17) remark that the occurrence of such copies could be considered evidence for both morphological movement and movement-related copying.

A crucial difference between the Chichewa data in (13)-(14) and the Potawatomi data in (10)-(11) is that in Chichewa, the moved item and the copy are identical in shape whereas in Potawatomi, they are not. It is therefore not entirely clear whether in Chichewa it is the copy closer to the stem or the copy farther away from the stem which is inserted by repair. In contrast, in Potawatomi, the distribution of /a/ and /UkO/ suggests that these are copies of exponent movement. In (1a-b), the exponent /$k_2$/, realising the less salient 3PL argument, is at the right edge, no matter whether the less salient argument is the object, as in (1a), or the subject, as in (1b). The same holds for the obviation marker /$n_2$/ in the obviative contexts in (1c-d). The distribution of the direct and inverse marker, on the other hand, does depend on whether the less salient argument is a subject or an object but does not depend on the person and obviation feature specification of the less salient marker itself. Moreover, there is potential evidence for the direct marker /a/ being a generic object marker and /UkO/ being a generic subject marker from underspecified object constructions and underspecified subject constructions (Andermann (2022: 40-43)). This suggests that unlike Chichewa exponent movement, which leaves a full copy, Potawatomi exponent movement leaves a minimal trace realised by the direct or inverse marker.

In syntax, full and minimal realisations of overt movement reflexes have been analysed by Pesetsky (1998), Hornstein (2000), and Bošković and Nunes (2007) within the copy theory of movement, and their occurrence is attributed to constraints on pronunciation rather than to movement types. All these analyses rely on the assumption that movement always leaves copies, and that in the unmarked case all but one of these copies are deleted to satisfy a constraint SILENT-*t* requiring all lower copies to be deleted in Pesetsky (1998: 25) or as a consequence of Kayne's (1991) Linearity Correspondence Axiom (LCA) in Hornstein (2000) and Bošković and Nunes (2007). It is furthermore assumed in these approaches that there is a general preference for pronouncing only the highest copy and deleting all lower copies (see Bošković and Nunes (2007: 29)).

In cases of multiple overt realisations of full copies, Nunes (2004) and

Bošković and Nunes (2007) assume that the lower of the overtly realised copies is invisible to the LCA because it has undergone a morphosyntactic fusion operation (as proposed by Halle and Marantz (1993)) with an adjacent constituent before linearisation applies. Minimal realisations of copies, on the other hand, have been taken to be repair items introduced by the grammar to minimise violation of SILENT-*t* (Pesetsky (1998)) or to repair a PF violation incurred by LCA-triggered chain reduction (Hornstein (2000)). Crucially, Hornstein (2000: 171) points out that pronominals only ever occur in repair contexts, for which he accounts by excluding them from the numeration and positing that they are introduced by grammar, analogous to Arnold's (1995) analysis of do-support, where *do* is likewise not assumed to be part of the numeration.

In my analysis of Potawatomi inverse marking, to derive the distribution of the direct marker, which realises subject movement, and the object marker, which realises object movement, I assume that the exponent /$k_2$/ realising the less salient argument, by moving to the right edge, splits a feature [+subject] or [-subject] off in violation of a constraint MAX ([±subject]), which requires subsequent realisation by a marker encoding either [+subject] or [-subject]. To account for the fact that the generic object and subject markers are not inserted in the Merge cycle, I assume, following Hornstein (2000: 171), that they are excluded from the numeration and introduced by grammar.

## 4. Analysis

Let us now look at the derivation of Potawatomi transitive animate forms in detail. The tableaux in (24)-(35) and (36)-(47) show the derivations of the forms in (15), where a first person singular subject acts on a third person plural object, and (16), where a third person plural subject acts on a first person plural object:

(15) n-wapm-a-k
1-see-DIR-3PL
'I see them'

(16) n-wapm-UkO-nan-k
1-see-INV-1PL.OBJ-3PL
'they see us' (3PL > 1PL)

In the Harmonic Serialism framework developed by Müller (2020), a lexical-realisational morphological theory in Stump's (2001) sense, morphology is presyntactic and takes place in the numeration (see Chomsky (2001)). A stem in the lexicon is assumed to bear a fully specified, language-specific, well-formed set of inherent features (see (19)-(20)). Non-inherent features, which are also fully specified, are added in the numeration (see (21)-(22)). The resulting set of features, henceforth referred to as feature structure, provides the context for underspecified inflection markers that form part of *morphological arrays* as defined in (17):

(17) *Morphological arrays* (Müller (2020: 126))
An exponent $\alpha$ is in a morphological array for a grammatical category X (MA$_X$) in the domain of a syntactic category (part of speech) W iff (i), (ii), or (iii) hold.
(i)$\alpha$ realises a grammatical category Y in the domain of W by a morpho-syntactic feature that is a (possibly underspecified) instantiation of X.
(ii)$\alpha$ realises a grammatical category Y in the domain of W c feature that is a (possibly underspecified) instantiation of Y, and there is an exponent in MA$_X$ that realises Y.
(iii)$\alpha$ is a unique radically underspecified exponent for X in the domain of W.

For each morphological array encoding a grammatical category (or fusion of categories) X, there is a structure-building feature [• X •] and a corresponding MERGE CONDITION MC(X), as defined in (18), which triggers morphological exponence. This feature [• X •] is part of the input but is discharged once the morphological array associated to it is accessed.

(18) MERGE CONDITION (Müller (2020: 14))
A structure-building feature [• X •] that is accessible in the input participates in (and is deleted by) a Merge operation in the output.

Transitive animate verbs in Potawatomi agree with both subject and object and therefore have two feature structures as well as two structure-building features [• Agr •](see (19))-(20) for the configurations 1SG > 3PL and 3PL > 1PL[2] Likewise, two morphological arrays are involved, each of which is associated to a feature structure. A constraint EXHAUST MORPHOLOGICAL ARRAY

---

[2] For ease of representation, the feature structure of the **less salient** argument is listed first and that of the *more salient* argument is listed second in both (i) and (ii) as well as in both (21a)-(21b) and (22a)-(22b).

(EXMORAR) ensures that, once a morphological array has been accessed, all exponents in that array which are compatible with the corresponding feature structure have to be merged *before* the other morphological array can be accessed (see Müller (2020: 141), Andermann (2022: 29)).

The inherent features of feature structures and exponents in Potawatomi are [± 1], [±2], [±3] for person, [± pl(ural)] for number, [± obv] for obviation and [±obj(ect)] for case to account for the distribution of the suffix /nan/. Apart from /nan/ there are no markers in the morphological array whose distribution is sensitive to case / grammatical function.

(19) *Inherent feature structures:*
1SG > 3PL
[$_V$ wapm] : [•Agr•] [•Agr•]
[-1-2+3+pl-obv] [+1-2-3-pl-obv]

(20) *Inherent feature structures:*
3PL > 1PL
[$_V$ wapm] : [•Agr•] [•Agr•]
[-1-2+3+pl-obv] [+1-2-3+pl-obv]

Before Merge takes place, an operation comparable to the Major Reference assignment function in Stump (2001) determines which of the feature structures is **less salient** and which one is *more salient*, based on the salience hierarchy *1/2 > 3 >* OBV that has already been proposed for Algonquian languages (see Trommer (2001) on Menominee, Kushnir (2015) on Plains Cree, Bruening (2017) on Passamaquoddy-Maliseet, and Despić et al. (2019) on Cheyenne) and, in slightly modified versions, also for Potawatomi (see Wunderlich (1997), Stiebels (2002), Henze and Zimmermann (2011)). By this operation, the binary feature [± sal(ient)] is added to the feature structures, i.e. the **less salient** feature structure is assigned the feature [-sal] while the more salient feature structure is assigned [+sal], as exemplified in (19a-b) for the configuration 1SG > 3PL and in (20a-b) for the configuration 3PL > 1PL. In local contexts, both feature structures are [+sal].

(21) 1SG > 3PL*:*

a. *Assign [-sal] and [+sal]*
[$_V$wapm]:[•Agr•][•Agr•]
[-1-2+3+pl-obv-**sal**]
[+1-2-3-pl-obv+*sal*]

b. *Assign [-su] and [+su]*
[$_V$wapm]:[•Agr•][•Agr•]
[-1-2+3+pl-obv-**sal**-*su*]
[+1-2-3-pl-obv+*sal*+*su*]

(22) 3PL > 1PL*:*

a. *Assign [-sal] and [+sal]*
[$_V$wapm]:[•Agr•][•Agr•]
[-1-2+3+pl-obv-**sal**]
[+1-2-3+pl-obv+*sal*]

b. *Assign [-su] and [+su]*
[$_V$wapm]:[•Agr•][•Agr•]
[-1-2+3+pl-obv-**sal**+*su*]
[+1-2-3+pl-obv+*sal*-*su*]

The feature [±sal] then percolates onto the morphological array associated with the feature structure, e.g. if a feature structure is assigned [-sal], then every exponent in the morphological array associated with it is assigned [-sal]. The same holds for [+sal]. The feature [±sal] is discharged after movement triggered by an alignment constraint referring to [±sal], i.e. after the exponent has moved, the feature is not present on it any more.

Via an operation analogous to assigning the binary salience feature, a binary grammatical function feature [±su(bject)] is assigned to the respective feature structures (possibly based on the inherent [±ob(ject)] feature), as shown in (24), and also percolates onto the corresponding morphological arrays.

Unlike the [±sal] feature, however, the [±su] feature is not immediately active but has to be activated by movement, i.e. it is only active after the first exponent has moved. Moreover, there is no movement based on [±su] (as there are no alignment constraints that refer to it), so [± su] is not discharged by movement, but ends up stranded in the base position whenever a) it is active and b) salience-driven movement takes place.

Whenever a [±su] feature is stranded, the DIR/INV markers are inserted; the DIR marker realises stranded [-su] and the INV marker realises [+su]. As repair elements, the DIR and INV markers are not part of the morphological arrays associated with the feature structures but form a separate morphological array.

Merge is assumed to follow the functional sequence of grammatical categories (*f-seq*, see Starke (2001)) that is assumed to hold for both morphology and syntax. By *f-seq*, one might expect that exponents realising the object are merged before exponents realising the subject since objects are lower in the syntactic structure than subjects. However, if one argument is specified for [-sal] and the other one is specified for [+salient], as is the case in Potawatomi, exponents realising the argument specified as [-sal] have to be merged first, as exemplified in (23) for the underlying representation of *n-wapm-a-k* ('I see them').

(23) 1SG > 3PL*: Merge operations*

a. wapm **-w**

b. wapm **-w** -$\mathbf{k_2}$

c. wapm **-w** -$\mathbf{k_2}$ **-wa**

d. wapm **-w** -$\mathbf{k_2}$ **-wa** -$n_1$

### 4.1. Derivation of *N-wapm-a-k* ‘I See Them’

The tableaux in (24) - (35) show the derivation of the form *n-wapm-a-k* (‘I see them’), 1sg > 1pl. Note that, for reasons of space, not all constraints can be listed in all tableaux. Rather, in many tableaux, the only constraints indicated are those that are relevant for the current step in the derivation.

#### *4.1.1. First Cycle: Merge*

Given that the exponents realising the less salient argument are merged first and the ones realising the more salient argument are merged next, and given that all exponents are merged as suffixes, the first cycle where all Merge operations take place is predicted to yield the final output *wapm*-**w**-***k***$_2$-***wa***-*n*$_1$.

(24) *n-wapm-a-k (‘I see them’), Step 1: Merge* **w** *↔ [+3 -su -sal])*

| $I_0$ | wapm [•Agr•] [•Agr•]<br>**[-1 -2 +3 +pl +obj -obv (-su) -sal]**<br>[+1 -2 -3 -pl -obj -obv (+su) +sal]<br>{$_{Agr}$/n/↔[+1 (-su) -sal],<br>$_{Agr}$/$k_1$/↔[+2 (-su) -sal],<br>**$_{Agr}$/w/↔[+3 (-su) -sal ]**,<br>$_{Agr}$ ∅ ↔ [+3 -pl (-su) -sal]<br>$_{Agr}$/mUn/↔[+1 +pl (-su) -sal],<br>$_{Agr}$/nan/$_{[-AGR-]}$↔[+1 +pl +ob (-su) -sal],<br>$_{Agr}$/wa/↔[-1 +pl (-su) -sal],<br>$_{Agr}$/$k_2$/↔[+3 +pl -sal (-su) -sal],<br>$_{Agr}$/$n_2$/↔[ +3 +obv (-su) -sal],...} | MinSat | ExMorAr | Remove Condition | MC(Agr) | IdentFeature | -sal ⇒ R | num ⇒ R | L ⇐ pers | L ⇐ Rt |
|---|---|---|---|---|---|---|---|---|---|---|
| | $O_1$: wapm [•Agr•] | | | | *! | | | | | |
| | $O_2$: wapm-$n_1$ | | | | | *! | | | | |
| | $O_3$: wapm-$k_1$ | | | | | *! | | | | |
| ☞ | $O_4$: wapm-**w** | | | | | | | | | |
| | $O_5$: w-wapm | | | | | | | | | *! |
| | $O_6$: wapm-∅ | | | | | *! | | | | |
| | $O_7$: wapm-mUn | | | | | *! | | | | |
| | $O_8$: wapm-nan | | | | | *! | | | | |
| | $O_9$: wapm-wa | *! | | | | | | | | |
| | $O_{10}$: wapm-$k_2$ | *! | | | | | | | | |
| | $O_{11}$: wapm-$n_2$ | | | | | *! | | | | |

In the first step of the derivation in (24), candidate $O_4$ wins, where w ↔ [+3] is merged as a suffix, satisfying MC-AGR, IDENT-FEATURE, and L ⇐ ROOT. The alignment constraint L ⇐PERS[3] is not violated because of its two-level nature (in the sense of Trommer (2001)); for a person exponent to be able to violate this constraint, the exponent has to be *already present in the input*, which is not the case when the exponent is merged.

(25) *n-wapm-a-k (‘I see them’), Step 2: Merge* ***k₂*** *↔ [+3 +pl -su -sal]*

| $I_4$ | wapm-**w** [• Agr •]<br>**[-1 -2 +3 +pl +obj -obv (-su) -sal]**<br>[+1 -2 -3 -pl -obj -obv (+su) +sal]<br>{$_{\text{Agr}}$/n/↔[+1 (-su) -sal],<br>$_{\text{Agr}}$/k$_1$/↔[+2 (-su) -sal],<br>$_{\text{Agr}}$ ∅ ↔ [+3 -pl (-su) -sal]<br>$_{\text{Agr}}$/mUn/↔[+1 +pl (-su) -sal],<br>$_{\text{Agr}}$/nan/$_{[-\text{AGR}-]}$↔[+1 +pl +ob (+su) -sal],<br>$_{\text{Agr}}$/wa/↔[-1 +pl (-su) -sal],<br>$_{\text{Agr}}$**/k$_2$/↔[+3 +pl -sal (-su) -sal]**,<br>$_{\text{Agr}}$/n$_2$/↔[ +3 +obv (-su) -sal],...} | MINSAT | EXMORAR | REMOVE CONDITION | MC(AGR) | IDENTFEATURE | -SAL ⇒ R | NUM ⇒ R | L ⇐ PERS | L ⇐ RT |
|---|---|---|---|---|---|---|---|---|---|---|
| | $O_{41}$:wapm-w | | *! | | | | | | | |
| | $O_{42}$: wapm-w-n$_1$ | | | | | *! | * | | | |
| | $O_{43}$: wapm-w-k$_1$ | | | | | *! | * | | | |
| | $O_{44}$: wapm-w-∅ | | | | | *! | * | | | |
| | $O_{45}$: wapm-w-mUn | | | | | *! | * | | | |
| | $O_{46}$: wapm-w-nan | | | | | *! | * | | | |
| | $O_{47}$: wapm-w-wa | *! | | | | | * | | | |
| ☞ | $O_{48}$: wapm-**w-k$_2$** | | | | | | * | | | |
| | $O_{49}$: wapm-w-n$_2$ | | | | | *! | * | | | |

In the second step of the derivation in (25), the exponent /k$_2$/ ↔ [+3 +pl] is merged. In theory, both /k$_2$/ ↔ [+3 +pl] and /wa/ ↔ [-1 +pl] are compatible, but a constraint MINIMIZE SATISFACTION (MINSAT) requires that of multiple compatible exponents, the exponent that should always be merged is the one that realises the least amount of "new" features (that are not yet realised by some other exponent) and that therefore incurs the least number of constraint

---

[3] L ⇐PERS should be properly understood as L ⇐ +3. The constraints L ⇐ +2 » ⇐ +1 ⇐ +3 are subsumed under L ⇐ PERS in some tableaux for reasons of space. The same holds for the corresponding MAX constraints.

satisfactions while still improving the constraint profile (hence the constraint‘s name).

(26) *n-wapm-a-k (‘I see them’), Step 3: Merge* ***wa*** *↔ [-1 +pl +su -sal]*

| $I_{48}$ | wapm-w-$k_2$ [• Agr •]<br>**[-1 -2 +3 +pl +obj -obv (-su) -sal]**<br>[+1 -2 -3 -pl -obj -obv (+su) +sal]<br>{$_{Agr}$/$n_1$/↔[+1 (+su) -sal],<br>$_{Agr}$/$k_1$/↔[+2 (-su) -sal],<br>$_{Agr}$ ∅ ↔ [+3 -pl (-su) -sal]<br>$_{Agr}$/mUn/↔[+1 +pl (-su) -sal],<br>$_{Agr}$/nan/$_{[-Agr-]}$↔[+1 +pl +ob (-su) -sal],<br>**$_{Agr}$/wa/↔[-1 +pl (-su) -sal]**,<br>$_{Agr}$/$n_2$/↔[ +3 +obv (-su) -sal],...} | MinSat | ExMorAr | Remove Condition | MC(Agr) | IdentFeature | -sal ⇒ R | num ⇒ R | L ⇐ pers | L ⇐ Rt |
|---|---|---|---|---|---|---|---|---|---|---|
| | $O_{481}$:wapm-w-$k_2$ | | *! | | | | * | | | |
| | $O_{482}$: wapm-w-$k_2$-$n_1$ | | | | | *! | ** | * | | |
| | $O_{483}$: wapm-w-$k_2$-$k_1$ | | | | | *! | ** | * | | |
| | $O_{484}$: wapm-w-$k_2$-∅ | | | | | *! | ** | * | | |
| | $O_{485}$: wapm-w-$k_2$-mUn | | | | | *! | ** | * | | |
| | $O_{486}$: wapm-w-$k_2$-nan | | | | | *! | ** | * | | |
| ☞ | $O_{487}$: wapm-**w-$k_2$-wa** | | | | | | ** | * | | |
| | $O_{488}$: wapm-w-$k_2$-$n_2$ | | | | | *! | ** | * | | |

In (26), the exponent /wa/ ↔ [-1 +pl] is merged. The fact that this exponent does not show up in the final output is accounted for in the second cycle, where both /wa/ and the third person plural marker $k_2$ compete for a position at the right edge, and /wa/, which loses the competition, is deleted due to low-ranked Max(-1) and Max(+pl) constraints.

After Merge of /wa/, the morphological array for the 3pl argument is exhausted as there are no more exponents compatible with the feature structure of that argument. As a consequence, the feature structure of the more salient 1sg argument is accessed, and the first person prefix /$n_1$/ is merged, as shown in (27). As the feature structure of the subject is specified for [+1] and [-pl], neither the 1pl markers /mUn/ and /nan/ nor the 3sg marker /wa/ can be merged, which means the second morphological array is exhausted after Merge of /$n_1$/, and the first morphological cycle is completed.

(27) *n-wapm-a-k (‘I see them’), Step 4: Merge n$_1$ ↔ [+1 -su +sal]*

| I$_{487}$ wapm-w-k$_2$-wa [• Agr •]<br>[-1 -2 +3 +pl +obj -obv (-su) -sal]<br>*[+1 -2 -3 -pl -obj -obv (+su) +sal]*<br>{$_{Agr}$*/n/↔[+1 (+su) +sal],*<br>$_{Agr}$/k$_1$/↔[+2 (+su) +sal],<br>$_{Agr}$/w/↔[+3 (+su) +sal ],<br>$_{Agr}$ ∅ ↔ [+3 -pl (+su) +sal]<br>$_{Agr}$/mUn/↔[+1 +pl (+su) +sal],<br>$_{Agr}$/nan/$_{[-AGR-]}$↔[+1 +pl +ob (+su) +sal],<br>$_{Agr}$/wa/↔[-1 +pl (+su) +sal],<br>$_{Agr}$/k$_2$/↔[+3 +pl (+su) +sal],<br>$_{Agr}$/n$_2$/↔[ +3 +obv (+su) +sal],...} | MINSAT | EXMORAR | MC(AGR) | IDENTFEATURE | -SAL ⇒ R | NUM ⇒ R | L ⇐ PERS | L ⇐ RT |
|---|---|---|---|---|---|---|---|---|
| $O_{4871}$: wapm-w-k$_2$-wa [• Agr •] | | | *! | | ** | | | |
| ☞ $O_{4872}$: wapm-**w-k$_2$-wa**-*n$_1$* | | | | *** | | ** | | |
| $O_{4873}$: wapm-w-k$_2$-wa-k$_1$ | | | | *! | *** | * | *** | |
| $O_{4874}$: wapm-w-k$_2$-wa-w | | | | *! | *** | ** | | |
| $O_{4875}$: wapm-w-k$_2$-wa-∅ | | | | *! | *** | ** | | |
| $O_{4876}$: wapm-w-k$_2$-wa-mUn | | | | *! | *** | ** | | |
| $O_{4877}$: wapm-w-k$_2$-wa-nan | | | | *! | *** | ** | | |
| $O_{4878}$: wapm-w-k$_2$-wa-wa | | | | *! | *** | ** | | |
| $O_{4879}$:wapm-w-k$_2$-wa-k$_2$ | | | | *! | *** | ** | | |
| $O_{48710}$: wapm-w-k$_2$-wa-n$_2$ | | | | *! | *** | ** | | |

The final output of this cycle, *wapm-**w-k$_2$-wa**-n$_1$*, consists of nothing but person and person-number exponents that, as we have seen, are merged in a perfectly regular manner, first the exponents realising the **less salient** argument, then the exponent realising the *more salient* one. It is only in the second cycle that complications arise: exponents move to different positions, the direct marker /a/ is inserted as a reflex of exponent movement, and the markers /w/ and /wa/ are deleted.

### *4.1.2. Second Cycle: Movement*

After all Merge operations have been carried out and the first morphological cycle is terminated, a second cycle takes place in which alignment-driven movement and movement-related repair operations take place. The final output from the first cycle, *wapm*-***w***-***k$_2$***-***wa***-*$n_1$*, serves as input to the second cycle, whose final output is *$n_1$-wapm*-***a***-***k$_2$***.

In the first step of the second-cycle derivation, driven by high-ranked -SAL ⇒ R, the first-merged exponent /w/ moves to the right edge and discharges its non-inherent feature **[-sal]**.

(28) *n-wapm-a-k (‘I see them’), Step 5: Move* ***w*** *right and discharge [-sal]*

| I$_{4872}$ wapm-**w**-**k$_2$**-**wa**-$n_1$ | -SAL ⇒ R | NUM ⇒ R | L⇐ +2 | L⇐ +1 | L⇐ +3 | MAX (±SU) | MAX +2 | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|---|---|---|
| O$_{48721}$ wapm-**w**-**k$_2$**-**wa**-$n_1$ | **! | ** | | * | *** | | | | | | |
| O$_{48722}$ **w**-wapm-**k$_2$**-**wa**-$n_1$ | **! | ** | | * | *** | | | | | | |
| ☞ O$_{48723}$ wapm-**k$_2$**-**wa**-$n_1$-w | ** | ** | | * | *** | | | | | | |

After having discharged [-sal], /w/ moves to the left edge to satisfy the constraint L ⇐ [+3], as represented in (29), and remains there until it is later deleted by the first person exponent /n$_1$/.

(29) *n-wapm-a-k (‘I see them’), Step 6: Move w left*

| I$_{48723}$ wapm-**k$_2$**-**wa**-$n_1$-w | -SAL ⇒ R | NUM ⇒ R | L⇐ +2 | L⇐ +1 | L⇐ +3 | MAX (±SU) | MAX +2 | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|---|---|---|
| O$_{487231}$ wapm-**k$_2$**-**wa**-$n_1$-w | ** | ** | | * | ***! | | | | | | |
| ☞ O$_{487232}$ w-wapm-**k$_2$**-**wa**-$n_1$ | ** | ** | | * | ** | | | | | | |

Crucially, movement of /w/ to the left is only possible at this point because the order of movement operations is determined by the CYCLIC PRINCIPLE in (7), according to which no other exponent can be moved until the exponent currently targeted has reached its final landing site. If the order of movement operations were driven by the ranking of alignment constraints alone, the marker /w/ would be predicted to remain in its suffix position while /k$_2$/ and /wa/ would move past it, as these exponents have not yet discharged their [-sal] feature and-SAL ⇒ R is ranked higher than L ⇐ [+3]. /w/ could not even be deleted by entering into competition with /k$_2$/ and /wa/ for the rightmost position as this competition is triggered by NUM ⇒ R, and /w/ only encodes person but not number. Deleting /w/ to repair the violation of L ⇐ 1 after all other movement steps have been carried out would violate the weakened version of the STRICT CYCLE CONDITION introduced in (6b), which requires deletion to be directly related to the immediately preceding operation (in this case, to result from competition with the previously-moved exponent) and requires any deleted exponent to be either at the left edge, right edge, or adjacent to the element at the left or right edge. The predicted final output would therefore be $*n_1$*-wapm-a-w-*$k_2$.

(30) *n-wapm-a-k ('I see them'), Step 7: Move* $\mathbf{k_2}$ *right, discharge [-sal] and strand [-su] in the base position*

| I$_{487232}$ wapm-$\mathbf{k_2}$-**wa**-$n_1$-*nan*-w | -SAL $\Rightarrow$ R | NUM $\Rightarrow$ R | L $\Leftarrow$ +2 | L $\Leftarrow$ +1 | L $\Leftarrow$ +3 | MAX (±SU) | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|---|
| O$_{4872321}$ w-wapm-$\mathbf{k_2}$-**wa**-$n_1$ | **! | ** | | * | | | | |
| O$_{4872322}$ $\mathbf{k_2}$-w-wapm-[-su]-**wa**-$n_1$ | **! | ** | | * | | * | | |
| ☞ O$_{4872323}$ w-wapm-[-su]-**wa**-$n_1$-k$_2$ | * | * | | * | | * | | |

Now that the exponent /w/ has moved and discharged its **[-sal]** feature, the grammatical function feature [-su] becomes active, i.e. visible for morphology, on the exponent /k$_2$/. As /k$_2$/ moves to the right edge (represented in (31)), it splits off and strands the grammatical function feature [-su], which is required by a constraint MAX(±SU) to be realised by an exponent: in this case, the underspecified object marker /a/ (shown in (32)). As a repair element, this marker is not part of the numeration but is introduced by grammar.

(31) *n-wapm-a-k ('I see them'), Step 8: Insert generic object marker* ***a*** *to repair violation of* MAX *(±su)*

| I$_{4872323}$ w-wapm-[+su]-**wa**-$n_1$-k$_2$ | -SAL $\Rightarrow$ R | NUM $\Rightarrow$ R | L $\Leftarrow$ +2 | L $\Leftarrow$ +1 | L $\Leftarrow$ +3 | MAX (±SU) | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|---|
| O$_{48723231}$ w-wapm-[-su]-**wa**-$n_1$-k$_2$ | * | ** | | * | ** | *! | | |
| ☞ O$_{48723232}$ w-wapm-**a**-**wa**-$n_1$-k$_2$ | * | ** | | * | * | | | |

Again, if the order of movement and movement-related operations were determined by the ranking of alignment and MAX constraints alone, insertion of /a/ would be predicted to take place *after* the other exponents have moved as MAX(±SU) is ranked lower than all alignment constraints. A later insertion of /a/, however, would violate the weakened STRICT CYCLE CONDITION, according to which repair-driven insertion must be a direct consequence of the immediately preceding (movement) operation, i.e. /a/ must be inserted immediately after /$k_2$/ has moved and stranded the feature [-su] in the base position. Crucially, insertion of /a/ is only possible at this point because the order of movement and movement-related operations is *not* determined by the ranking of alignment and MAX constraints alone but, first and foremost, by the CYCLIC PRINCIPLE.

While the insertion of /a/ under this approach is relatively unproblematic in that it respects the weak version of the SCC assumed for insertion, the process of *splitting* the feature [-su] *off* gives rise to a problem: this process must either occur simultaneously with movement of /$k_2$/, in analogy to Obata and Epstein's (2008) feature-splitting internal Merge in syntax, or there must be a designated feature splitting operation that precedes movement of /$k_2$/. The first option involving simultaneous application of feature splitting and movement is not compatible with the principle of Harmonic Serialism according to which only one process may apply in one step. The second option, where feature splitting precedes movement, violates the SCC.

This problem, however, can be solved by assuming what Müller (2023, this volume) refers to as *derivational branching*. In analogy to Müller's (2014) account of resumption (which, unlike the syntactic accounts of overt movement reflexes mentioned above, is not based on the copy theory of movement but on a generative approach to copying involving a designated operation[4]), the feature [±su] is either already split off and realised by the direct or inverse marker after *Merge* of /$k_2$/ or not split off and realised at all.

In the former case, the information that the feature has been split off is registered as a feature on a buffer (a list of movement-related features) on /$k_2$/. This feature is deleted if /$k_2$/ moves to the right edge but causes the derivation to crash, yielding ungrammaticality if /$k_2$/ does not move (which is predicted to lead to ungrammaticality anyway given high-ranked -SAL ⇒ R and the fact that /$k_2$/ is always specified for [-sal]). In the latter case (in which /$k_2$/ does not split [±su] off, which is then realised by a direct/inverse marker), on the

---

[4] For arguments in favour of such a generative approach to copying and against the copy theory of movement, see e.g. Müller (2016).

other hand, the result would be predicted to be grammatical if /k$_2$/ *does not* move. However, this is ruled out by high-ranked -SAL ⇒ R.

Note that in such a branching derivation, the CYCLIC PRINCIPLE would already be relevant at the Merge level, causing /a/ and /UkO/, which belong to the cyclic domain of /k$_2$/, to be inserted before the next exponent (in this case /wa/) is merged. Moreover, insertion of /a/ and /UkO/, if at all, applies immediately after Merge of /k$_2$/. This means that it targets the rightmost position, respecting the strongest version of the SCC and obliterating the need for its weakest version postulated for repair-driven insertion in (6c).

(32) *n-wapm-a-k (‘I see them’), Step 9: Move* ***wa*** *to the right edge and discharge* ***[-sal]***

| I$_{48723232}$ w-wapm-**a**-**wa**-$n_1$-k$_2$ | -SAL ⇒ R | NUM ⇒ R | L⇐ +2 | L⇐ +1 | L⇐ +3 | MAX +SU | MAX +2 | MAX +1 | MAX +3 | MAX -1 |
|---|---|---|---|---|---|---|---|---|---|---|
| O$_{487232321}$ w-wapm-**a**-**wa**-$n_1$-k$_2$ | *! | * | | * | * | | | | | |
| ☞ O$_{487232322}$ w-wapm-**a**-$n_1$-k$_2$-wa | | * | | * | * | | | | | |

Once the generic object marker /a/ is inserted and the cyclic domain of the second exponent /k$_2$/ is completed, the next exponent, /wa/, can move to the right edge and discharge its **[-sal]** feature (see (32)).

(33) *n-wapm-a-k (‘I see them’), Step 10: Resolve competition of* ***wa*** *and* ***k$_2$*** *for the right edge by deleting* ***wa***

| I$_{487232323}$ w-wapm-**a**-$n_1$-k$_2$-wa | -SAL ⇒ R | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX +SU | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|---|
| O$_{487232321}$ w-wapm-**a**-$n_1$-k$_2$-wa | | *! | * | | | | | | |
| O$_{487232322}$ w-wapm-**a**-$n_1$-wa-k$_2$ | | *! | * | | | | | | |
| O$_{487232323}$ w-wapm-**a**-$n_1$-□-wa | | | * | | | | *! | | * |
| ☞ O$_{487232324}$ w-wapm-**a**-$n_1$-k$_2$-□ | | | * | | | | | * | * |

Now both /wa/ and /k$_2$/ compete for the position at the right edge, as they both encode number and the constraint NUM ⇒ R requires number exponents to be right-aligned. Since NUM ⇒ R is ranked higher than all MAX constraints including MAX +PL, deleting one of the exponents improves the constraint profile, and given the ranking Max +3 » Max -1, deletion of /wa/ ↔ [-1 +pl] wins over deletion of /$k_2$/ ↔ [+3 +pl] (see (33)).

(34) *n-wapm-a-k (‘I see them’), Step 11: Move* n$_1$ *to the left edge*

| I$_{487232324}$ w-wapm-**a**-$n_1$-$k_2$ | -SAL ⇒ R | NUM ⇒ R | L⇐ +1 | L⇐ +3 | Max +SU | Max +1 | Max +3 | Max -1 |
|---|---|---|---|---|---|---|---|---|
| O$_{4872323241}$ w-wapm-**a**-$n_1$-$k_2$ | | | *! | | | | | |
| ☞ O$_{4872323242}$ $n_1$-w-wapm-**a**-$k_2$ | | | | * | | | | |

Finally, the marker /$n_1$/ moves to the left edge to satisfy L ⇐ +1 (see (34)). This movement, however, incurs a violation of L ⇐ +3, as /w/ ↔ [+3] is now to the right of /$n_1$/ and therefore not at the left edge any more.

(35) *n-wapm-a-k (‘I see them’), Step 12: Resolve competition of w and* n$_1$ *for the left edge by deleting w*

| I$_{4872323242}$ $n_1$-w-wapm-**a**-$k_2$ | -SAL ⇒ R | NUM ⇒ R | L⇐ +1 | L⇐ +3 | Max +SU | Max +2 | Max +1 | Max +3 | Max -1 |
|---|---|---|---|---|---|---|---|---|---|
| O$_{48723232421}$ $n_1$-w-wapm-**a**-$k_2$ | | | | *! | | | | | |
| O$_{48723232422}$ w-$n_1$-wapm-**a**-$k_2$ | | | *! | | | | | | |
| O$_{48723232423}$ □-w-wapm-**a**-$k_2$ | | | | | | | *! | | |
| ☞ O$_{48723232424}$ $n_1$-□-wapm-**a**-$k_2$ | | | | | | | | * | |

Like /$k_2$/ and /wa/ in (32)-(33), the markers /$n_1$/ and /w/ compete for an edge position, except for it being leftmost one this time. Given the ranking L ⇐ +1 » L ⇐ +3 » Max (+1) » Max (+3), resolving the competition by deleting /w/ ↔ [+3] yields the best constraint profile. In fact, after /w/ is deleted, the derivation converges on the output *$n_1$-wapm-a-$k_2$* (see (35)).

The derivation of *$n_1$-wapm-a-$k_2$* has shown three things. Firstly, by assuming two separate morphological cycles, one for Merge and one for movement, one can see that Potawatomi TA verb forms are perfectly regular as far as Merge operations are concerned, and both unexpected exponence (insertion of the direct marker /a/) and unexpected non-exponence (deletion of /w/ and /wa/) arise only in the second cycle as a consequence of movement. Secondly, for insertion and deletion operations to not be ruled out by the weak versions of the SCC in (6b) and (6c), the order of movement operations has to follow the

CYCLIC PRINCIPLE in (7). And finally, adopting a derivational branching approach strengthens the SCC by removing the need for its weakest version in (6c). The need for the SCC itself is demonstrated in section 4.2.

## 4.2. Derivation of *N-wapm-UkO-nan-k* ‘They See Us’

After having seen how the direct form *n-wapm-a-k* (‘I see them’, 1SG > 3PL) is derived, let us now consider the derivation of the inverse form *n-wapm-UkO-nan-k* (‘They see us’), where a 3PL subject acts on a 1PL object. As shown in (22a) and (22b), it is the subject that bears the feature value **[-sal]** in inverse forms. As the exponent /$k_2$/, this time specified for [+su -sal], moves to the right edge, it strands the feature [+su], which is then overtly realised by the generic subject marker /UkO/.

### *4.2.1. First Cycle: Merge*

In the Merge cycle, first all exponents realising the **less salient** 3PL subject and then all exponents realising the *more salient* 1PL object are merged, yielding the final output of the Merge cycle: *wapm*-**w**-***k***$_2$-***wa***-*n*$_1$-*nan*. Again, a 3PL argument and a first person argument are involved, the first Merge operations are almost identical to those in (24)-(27) except that on every exponent, the feature specifications for [±su] and [± sal] now have opposite feature values. However, after Merge of *n*$_1$ ↔ [+1 *-su +sal*], additional Merge steps are required, as this time, the first person argument is a first person plural object. Recall from section 2 that there are two first person plural markers, the generic /mUn/ and the more specific /nan/, which appears only in object contexts. The high-ranked constraint MINSAT which, of all compatible exponents, requires the most generic one to be merged, predicts that /mUn/ is merged first and only afterwards is Merge of /nan/ possible.

(36) *n-wapm-UkO-nan-k ('they see us'), Step 5:Merge* /mUn/ *↔ [+1 +pl -su +sal]*

| $I_{4872}$ wapm-w-k$_2$-wa-n$_1$<br>[-1 -2 +3 +pl +obj -obv (+su)-sal]<br>$\boxed{\textit{[+1 -2 -3 +pl -obj -obv (-su) +sal]}}$<br>{ $_{Agr}$/k$_1$/↔[+2 (-su) +sal],<br>$_{Agr}$/w/↔[+3 (-su) +sal ],<br>$_{Agr}$ ∅ ↔ [+3 -pl (-su) +sal]<br>$_{Agr}$*/mUn/↔[+1 +pl (-su) +sal],*<br>$_{Agr}$/nan/$_{[-\text{Agr}-]}$↔[+1 +pl +ob (-su) +sal],<br>$_{Agr}$/wa/↔[-1 +pl (-su) +sal],<br>$_{Agr}$/k$_2$/↔[+3 +pl (-su) +sal],<br>$_{Agr}$/n$_2$/↔[ +3 +obv (-su) +sal],...} | MinSat | ExMorAr | Remove Condition | MC(Agr) | IdentFeature | -sal ⇒ R | num ⇒ R | L ⇐ pers | L ⇐ Rt |
|---|---|---|---|---|---|---|---|---|---|
| O$_{48721}$:wapm-w-k$_2$-wa-n$_1$ | | *! | | | | *** | ** | | |
| O$_{48722}$:wapm-w-k$_2$-wa-n$_1$-k$_1$ | | | | | *! | *** | ** | | |
| O$_{48723}$:wapm-wk$_2$-wa-n$_1$-w | | | | | *! | *** | ** | | |
| O$_{48724}$: wapm-w-k$_2$-wa-n$_1$-∅ | | | | | *! | *** | ** | | |
| ☞ O$_{48725}$: wapm-**w-k$_2$-wa**-*n$_1$-mUn* | | | | | | *** | ** | | |
| O$_{48726}$: wapm-w-k$_2$-wa-n$_1$-nan$_{[-\text{Agr}-]}$ | *! | | | | | *** | ** | | |
| O$_{48727}$: wapm-w-k$_2$-wa-n$_1$-wa | | | | | *! | *** | ** | | |
| O$_{48728}$: wapm-w-k$_2$-wa-n$_1$-k$_2$ | | | | | *! | *** | ** | | |
| O$_{48729}$: wapm-w-k$_2$-wa-n$_1$-n$_2$ | | | | | *! | *** | ** | | |

However, in the final output, only /nan/ appears, not /mUn/. Deleting /mUn/ in the 2nd morphological cycle as a result of competition for the position at the right edge of the word is not possible, since when /mUn/ rightfully occurs in 1PL subject contexts, it triggers exponent drop of /k$_2$/ and /n$_2$/ and must therefore be assumed to win the competition for the position at the right edge. Specifying /mUn/ as [+1 +pl -obj] to prevent it from being merged here is also not possible, as it would make wrong predictions for the preterite, where /mUn/ occurs in 1PL > 3 as well as 3 > 1PL contexts.

(37) *n-wapm-UkO-nan-k ('they see us'), Step 6: Merge /nan*$_{[-R-]}$*/ ↔ [+1 +pl +ob -su +sal]*

| | | MINSAT | EXMORAR | REMOVE CONDITION | MC(AGR) | IDENTFEATURE | -SAL ⇒ R | NUM ⇒ R | L ⇐ PERS | L ⇐ RT |
|---|---|---|---|---|---|---|---|---|---|---|
| $I_{48725}$ | wapm-**w**-**k$_2$**-**wa**-*n$_1$-mUn*<br>[-1 -2 +3 +pl +obj -obv (+su)-sal]<br>$\boxed{\textit{[+1 -2 -3 +pl -obj -obv (-su) +sal]}}$<br>{ $_{Agr}$/k$_1$/↔[+2 (-su) +sal],<br>$_{Agr}$/w/↔[+3 (-su) +sal ],<br>$_{Agr}$ ∅ ↔ [+3 -pl (-su) +sal]<br>*$_{Agr}$/nan/$_{[-AGR-]}$↔[+1 +pl +ob (-su) +sal],*<br>$_{Agr}$/wa/↔[-1 +pl (-su) +sal],<br>$_{Agr}$/k$_2$/↔[+3 +pl (-su) +sal],<br>$_{Agr}$/n$_2$/↔[ +3 +obv (-su) +sal],...} | | | | | | | | | |
| | $O_{487251}$:wapm-w-k$_2$-wa-n$_1$-mUn | | *! | | | *! | ** | *** | | |
| | $O_{487252}$:wapm-w-k$_2$-wa-n$_1$-mUn-k$_1$ | | | | | *! | *** | *** | | |
| | $O_{487253}$:wapm-w-k$_2$-wa-n$_1$-mUn-w | | | | | *! | *** | *** | | |
| | $O_{487254}$: wapm-w-k$_2$-wa-n$_1$-mUn-∅ | | | | | *! | *** | *** | | |
| ☞ | $O_{487255}$: wapm-**w**-**k$_2$**-**wa**-*n$_1$-mUn-nan$_{[-AGR-]}$* | | | | | | *** | *** | | |
| | $O_{487256}$: wapm-w-k$_2$-wa-n$_1$-wa | | | | | *! | *** | *** | | |
| | $O_{487257}$: wapm-w-k$_2$-wa-n$_1$-k$_2$ | | | | | *! | *** | *** | | |
| | $O_{487258}$: wapm-w-k$_2$-wa-n$_1$-n$_2$ | | | | | *! | *** | *** | | |

As /mUn/ has to be merged immediately before /nan/, though, one can assume that /nan/ has a remove feature [-AGR-] that removes /mUn/ to satisfy the REMOVE CONDITION (Müller (2020: 168), Andermann (2022: 31)), a constraint that, in analogy to the MERGE CONDITION in (18), requires remove features to participate and be deleted in a remove operation. The possibility of /nan/ removing exponents other than /mUn/ is ruled out by the weakened version of the STRICT CYCLE CONDITION (SCC), according to which deletion need not target the leftmost or rightmost position but an exponent may only remove an exponent that has been merged immediately before, and /nan/ is always merged immediately after /mUn/ (at least in the verbal domain). Crucially, this proves that the SCC, at least in the weakened version proposed in (6b) for deletion operations, is necessary to derive the Potawatomi Transitive Animate paradigm. Merge of /nan/ and removal of /mUn/ are represented in (37)-(38).

(38) *n-wapm-UkO-nan-k (‘they see us’), Step 7: Discharge [–AGR–] on /nan/ by removing mUn*

| I$_{487255}$ wapm-**w**-**k$_2$**-**wa**-*n$_1$*-*mUn*-*nan$_{[-\text{AGR}-]}$*<br>[-1 -2 +3 +pl +obj -obv (+su)-sal]<br>$\boxed{\textit{[+1 -2 -3 +pl -obj -obv (-su) +sal]}}$<br>{ $_{\text{Agr}}$/k$_1$/↔[+2 (-su) +sal],<br>$_{\text{Agr}}$/w/↔[+3 (-su) +sal ],<br>$_{\text{Agr}}$ ∅ ↔ [+3 -pl (-su) +sal]<br>$_{\text{Agr}}$/wa/↔[-1 +pl (-su) +sal],<br>$_{\text{Agr}}$/k$_2$/↔[+3 +pl (-su) +sal],<br>$_{\text{Agr}}$/n$_2$/↔[ +3 +obv (-su) +sal],...} | MINSAT | EXMORAR | REMOVE CONDITION | MC(AGR) | IDENTFEATURE | -SAL ⇒ R | NUM ⇒ R | L ⇐ PERS | L ⇐ RT |
|---|---|---|---|---|---|---|---|---|---|
| O$_{4872551}$:wapm-**w**-**k$_2$**-**wa**-*n$_1$*-*mUn*-*nan$_{[-\text{AGR}-]}$* | | | *! | | | *** | *** | | |
| ☞ O$_{4872552}$:**w**-**k$_2$**-**wa**-*n$_1$*-*nan* | | | | | | *** | ** | | |

### *4.2.2. Second Cycle: Movement*

Now that all compatible exponents have been merged, the first morphological cycle is completed and the second cycle can take place, taking as its input the final output from the first cycle, ***w**-**k$_2$**-**wa**-n$_1$-nan*, and yielding the final output *n$_1$-wapm-***UkO***-*nan-**k$_2$**.

(39) *n-wapm-UkO-nan-k (‘they see us’), Step 8: Move /w/ right and discharge [-sal]*

| I$_{4872552}$ wapm-**w**-**k$_2$**-**wa**-*n$_1$*-*nan* | -SAL ⇒ R | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|
| O$_{48725521}$ wapm-**w**-**k$_2$**-**wa**-*n$_1$*-*nan* | **! | ** | *** | * | ** | | |
| O$_{48725522}$ **w**-wapm-**k$_2$**-**wa**-*n$_1$*-*nan* | **! | ** | *** | ** | ** | | |
| ☞ O$_{48725523}$ wapm-**k$_2$**-**wa**-*n$_1$*-*nan*-w | ** | ** | *** | ** | ** | | |

Recall that movement operations are subject to the same cyclic domains and take place in the same order as Merge operations. The first exponent to move is therefore /w/ ↔ [+3 +*su* **-sal**], which first moves to the right edge to satisfy -SAL ⇒ R, as represented in (39), discharge its **[-sal]** feature, and then moves to the left to satisfy L ⇐ +3, as represented in (40).

(40) *n-wapm-UkO-nan-k (‘they see us’), Step 9: Move /w/ left*

| $I_{48725523}$ wapm-**$k_2$**-**wa**-*$n_1$*-*nan*-w | -SAL ⇒ R | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|
| $O_{487255231}$ wapm-**$k_2$**-**wa**-*$n_1$*-*nan*-w | ** | ** | *** | ** | **! | | |
| ☞ $O_{487255232}$ w-wapm-**$k_2$**-**wa**-*$n_1$*-*nan* | ** | ** | ** | ** | * | | |

As a consequence of the CYCLIC PRINCIPLE in (7), it is only after /w/ has moved to its final landing site that the next exponent, /$k_2$/ can, in its turn, move to the rightmost position, as represented in (41).

(41) *n-wapm-UkO-nan-k (‘they see us’), Step 10: Move /**$k_2$**/ right, discharge [-sal] and strand [+su] in the base position*

| $I_{487255232}$ wapm-**$k_2$**-**wa**-*$n_1$*-*nan*-w | -SAL ⇒ R | RT ⇐ +SAL | NUM ⇒ R | L⇐ PERS | MAX ± SU | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|
| $O_{4872552321}$ w-wapm-**$k_2$**-**wa**-*$n_1$*-*nan* | **! | ** | ** | *** | | | |
| $O_{4872552322}$ **$k_2$**-w-wapm-[+su]-**wa**-*$n_1$*-*nan* | **! | ** | ** | *** | * | | |
| ☞ $O_{4872552323}$ w-wapm-[+su]-**wa**-*$n_1$*-*nan*-$k_2$ | * | ** | ** | *** | * | | |

Again, after the first exponent has discharged its **[-sal]** feature by movement, the grammatical function feature [+su] is activated on the exponent /$k_2$/. As /$k_2$/ moves to the right edge, it strands the grammatical function feature, which must be realised by an exponent in order to satisfy MAX (±SU).

(42) *n-wapm-UkO-nan-k (‘they see us’), Step 11: Insert generic subject marker **UkO** to repair violation of* MAX *(± su)*

| $I_{4872552323}$ wapm-**$k_2$**-**wa**-*$n_1$*-*nan*-w | -SAL ⇒ R | RT ⇐ +SAL | NUM ⇒ R | L⇐ PERS | MAX ± SU | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|
| $O_{48725523231}$ w-wapm-[+su]-**wa**-*$n_1$*-*nan*-$k_2$ | * | ** | ** | *** | *! | | |
| ☞ $O_{48725523232}$ w-wapm-**UkO**-**wa**-*$n_1$*-*nan*-$k_2$ | * | ** | ** | *** | | | |

In this case, since the stranded feature is [+su], it is the underspecified subject marker /UkO/ that is inserted, as shown in (42). Again, in a derivational branching approach, /k$_2$/ either does not split its [+su] feature off at all or already splits it off after Merge of /k$_2$/. In the latter case, [-su] is realised by /UkO/ immediately afterwards to satisfy MAX (± SU) as well as both the SCC and the CYCLIC PRINCIPLE, and the information that [+su] has been split off and /UkO/ has been inserted is registered on a buffer on /k$_2$/, yielding ungrammaticality if /k$_2$/ does not move. This is again independently ruled out by high-ranked -SAL ⇒ R. Likewise, /k$_2$/ not splitting [+su] off must lead to ungrammaticality due to high-ranked -SAL ⇒ R as it could only ever yield a grammatical result if /k$_2$/ did not have to move.

(43) *n-wapm-UkO-nan-k (‘they see us’), Step 12: Move* ***/wa/*** *right and discharge [-sal]*

| I$_{48725523232}$ w-wapm-**UkO-wa**-*n$_1$*-*nan*-k$_2$ | -SAL ⇒ R | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|---|
| O$_{487255232321}$ w-wapm-**UkO-wa**-*n$_1$*-*nan*-k$_2$ | *! | ** | ** | ** | * | | |
| O$_{487255232322}$ **wa**-w-wapm-**UkO**-*n$_1$*-*nan*-k$_2$ | *! | ** | ** | ** | * | | |
| ☞ O$_{487255232323}$ w-wapm-**UkO**-*n$_1$*-*nan*-k$_2$-wa | | ** | ** | ** | * | | |

After all movement and movement-related insertion operations concerning k$_2$ have been carried out, /wa/ moves to the right edge to discharge its **[-sal]** feature (see (43)).

(44) *n-wapm-UkO-nan-k (‘they see us’), Step 13: Resolve competition between /wa/ and /k$_2$/ for the rightmost position by deleting wa*

| I$_{487255232323}$ w-wapm-**UkO**-*n$_1$*-*nan*-k$_2$-wa | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|
| O$_{4872552323231}$ w-wapm-**UkO**-*n$_1$*-*nan*-k$_2$-wa | ** | **! | ** | * | | | | |
| ☞ O$_{4872552323232}$ w-wapm-**UkO**-*n$_1$*-*nan*-k$_2$-□ | * | ** | ** | * | | | * | * |
| O$_{4872552323233}$ w-wapm-**UkO**-*n$_1$*-*nan*-□-wa | ** | * | ** | * | | *! | | * |

Now that the **[-sal]** feature is discharged, /wa/ and /k$_2$ compete for the position at the right edge due to NUM ⇒ R. The exponent /k$_2$/ ↔ [+3 +pl] wins over /wa/ ↔ [-1 +pl] because of the ranking MAX(+3) » MAX(-1), and /wa/ is deleted (see (44)).

(45) *n-wapm-UkO-nan-k (‘they see us’), Step 14: Move /*$\text{n}_1$*/ left*

| I$_{4872552323232}$ w-wapm-**UkO**-*n*$_1$-*nan*-k$_2$ | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|
| O$_{48725523232321}$ w-wapm-**UkO**-*n*$_1$*nan*-k$_2$ | ** | * | **! | * | | | | |
| ☞ O$_{48725523232322}$ *n*$_1$-w-wapm-**UkO**-*nan*-k$_2$ | ** | * | * | ** | | | | |

As all exponents realising the less salient arguments have either reached their final landing site or have been deleted, /$\text{n}_1$/ can now be moved to the left edge, driven by L ⇐ [+1]. Here, /$\text{n}_1$/ competes with /w/, and due to the ranking of MAX constraints, /w/ loses (see (45)).

(46) *n-wapm-UkO-nan-k (‘they see us’), Step 15: Resolve competition between /*$\text{n}_1$*/ and /w/ for the leftmost position by deleting w*

| I$_{48725523232322}$ w-wapm-**UkO**-*n*$_1$-*nan*-k$_2$ | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX +1 | MAX +3 | MAX -1 | MAX +PL |
|---|---|---|---|---|---|---|---|---|
| O$_{487255232323221}$ *n*$_1$-w-wapm-**UkO**-*nan*-k$_2$ | ** | * | * | **! | | | | |
| O$_{487255232323222}$ w-*n*$_1$-wapm-**UkO**-*nan*-k$_2$ | * | * | * | * | | | | |
| O$_{487255232323223}$ □-w-wapm-**UkO**-*nan*-k$_2$ | | * | * | | *! | | | |
| ☞ O$_{487255232323224}$ *n*$_1$-□-wapm-**UkO**-*nan*-k$_2$ | | * | * | | | * | | |

Now the question arises why the first person plural object marker /nan/[5] does not end up being deleted or deleting /$\text{k}_2$/ by trying to move rightwards. The answer is that /nan/ is specified as *[+salient]* and there is a constraint requiring *[+salient]* exponents to be as close to the stem as possible. This constraint, however, is always violated for /nan/ as the INV marker intervenes between it and the stem. Therefore, under a categoric interpretation, moving /nan/ to the right would not make any difference for that constraint, but it would improve the constraint profile w.r.t. the plural alignment constraint. For the constraint RT ⇐ *[+salient]* to prevent /nan/ from moving, this constraint must be gradient(see (47)).

---

[5] The same question applies to marker /wa/ realising the 2PL argument in 2PL ↔ 3PL contexts.

(47) *n-wapm-UkO-nan-k (‘they see us’), Step 16: Convergence (Attention: Gradient interpretation of alignment constraints)*

| $\mathrm{I}_{487255232323224}$ $n_1$-wapm-**UkO**-*nan*-$\mathrm{k}_2$ | RT ⇐ +SAL | NUM ⇒ R | L⇐ +1 | L⇐ +3 | MAX (PERS) | MAX +PL |
|---|---|---|---|---|---|---|
| ☞ $\mathrm{O}_{4872552323232241}$ $n_1$-wapm-**UkO**-*nan*-$\mathrm{k}_2$ | * | *** | **** | ***** | | |
| $\mathrm{O}_{4872552323232242}$ $n_1$-wapm-**UkO**-$\mathrm{k}_2$-*nan* | **! | *** | ***** | **** | | |

The derivation of the inverse form *$n_1$-wapm-UkO-nan-$k_2$* has once more illustrated the movement-based derivation of inverse marking proposed here and shown the need for such a derivation to obey the CYCLIC PRINCIPLE in (7). The main insight to be gained from this derivation, however, is that the SCC as defined in (6) is necessary to correctly predict that the Remove feature on /nan/ does not trigger removal of any exponent other than /mUn/, as shown in (36)-(38).

## 5. Conclusion

I have shown that an analysis of Potawatomi direct and inverse marking as minimally realised overt reflexes of morphological movement, which I have provided evidence for in section 3, obliterates the need for assuming two Voice heads in the syntax, nominative-accusative and absolutive-ergative alignment at the same time, or one exponent encoding both arguments. In a derivational optimality-theoretic approach, such as Harmonic Serialism, morphological movement does not have to be derived via an additional operation type such as local dislocation or metathesis, but follows without further ado from the interaction of MERGE CONDITION, MAX, and alignment constraints, with the exception of deletion in the context of extended exponence, i.e. removal of /mUn/ by /nan/ in section 4.2.1. I have also shown that assuming two morphological cycles, one for Merge operations and one for movement operations, offers new insight into the Potawatomi transitive animate paradigm, namely that it is underlyingly regular and well-behaved: all exponents are merged neatly in a row, first the markers encoding the **less salient** argument, then the markers realising the *more salient* one. All unexpected exponence (direct/inverse marking) or unexpected non-exponence (participant reduction) is a consequence of movement and movement-related repair operations that take place in the second morphological cycle (or, under a derivational branching approach, that are prepared in the first and completed in the second cycle).

My analysis crucially relies on two concepts of cyclicity: the STRICT CYCLE CONDITION (SCC) and the CYCLIC PRINCIPLE. The SCC comes in three degrees of strength listed in (6) and repeated in (48):

(48) a. *Merge* and *movement* may only target the left or the right edge.
b. *Deletion* must target a position adjacent to the left- or rightmost position and must be the consequence of an immediately preceding Merge or movement operation (that has targeted the left or right edge, as per (48a))
c. *Repair-driven insertion* may apply to any position but must be a direct consequence of an immediately preceding Merge or movement operation (that has targeted the left or right edge, as per (48a)).

As I have shown in section 4.2.1, the versions of the SCC in (48a-b) prevent Remove from overapplying; the 1PL.OBJ marker /nan/ bears a generic Remove feature [–R–] and could therefore in theory remove any exponent and not just the generic 1PL marker /mUn/. However, given the version of the SCC in (48b), /nan/ may only remove the exponent adjacent to it. As an effect of (48a)[6] at the point where /nan/ is merged, the only exponent adjacent to /nan/ is the generic 1PL marker /mUn/ that has been merged immediately before. The weakest version of the SCC in (48c) finally ensures that the trace of /$k_2$/ is overtly realised in the base position by the generic object marker /a/ and the generic subject marker /UkO/ immediately after movement of /$k_2$/ to the right edge. The CYCLIC PRINCIPLE, in turn, must hold for movement operations in order for competition-driven deletion and repair-driven insertion operations to not violate the weaker versions of the SCC in (48b-c).

Interestingly, for the movement operations themselves, it is not so much the SCC as the CYCLIC PRINCIPLE which predicts their order. In (28)-(30), for example, moving /$k_2$/ right immediately after rightward movement of /w/, without moving /w/ left first, would be compatible with the SCC but not with the CYCLIC PRINCIPLE. It is but at a later stage, in (35), where such a

---

[6] (48a) here must be taken together with MINSAT, which ensures that in cases of extended exponence, the more generic exponent is merged first, and L $\Leftarrow$ RT, which requires all affixes to be merged as suffixes.

derivation leads to an ungrammatical result since it is impossible to delete /w/ as a result of competition with /$n_1$/ for the left edge unless /w/ is adjacent to the leftmost position. However, as this is not yet clear at the stage where /w/ needs to move left, its movement must be predicted by an independent principle, namely the CYCLIC PRINCIPLE, in order for the derivation to be myopic. For insertion and deletion operations, on the other hand, it is the SCC which is relevant.

Furthermore, as we have seen in sections 4.1.2 and 4.2.2, by adopting a derivational branching approach, the weakest version of the SCC in (48c) may be abandoned, leaving us with the two stronger versions. An attempt to strengthen the SCC even further by deriving deletion in terms of derivational branching (where information about the deleted item is possibly stored on the adjacent item and deleted under adjacency with the next merged or moved item) might be worth considering in the future.

## References

Andermann, Felicitas (2022): Perspectives on Inverse Marking. Master's thesis, Universität Leipzig.

Anderson, Stephen (1992): *A-Morphous Morphology*. Cambridge University Press.

Arnold, M.D (1995): Case, Periphrastic Do and the Loss of Verb Movement in English. PhD thesis, University of Maryland.

Arregi, Karlos and Andrew Nevins (2012): *Morphotactics: Basque auxiliaries and the structure of spellout*. Vol. 86, Springer Science & Business Media.

Bloomfield, Leonard (1946): Algonquian. *In:* H. Hoijer, ed., *Linguistic Structures of Native America*. Viking Fund Publications in Anthropology, p. 85–129.

Bobaljik, Jonathan David (2000): 'The ins and outs of contextual allomorphy', **10**, 35–71.

Bošković, Željko and Jairo Nunes (2007): The Copy Theory of Movement: A view from PF. In: *The Copy Theory of Movement*. John Benjamins, p. 13–74.

Branigan, Phil and Marguerite MacKenzie (2002): How much syntax can you fit into a word? Late Insertion and verbal agreement in Innu-aimûn. In: *Workshop on Structure and Constituency in Languages of the Americas 5*. University of Toronto.

Brown, Dunstan and Andrew Hippisley (2012): *Network Morphology*. Cambridge University Press.

Bruening, Benjamin (2017): *Consolidated Morphology*. University of Delaware.

Chomsky, Noam (1957): *Syntactic structures*. Mouton.

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 (2001): Derivation by Phase. *In:* K. H. M. Kenstowicz, ed., *A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 1–52.

Crysmann, Berthold and Olivier Bonami (2016): 'Variable Morphotactics in Information-Based Morphology', **52**, 311–374.

Despić, Miloje, David Hamilton and Sarah E. Murray (2019): 'A Cyclic and Multiple Agree account. Person/ number marking in Cheyenne', *Natural Language and Linguistic Theory* **37**, 51–89.

Déchaine, Rose-Marie (1999): What Algonquian Morphology is Really Like: Hockett Revisited. In: *MIT Occasional Papers in Linguistics*. Vol. 17, MIT Press, pp. 25–72.

Embick, David and Rolf Noyer (2001): 'Movement operations after syntax', *Linguistic inquiry* **32**(4), 555–595.

Gleim, Daniel, Gereon Müller, Mariia Privizentseva and Sören E. Tebay (2023): 'Reflexes of Exponent Movement in Inflectional Morphology: A Study in Harmonic Serialism', *Natural language & Linguistic Theory* **41**, 103–158.

Goddard, Ives (1969): Delaware Verbal Morphology: A Descriptive and Comparative Study. PhD thesis, Harvard.

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 Massachusetts, pp. 111–176.

Henze, Daniela and Eva Zimmermann (2011): Collateral Feature Discharge. *In:* A. Black and M. Louie, eds, *Proceedings of the Sixteenth Workshop on the Structure and Constituency of the Languages of the Americas (WSCLA 16)*. Vol. 31, p. 74–91.

Hockett, Charles (1948): 'Potawatomi III: The verb complex', **14**(3), 139–149.

Hornstein, Norbert (2000): *Move! A minimalist theory of construal*. Wiley-Blackwell.

Hyman, Larry M (2003): Suffix ordering in Bantu: A morphocentric account. In Geert Booij and Jaapvan Marle. In: *Yearbook of morphology*. Kluwer, p. 245–281.

Hyman, Larry M. and Sam Mchombo (1992): 'Morphotactic constraints in the Chichewa verb stem', **18**, 350–364.

Kayne, Richard S (1991): 'Romance clitics, verb movement, and PRO', *Linguistic inquiry* **22**(4), 647–686.

Kushnir, Yuriy (2015): Inverse marking in Plains Cree. Bachelor's thesis, Universität Leipzig.

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

Müller, Gereon (2016): Predicate doubling by phonological copying. Replicative Processes in Grammar. In: *Linguistische Arbeitsberichte*. Universität Leipzig.

Müller, Gereon (2020): *Inflectional Morphology in Harmonic Serialism*. Equinox.

Müller, Gereon (2023): Challenges for Cyclicity. In: *Linguistische Arbeitsberichte*. Vol. 95, Universität Leipzig.

Nunes, Jairo (2004): *Linearizations of Chains and Sideward Movement*. MIT Press.

Obata, Miki and Samuel David Epstein (2008): Deducing improper movement from phase-based C-to-T phi transfer: Feature-splitting internal merge. In: *Proceedings of the 27th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project Somerville, MA, pp. 353–360.

Oxford, Will (2018): 'Inverse marking and Multiple Agree in Algonquin: Complementarity and variability', **37**, 955–996.

Oxford, Will (2022): 'Probe specification and agreement variation: Evidence from the Algonquian inverse', https://ling.auf.net/lingbuzz/006486.

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

Pesetsky, David (1989): 'Language-particular processes and the Earliness Principle', *ms., MIT* .

Pesetsky, David (1998): Some optimality principles of sentence pronunciation. In: *Is the Best Good Enough? Optimality and Competition in Syntax*. MIT Working Papers in Linguistics, p. 337–383.

Pesetsky, David and Esther Torrego (2001): 'T-to-C movement: Causes and consequences', *Current Studies in Linguistics Series* **36**, 355–426.

Preminger, Omer (2018): How to tell a syntactic phenomenon when you see it. In: *Talk given at Groningen Syntax Workshop*. Vol. 17.

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

Starke, Michal (2001): Move Dissolves Into Merge. PhD thesis, University of Geneva.

Steele, Susan (1995): 'Towards a Theory of Morphological Information', **2**(71), 260–309.

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

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

Trommer, Jochen (2001): Distributed Optimality. PhD thesis, University of Potsdam.

Trommer, Jochen (2003): Participant Reduction and Two-Level Markedness. In: *Variation within Optimality Theory. Proceedings of the Stockholm Workshop*. .

Trommer, Jochen (2006): 'Direction marking and case in Menominee', *Case, Valency and Transitivity* pp. 91–111.

Wunderlich, Dieter (1997): A Minimalist Model of Inflectional Morphology. In: *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, p. 267–298.

# Cyclicity and Extended Exponence

Jelena Grofulović & Gereon Müller*

## Abstract

The main goal of this paper is to derive, in a principled way, the Partially Superfluous Extended Exponence Generalization, according to which the more general one of two morphological exponents whose specifications for morpho-syntactic features are in a subset relation must always precede the more specific one at the base level of morphological organization. This issue has been addressed in optimality-theoretic approaches to morphology, where it has been argued that an account of the generalization requires a stratal or derivational approach to optimization (Caballero & Inkelas (2013), Stiebels (2015), Müller (2020)). In the present paper, we show that the generalization can also be derived without further ado in a Distributed Morphology approach, given that extended exponence requires feature copying (enrichment), and morphological realization obeys cyclicity: At the point where the derivation in which the more general exponent comes second could generate the required copy of a feature without violating cyclicity, the feature is already gone, due to prior insertion of the more specific exponent.

## 1. The Phenomenon

The concept of extended (or multiple) exponence as an issue in morphology goes back to Matthews (1972, 1974). Extended exponence refers to cases of morphological realization where a single morpho-syntactic feature seems to be expressed by more than one exponent. Thus, number is realized twice in the Archi nouns in (1ad), with a plural exponent (*um* or *or*) that directly follows the stem accompanied by an ergative case exponent (*čaj*) that is also specified for plural (as shown in (1be), the singular number realization is *li*

---

*For helpful comments and discussion, we would like to thank Artemis Alexiadou, Felicitas Andermann, Peter Arkadiev, Paula Fenger, Johannes Hein, Maria Kouneli, Elango Kumaran, Masha Privizentseva, Barbara Stiebels, Philipp Weisser, and audiences at Universität Leipzig and the *Multiple Exponence* workshop at ZAS Berlin. Research for this article was supported by a DFG grant for the Research Training Group *Interaction of Grammatical Building Blocks* (GRK 2011) and by a DFG grant for the project *Prospects of Inflectional Morphology in Harmonic Serialism* as part of the Research Unit *Cyclic Optimization* (FOR 5175).

*Cyclicity*, 159–198
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

or *i*). Furthermore, extended exponence of number is obligatory here; even though *čaj* realizes both plural and ergative case, the pure plural marker cannot be left out (cf. (1cf)). Exactly the same pattern shows up with number and case exponents in dative plural inflections of German nouns in (2): In (2ad), a pure plural exponent (*er* or *e*) is followed by a dative plural exponent *n* (the singular has a different realization as *Ø* or, in slightly archaic style, *e*; see (2cf)); this does not render the pure plural exponent superfluous (cf. (2be)).

(1) *Number in Archi Nouns*

a. gel-um-čaj
cup-PL-ERG.PL
b. gel-li
cup.SG-ERG
c. *gel-čaj
cup.SG-ERG.PL
d. qIinn-or-čaj
bridge-PL-ERG.PL
e. qIonn-i
bridge.SG-ERG
f. *qIonn-čaj
bridge.SG-ERG.PL

(2) *Number in German Nouns*:

a. Kind-er-n
child-PL-DAT.PL
b. *Kind-n
child.SG-DAT.PL
c. Kind-(e)
child.SG-DAT
d. Tisch-e-n
table-PL-DAT.PL
e. *Tisch-n
table.SG-DAT.PL
f. Tisch-(e)
table.SG-DAT

There are different types of extended exponence. A taxonomy going back to Caballero and Harris (2012) distinguishes between partially superfluous extended exponence, overlapping extended exponence, and fully superfluous extended exponence. Two morphological exponents /a/, /b/ co-occuring in a word exhibit partially superfluous extended exponence when their feature specifications are in a subset relation, as schematically depicted in (3).

(3) *Partially superfluous extended exponence*:
The feature specifications associated with two exponents are in a proper subset relation.

a. /a/ $\leftrightarrow [f_1]$
b. /b/ $\leftrightarrow [f_1,f_2]$

This is the scenario that shows up with multiple exponence of number in Archi (see (1)) and German (see (2)). Another example of such a pattern is instantiated by Meskwaki person agreement marking on verbs; see (4ab) (cf.

Dahlstrom (2000)). As observed by Caballero and Harris (2012), the subject person information is provided twice in the verb – both by the more general exponent (here the prefixes *ne* and *ke*), and by the suffixes (*pena* and *pwa*), which are more specific as they also encode subject number information in addition.

(4) *Person in Meskwaki*:

a. ne-nowi:-pena
1-go.out-1.PL
'We (excl.) go out'
b. ke-nowi:-pwa
2-go.out-2.PL
'You (pl.) go out'

A fourth and final example of partially superfluous extended exponence comes from Mari (see Alhoniemi (1993)). In Mari, the standard marker for second person singular in all tense/mood-combinations is *t*; see (5a). In past contexts, in which a pure tense exponent *šə̂* shows up, the second person singular exponent *t* is replaced by *č́*; see (5b). The restriction to past contexts (cf. (5c)) implies that *č́* is specified for both person/number and tense.

(5) *Tense in Mari*:

a. kole-t
die-2SG
'You die'
b. kolə̂-šə̂-č́
die-PST-2SG.PST
'You died'
c. *kole-č́
die-2SG.PST
'You die'

Next, with overlapping extended exponence, two exponents in a word share a morpho-syntactic feature, but they also each have some morpho-syntactic feature that is not shared. A schematic illustration is given in (6); here, both abstract exponents realize a separate piece of morpho-syntactic information on the verb (viz., features [$f_2$] and [$f_3$], respectively).

(6) *Overlapping extended exponence*:
Two exponents share some morpho-syntactic feature, but their morpho-syntactic features are not in a subset relation.

a. /a/ ↔ [$f_1$,$f_2$]
b. /b/ ↔ [$f_1$,$f_3$]

One of the examples discussed by Caballero and Harris (2012) comes from Filomeno Mata Totonaco (cf. Inkelas et al. (2006), McFarland (2009)). As can be seen from (7b), in second person singular progressive contexts, Totonacan morphology realizes second person subject information on the stem, on the progressive marker, and on the number marker; thus, there is a massive overlap of second person exponence. Still, none of the feature sets associated with the three morphological exponents is a subset of the feature set of any other exponent.

(7) *Person in Totonacan*:

a. min-maa
come-PROG
'he is coming'
b. tan-paa-ti
come.2SUBJ-PROG.2SUBJ-2SUBJ.SG
'you are coming'
c. *min-maa-ti
come-PROG-2SUBJ.SG

Finally, Caballero and Harris (2012) recognize fully superfluous extended exponence as a third pattern. As shown in (8), here the sets of morpho-syntactic features associated with two (or more) exponents in a word are identical.

(8) *Fully superfluous extended exponence*:
Two exponents have identical feature specifications.

a. /a/ ↔ [$f_1$,$f_2$]
b. /b/ ↔ [$f_1$,$f_2$]

However, for many of the relevant examples, closer inspection reveals that they do not instantiate extended exponence after all, e.g., because what at first sight looks like two separate exponents actually qualifies as a single discontinuous exponent, or because the features that are involved are not in fact identical, or because copying (of the whole exponent, i.e., including the form) is involved (see Stiebels (2015, 2016) and Müller (2020) for some case studies).

Causative formation in Sinhala (see Fenger and Weisser (2023)) might be an instance of discontinuous exponence. (9a) shows the base form of a verb. In (9b), causativization has applied, and a causative exponent shows up. With a verb of this type, the causative exponent takes the form of a zero item that triggers gemination of the stem-final consonant (plus schwa epenthesis), leading to *də*. However, in addition, there is also a second verb class that handles causative formation differently, viz., by adding a causative affix *wə*. Importantly, this segmental exponent can also be added to the non-segmental exponent of the first verb class without a change in meaning; see (9c). This might then instantiate a scenario of the type in (8). Alternatively, and this is the analysis that we would like to adopt here, the (optional) co-occurence of a segmental and a suprasegmental marking in (9c) can be understood in the same way as, say, plural markers like *e* and *er* in German (see (2) above), which may in addition trigger Umlaut on a preceding syllable (cf., e.g., *Schaf-e* vs. **Schä-f-e* ('sheep-pl') and **Hand-e* vs. *Händ-e* ('hand-pl')). In both the Sinhala and the German case, the most straightforward analysis would presumably postulate a single segmental exponent that can or must be accompanied by abstract supra-segmental information that subsequently phonologically modifies the stem (see, e.g., Wiese (2000) and Trommer (2011)).

(9) *Causative in Sinhala*:

a. adi-nə-wa
pull-NPST-IND
'pull'

b. ad-də-nə-wa
pull-CAUS-NPST-IND
'make somebody pull'

c. ad-də-wə-nə-wa
pull-CAUS-CAUS-NPST-IND
'make somebody pull'

Assuming that this result can be generalized, we will postulate, here and henceforth, that fully superfluous extended exponence does not in fact exist. It can also be noted that such a phenomenon would be unexpected under many theories of inflectional morphology; and the approach that we will develop also has this property. This leaves overlapping extended exponence and partially superfluous extended exponence as explananda for morphological theories. While overlapping extended exponence turns out to be unproblematic under most theories of inflectional morphology, the case is different with partially superfluous extended exponence, which raises problems for various restrictive theories of inflectional morphology: Essentially, the question is why the availability (and presence) of a more specific exponent like /b/ $\leftrightarrow [f_1,f_2]$ does not block a more general exponent /a/ $\leftarrow [f_1]$; given the availability of /b/, the occurrence of /a/ looks redundant, and might be expected to be blocked for this reason. It is the primary goal of the present paper to give a principled answer to this question on the basis of Distributed Morphology (cf. Halle and Marantz (1993)).

We will proceed as follows. In section 2, we discuss the nature of this problem, and some solutions that have been advanced, against the background of a morphological theory in which it shows up in a particularly obvious way, viz., Optimality Theory. All existing solutions have a common core, which centers around what we call the Partially Superfluous Extended Exponence Generalization: The more general exponent must be closer to the stem than the more specific exponent. In section 3, we then turn to Distributed Morphology. We show that, as it stands, the Partially Superfluous Extended Exponence Generalization cannot yet be derived under any existing approach based on Distributed Morphology; in fact, the only such approach that might have anything to say about the phenomenon (viz., Bobaljik (2000)) turns out to make predictions that are diametrically opposed to the ones covered by the generalization. However, we show that by clarifying the nature of context features, by treating every instance of extended exponence as a consequence of a post-syntactic feature copy operation (‘enrichment’; Müller (2007)), and, most importantly, by invoking cyclicity (the Cyclic Principle and the Strict Cycle Condition), the generalization can be derived. Finally, section 4 discusses some empirical challenges to the generalization and the Distributed Morphology account from which it follows.

## 2. Optimality Theory

Optimality Theory (cf. Prince and Smolensky (2004)) highlights the general problem with partially superfluous extended exponence because economy of representation is straightforwardly derived in this approach: Every grammatical operation automatically incurs some violation (e.g., of a faithfulness constraint), and this implies that, ceteris paribus, if all relevant constraints can be satisfied without this operation, it will be precluded. Hence, in a partially superfluous extended exponence scenario like (3), the more general exponent /a/ should always be blocked by the more specific exponent /b/: The presence of /a/ is per se costly; it invariably violates some (low-ranked) constraint. Therefore, it seems that /a/ has nothing to contribute that could not be obtained with /b/ alone: The constraint profile of a candidate with /a/ and /b/ must be worse than the constraint profile of a candidate with just /b/.

This reasoning is illustrated for the competition underlying extended exponence of number with Archi nouns (recall (1)) in the tableau in (10). For the purposes of the present discussion, let us make the following assumptions about optimality-theoretic morphology:[1] First, suppose that the input for morphological exponence is a stem, together with a fully specified set of morpho-syntactic features that need to be realized by exponents – in the case at hand, $I_1$ has a [+pl] number feature and a case specification [–obl(ique),+gov(erned)] that represents the ergative. Second, the competing output candidates $O_{11}$, $O_{12}$, etc., have carried out morphological realization to different degrees, and with different exponents.[2] Third, faithfulness constraints derive the compatibility and specificity requirements that are stipulated as parts of a constraint like the Subset Principle (cf. Halle (1997)) or Panini's Principle (cf. Stump (2001)) in other morphological theories, like Distributed Morphology or Paradigm Function Morphology. More specifically, ID(ENT)-F(EATURE) ensures compatbility (i.e., exponents have feature specifications that are subsets of the target specification of the input), and MAXNUM and MAXCASE demand

---

[1] What follows is an amalgamation and simplification of various different optimality-theoretic approaches to morphology; see, e.g., Grimshaw (2001), Trommer (2001), Don and Blom (2006), Ortmann (2004), and Stiebels (2006), among many others.

[2] I.e., a *realizational*, rather than *incremental*, approach to morphology is adopted; cf. Stump (2001). In fact, the problems with licensing partially superfluous extended exponence are exacerbated in an incremental approach, according to which there are no morpho-syntactic features in a word except for those contributed by morphological exponents; see, e.g., Wunderlich (1997).

realization of number and case features of the input by output exponents. Finally, a low-ranked *Struc(ture) is here assumed to stand for whatever constraints ensure that no grammatical operation comes for free (including, of course, morphological exponence), and that thus derive the general economy effect in optimality theory. On this basis, the competition in (10) makes it clear that the intended winner (i.e., $O_{14}$, which exhibits partially superfluous extended exponence) has no chance to ever become optimal (signalled here by ☆); it will always be blocked by a more economical output candidate (viz., $O_{11}$, which dispenses with the gratuitous more general exponent that realizes number but not case); the wrong winner is indicated by ☛ here.

(10) *Extended exponence as a problem* (standard parallel optimality theory):

| $I_1$: /[$_N$ gel:[+pl,–obl,+gov]]/ | Id-F | MaxNum | MaxCase | *Struc |
|---|---|---|---|---|
| ☛$O_{11}$: gel[+pl,–obl,+gov]-čaj[+pl,–obl,+gov] | | | | * |
| $O_{12}$: gel[+pl,–obl,+gov]-um[+pl] | | | *!* | * |
| $O_{13}$: gel[+pl,–obl,+gov]-li[–obl,+gov] | | *! | | * |
| ☆$O_{14}$: gel[+pl,–obl,+gov]-um[+pl]-čaj[+pl,–obl,+gov] | | | | **! |

To the best of our knowledge, there is no solution to this problem in standard parallel optimality theory, as devised in Prince and Smolensky (2004). However, various solutions have been proposed that rely on versions of optimality theory that either invoke strata, or that are inherently derivational; stratal and derivational approaches have in common that they presuppose that grammatical operations (like morphological exponence, in the case at hand) can be ordered with respect to one another, such that an operation can become *opaque* (cf. Kiparsky (1973)), in the sense that it would be bled by another operation but is not factually bled because that other operation applies too late (i.e., counter-bleeding takes place).

More specifically, the solutions to the problem in (10) that have been suggested in Caballero and Inkelas (2013), Stiebels (2015), and Müller (2020) all take the following form: Assuming that there are several optimization procedures, which are organized sequentially, these approaches converge on the assumption that for the first optimization procedure in the scenario in (3), only /a/ ↔ [$f_1$] is available, satisfying the constraint demanding realization of [$f_1$] (but not the constraint demanding realization of [$f_2$]). Subsequently, /b/ ↔ [$f_1$,$f_2$] becomes available, and is selected so as to also satisfy the constraint demanding realization of [$f_2$]. The systems are *myopic*: An earlier selection of

/b/ would have made selection of /a/ impossible (bleeding), but since /b/ is not initially available, selection of /a/ is *counter-bled* by subsequent selection of /b/.

While these stratal/derivational approaches all share a common core, the concrete implementations differ substantially; most importantly, the answer given to the question of *why* a more specific /b/ $\leftrightarrow$ [$f_1$,$f_2$] is not initially available, so that less specific /a/ $\leftrightarrow$ [$f_1$] can become optimal at an early stage, is addressed in diverging ways. To begin with, the approach proposed by Caballero and Inkelas (2013) presupposes *strata* (see Kiparsky (1982)): On this view, /a/ belongs to stratum 1, /b/ belongs to stratum 2, and optimization in stratum one (where /b/ is not yet available) precedes optimization in stratum 2. Second, Stiebels' (2015) analysis makes use of *f-seq* (see Wunderlich (1997), Starke (2001)): The order of exponent selection follows the functional sequence of grammatical categories. If [$f_2$] outranks [$f_1$] on f-seq, the exponent /b/ that (also) realizes [$f_2$] must come after the exponent /a/ that (only) realizes [$f_1$]. Finally, the approach developed in Müller (2020) relies on a constraint *Minimize Satisfaction* that is independently designed to capture effects like those covered by Chomsky's (2001) Merge over Move constraint. *Minimize Satisfaction* is an overarching, inviolable constraint demanding (non-zero) minimization of new constraint satisfactions in a derivational version of optimality theory (*harmonic serialism*; cf. McCarthy (2016)). In this approach to inflectional morphology, partially superfluous extended exponence is possible since the more general exponent /a/ (yielding fewer new constraint satisfactions) is optimal at an early stage but selection of more specific /b/ is both required and permitted at a later stage (because it improves the constraint profile, and there is no exponent left that would do so with fewer new constraint satisfactions at this point).

All these optimality-theoretic approaches relying on derivational order derive the generalization in (11).

(11) *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/.

Based on a preliminary investigation of the typological record, Stiebels (2015) ventures the hypothesis that this prediction is corroborated by the empirical evidence in the world's languages. In what follows, we will postulate that this is indeed the case (we will address some pieces of apparent counter-evidence in section 4). Given this state of affairs, the question arises of whether the Partially Superfluous Extended Exponence Generalization can also be derived in other, non-optimality-theoretic approaches to morphology. The prospects would seem to be bleak for any theory that is (a) non-derivational, and that (b) intrinsically permits unlimited feature realization by multiple exponents, like Paradigm Function Morphology (see Stump (2001)) or Network Morphology (see Brown and Hippisley (2012)). However, things might be different with Distributed Morphology, which is derivational in nature, and which associates each morpho-syntactic feature with a designated functional head.

## 3. Distributed Morphology

### 3.1. State of the Art

Can the Partially Superfluous Extended Exponence Generalization in (11) also be derived in Distributed Morphology? As a first step towards an answer, it can be noted that the phenomenon of extended exponence is typically addressed by recourse to *secondary, contextual features* in Distributed Morphology. Such contextual features are associated with exponents just like primary, “core” features, but in contrast to the latter, they are usually put in brackets. Thus, instead of the two exponents /a/ $\leftrightarrow$ [$f_1$] and /b/ $\leftrightarrow$ [$f_1$,$f_2$] in a partially superfluous extended exponence scenario (cf. (3)), we get /a/$'$ $\leftrightarrow$ [$f_1$] (as before) and /b/$'$ $\leftrightarrow$ [$f_2$] ([$f_1$]), where [$f_2$] counts as primary and ([$f_1$]) counts as secondary. Furthermore, a contextual feature of an exponent like /b/$'$ is not matched by the functional head X into which /b/$'$ is inserted; rather, the matching [$f_2$] feature is located on some other functional head Y in the vicinity of X. Thus, on this view, extended exponence qualifies as *contextual allomorphy*.

For this reason, one might expect that restrictions for contextual allomorphy as they have been proposed in Distributed Morphology have some bearing on patterns of extended exponence. There are two relevant concepts. The first one is that of a *morphological phase* (see Marvin (2002), Embick (2010), and Bermúdez-Otero (2011), among others). Morphological phases act as locality domains for morphological realization, and thus ensure that a secondary, contextual feature of an exponent must be matched within this domain. However, irrespective of the exact definition of morphological phases, it seems clear that due to their size, they cannot systematically restrict patterns of partially superfluous extended exponence in an interesting way, let alone derive the Partially Superfluous Extended Exponence Generalization. The second relevant concept that has been argued to govern morphological exponence is that of *cyclicity*. And indeed, as argued by Bobaljik (2000), subjecting morphological exponence to cyclicity potentially can successfully restrict extended exponence.

Bobaljik's (2000) approach to contextual allomorphy rests on three basic assumptions. First, there is what he calls *separation*: Morphology interprets syntactic structures, rather than feeding them; i.e., a realizational approach is adopted (see footnote 2), as it is standardly assumed in Distributed Morphology. Second, the operation of vocabulary insertion that brings about morphological realization in Distributed Morphology implies feature *discharge* (or 'rewriting', in Bobaljik's terminology) in the target functional head (cf. Noyer (1997) and Trommer (1999)): Matched features are used up by vocabulary insertion and no longer a part of the representation. Third, morphological realization is subject to *cyclicity*, in the sense that vocabulary insertion proceeds root-outwards.

This system makes the following two predictions, one for morpho-syntactic features, and a contrary one for morpho-phonological diacritic features. First, outwards-sensitivity to morpho-syntactic features on functional heads in a complex word is possible because the heads hosting those features have not yet been subject to vocabulary insertion (and, hence, feature discharge); however, inwards-sensitivity to such morpho-syntactic features is not possible because, given cyclicity, they have already been discharged as a consequence of earlier vocabulary insertion. Second, the inwards-sensitivity to morpho-phonological diacritic features (like inflection class) on vocabulary items in a complex word is possible (because these features, by assumption, were originally brought into the structure by the root vocabulary item, in minimal violation of a the tenets of a realizational approach); in contrast, outwards-sensitivity to such diacritics is not possible (because the items that introduce them into the structure are not yet present, given cyclicity).

Unfortunately, this approach does not derive the Partially Superfluous Extended Exponence Generalization in (11); in fact, it predicts more or less the opposite of what is covered by it: Given the generalization, it should be the case that the outer (more specific) exponent /b/′ $\leftrightarrow$ [f$_2$] ([f$_1$]) requires the contextual presence of the morpho-syntactic feature [f$_1$] matched by the inner (more general) exponent /a/′ $\leftrightarrow$ [f$_1$]; however, once /a/′ has been inserted into a head X, $[f_1]$ is gone from X, and subsequent insertion of $/b/' \leftrightarrow [f_2]$ $([f_1])$ in an outer head Y will be impossible because the latter exponent does not find the contextual feature $([f_1])$ in X anymore that it needs to satisfy the compatibility ('subset') requirement of the Subset Principle. Consequently, a form like *Kind-er-n* ('child-PL-DAT.PL') (or any other instance of partially superfluous extended exponence in the above examples) can never be generated under these assumptions: The exponent /er/ ↔ [+pl] is inserted first (given cyclicity) into a functional head X (which one may consider a number head #), thereby discharging (and removing) the number feature [+pl] from X, and subsequent insertion of /n/ ↔ [+obj,+obl] ([+pl]) into a higher functional head Y (which we may identify as K, for case, with [+obj,+obl] standing for dative) will fail because ([+pl]) does not find a matching feature in the syntactic representation anymore.

This problem can in principle be solved by postulating that morpho-syntactic features that are discharged by exponent insertion into a functional head are *deleted* (and thus not accessible for direct insertion of another vocabulary item into the same head), but not fully *erased* (see Chomsky (1995)), so that subsequent reference by more specific exponents is still permitted. Such a modification of the concept of discharge would imply that all existing cases of extended exponence can be covered, but since restrictions on exponence are now weakened, it does not come as a surprise that it would not get us any closer to deriving the Partially Superfluous Extended Exponence Generalization in (11). In fact, the resulting approach would be hardly distinguishable from one where all features are available everywhere (modulo phases), all the time, as in Paradigm Function Morphology or Network Morphology (see above). The only remaining restriction would be that outwards-sensitivity to morpho-phonological diacritic features would still predicted not to be possible. In view of this, we take it that there is every reason to develop a new approach to extended exponence in Distributed Morphology that captures the generalization in (11). We lay out such an approach in the next section.

### 3.2. A New Approach

The new approach relies on six assumptions. The first assumption concerns *disjunctive blocking*, and it is a standard one made in Distributed Morphology (see, e.g., Halle and Marantz (1993)): Only one vocabulary item can be inserted into a given functional head. As a consequence, two morphological exponents that participate in extended exponence in a word must have been inserted into two separate functional heads.

The second assumption is shared with Bobaljik's approach: Morphological exponence involves *discharge* (see Noyer (1997) and Trommer (1999)): The insertion of an exponent with a matching feature discharges, and thereby deletes, a feature in the locally accessible domain (the morphological phase).

We refer to the third assumption as *feature uniqueness*: There is no distinction between "primary" and "secondary/contextual" features on morphological exponents; all morpho-syntactic features of a vocabulary item are of the same kind. It follows from this assumption that the specific exponents in a partially superfluous extended exponence scenario must look as in (12a), and cannot take a form like the one in (12b).

(12) a. $\surd$/b/ $\leftrightarrow$ $[f_1,f_2]$
b. */b/$'$ $\leftrightarrow$ $[f_2]([f_1])$

The conclusion that there can be no meaningful ontological differences among the morpho-syntactic features characterizing a morphological exponent has been most forcefully defended in Stump (2001, ch. 5) (also cf. Müller (2020, ch 3.)). Problems with contextual features pointed out by Stump include *ambiguity* (How can it be that one and the same exponent may qualify as a primary exponent of a given morpho-syntactic property in one environment, and as a secondary exponent of the same morpho-syntactic property in another environment?); *learnability* (How can a child acquiring a language decide whether a given feature on some vocabulary item is a primary or a secondary (contextual) one?); *specificity* (To what extent do contextual features count for the specificity of exponents that bear them?);[3] and *locality* (How far away can a contextual feature on a functional head be located from the exponent that shares it?). We take the first two of these problems to be decisive.

Fourth, we assume that extended exponence never comes for free; it must always be brought about by *enrichment* (see Müller (2007)). Enrichment is a post-syntactic feature copying operation that is a mirror image of post-syntactic impoverishment; enrichment generates *local copies* of morpho-syntactic features on a functional head, which can then give rise to extended exponence

---

[3] However, see Arregi and Nevins (2012) and Hanink (2018) for some suggestions.

because there are now two (or more, if enrichment applies more than once) identical features that can be separately subject to discharge.[4]

Fifth, these four assumptions require a clarification of the core concepts of *compatibility and specificity* governing vocabulary insertion. A modified Subset Principle (cf. Halle (1997)) is called for that is made sensitive to contextual features in syntactic representations, i.e., features of the syntactic contexts which are accessed by morphological exponents that bear them, but that are not located in the functional head into which the exponent is inserted. We adopt the following version of the Subset Principle.

(13) *Subset Principle*:
A vocabulary item *V* is inserted into a functional morpheme *M* contained in a morphological phase *P* (thereby discharging all matching features in *P*) iff (a) and (b) hold:

a. *Compatibility*:
*V* realizes a feature of *M*, and the morpho-syntactic features of *V* are a subset of the morpho-syntactic features of *P*.

b. *Specificity*:
Among the vocabulary items that satisfy (a), there is no $V'$ that realizes more features of *M* than *V*.

Thus, an exponent can only be inserted into a given head if it shares a feature with it, and if all other features that the exponent may be equipped with are available for insertion (and, consequently, discharge) on some head (which may or may not be the same head) in the local domain (the morphological phase). If there is more than one exponent that satisfies this compatibility requirement, specificity selects the one(s) that realize(s) most features in the head into which insertion takes place. Note that it is only the features of this head *M* that play a role for specificity, and *not* the features in the syntactic context *P*. As we will see, this may in principle lead to situations where more than one vocabulary item could be inserted in accordance with the Subset Principle; and this not only if the feature specifications two exponents are identical (cf. Hein (2008) and Driemel (2018)), but also if they differ. However, such optionality is not ususally found in morphological paradigms; it is precluded by the final assumption to be mentioned here.

---

[4] Feature copying is a well-established operation in Distributed Morphology; see Halle and Marantz (1993), Embick and Noyer (2007), and Norris (2014), among many others.

Sixth and finally, the present approach incorporates *cyclicity*. Cyclicity manifests itself in two different but related constraints, both of which can be shown to be required in derivational systems based on cyclic application of operations.[5] The first constraint is the Cyclic Principle in (14) (see Perlmutter and Soames (1979) for the formulation; and Adger, Béjar and Harbour (2003), Embick (2010), and Kalin and Weisser (2021), next to Bobaljik (2000), for applications in morphology).[6]

(14) *Cyclic Principle*:
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.

Assuming the most restrictive concept where every projection in a tree qualifies as a cyclic domain, the Cyclic Principle ensures that post-syntactic morphological operations like vocabulary insertion and enrichment (i.e., feature copying) apply root-outwards, exactly as in Bobaljik's approach.

The second cyclicity constraint that we will adopt is the Strict Cycle Condition (cf. Chomsky (1973, 1995, 2015)); a simple version of the constraint is given in (15). As before, it can be assumed that every projection qualifies as a cyclic domain.

(15) *Strict Cycle Condition*:
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 $D_x$.

It remains to be shown how this set of assumptions derives the Partially Superfluous Extended Exponence Generalization in (11). In a nutshell, the underlying logic will be as follows. If vocabulary insertion is to lead to extended exponence, it is clear that feature copying (enrichment) is required; this is the only way how two exponents bearing this feature can show up in a word where there is initially one one occurrence of the feature. Thus, feature copying and vocabulary insertion are two post-syntactic operations that will

---

[5] Cf. Müller (2023) for an overview of the arguments for this claim; also see below.

[6] Also note that the widely adopted Earliness Principle (see Pesetsky and Torrego (2001)) and Featural Cyclicity (see Richards (2001) and Preminger (2018)) are basically just versions of the Cyclic Principle for syntactic derivations.

invariably interact; and the interaction of the two processes is governed by the the cyclicity constraints on the one hand, and by the compatibility and specificity requirements of the Subset Principle on the other hand. As we will see, the interaction ensures that the more *general vocabulary item* only has a chance to show up in a word if it is inserted *early*, in a position close to the root: At the point where the derivation could generate the required copy of the feature for a more general exponent that comes second without violating cyclicity, the feature is already gone as a consequence of insertion of the more specific exponent. If, on the other hand, the more general exponent comes first, no such problem arises: Enrichment can apply early to the crucial feature that is shared by the two exponents. Therefore, the more *specific vocabulary item*, which has more options to satisfy the Subset Principle since it is equipped with more features that permit more insertion sites, can be inserted *later*.

### 3.3. Deriving Partially Superfluous Extended Exponence

Let us assume, as before, that there are two vocabulary items, the more general exponent /a/ ↔ [$f_1$] and the more specific exponent /b/ ↔ [$f_1$,$f_2$]; and that, furthermore, X and Y are functional categories (here assumed to be suffixal) with morpho-syntactic features in need of realization by vocabulary insertion, where X hosts one of these features, and Y hosts the other feature, and X is closer to the root than Y.[7] Then, two basic scenarios need to be considered: In the first one, [$f_1$] is in X, and [$f_2$] is in Y. In the second scenario, it is the other way round: [$f_2$] is in X, and [$f_1$] is in Y. Thus, it follows that /a/ ↔ [$f_1$] can only ever have a chance to be inserted into X in the first scenario, and into Y in the second; a priori, there is no such restriction for /b/ ↔ [$f_1$,$f_2$], which is equipped with both features.

On this basis, let us start with the first scenario (which we will henceforth also refer to as "a→X"): The more general exponent /a/ bears a feature that shows up in the hierarchically lower head X. The abstract derivation in (16) shows how this scenario can lead to a successful instance of partially superfluous extended exponence, in accordance with (11).[8]

---

[7] However, everything that follows can be generalized to scenarios where more features than just two are involved.

[8] Some remarks on notation. Here and in what follows, the current cyclic domain is rendered in black, and the domain which is not yet affected by some operation in gray. $\sqrt{\ }$ designates the root, c the categorizing head, and $\alpha$ the root vocabulary item inserted into $\sqrt{\ }$.

(16) *Scenario a→X, derivation 1:* √

- 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:
  $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /a/\ ]]\ [_{Y}\ /b/\ ]]$

Starting with the complex head in (16a) (formed by earlier head movement of $\sqrt{\ }$ to c, of c to X, and of X to Y, in the syntax or at the beginning of post-syntax), the root is lexicalized first (indicated by $\alpha$) in (16b) (given cyclicity); this completes the c cycle. Next, the derivation moves to the X cycle. Suppose that the first operation that applies is enrichment in (16c); such feature copying is effected by a designated enrichment rule like (17).

(17) $\emptyset \rightarrow [f_1]/[f_1]\_\_$.

Now there are two features $[f_1]$ available in X in (16c). In the next step in (16d), the more general exponent /a/ $\leftrightarrow$ $[f_1]$ is inserted into X, thereby discharging one of the two $[f_1]$ features there. Note that this does not violate the Subset Principle in (13).[9] Finally, the derivation reaches the Y cycle in (16e), and inserts /b/ $\leftrightarrow$ $[f_1,f_2]$ in Y, in accordance both with the Subset Principle ($[f_2]$ on /b/ is matched and discharged in Y, and $[f_1]$ on /b/ is matched and discharged in X), the Cyclic Principle, and the Strict Cycle Condition (insertion of /b/ affects the embedded domain X by discharging $[f_1]$, but it does not do so exclusively since it also affects the Y domain).

The derivation in (16) thus gives rise to partially superfluous extended exponence. It essentially underlies all instances of the phenomenon discussed

---

[9] Does this imply that there could be optionality of /a/ in /b/ in a minimally different system where there is no enrichment, and hence no extended exponence? This is not the case if there is an overarching requirement demanding every head's morpho-syntactic features to be discharged if possible; also, in such a system, there would be arguably no good evidence for two separate heads X and Y in the morphology to begin with.

in section 1. This is shown for the case of dative plural nouns in German (cf. (2)); (18) parallels (16) in all relevant respects.

(18) *A well-formed derivation for dative plural nouns in German*:
Exponents:
(i) /er/ ↔ [+pl]
(ii) /n/ ↔ [+pl,+obl,+gov]
a. Initial structure:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\ }\ \mathrm{c}\ ]\ \mathrm{X}_{[+\mathrm{pl}]}\ ]\ \mathrm{Y}_{[+\mathrm{obl},+\mathrm{gov}]}\ ]$
b. Root lexicalization:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{n}}\ \mathrm{Kind}\ \text{Ø}\ ]\ \mathrm{X}_{[+\mathrm{pl}]}\ ]\ \mathrm{Y}_{[+\mathrm{obl},+\mathrm{gov}]}\ ]$
c. Feature copying on X cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{n}}\ \mathrm{Kind}\ \text{Ø}\ ]\ \mathrm{X}_{[+\mathrm{pl}],[+\mathrm{pl}]}\ ]\ \mathrm{Y}_{[+\mathrm{obl},+\mathrm{gov}]}\ ]$
d. Vocabulary insertion on X cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{n}}\ \mathrm{Kind}\ \text{Ø}\ ]\ [_{\mathrm{X}_{[+\mathrm{pl}]}}\ \mathrm{er}\ ]]\ \mathrm{Y}_{[+\mathrm{obl},+\mathrm{gov}]}\ ]$
e. Vocabulary insertion on Y cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{n}}\ \mathrm{Kind}\ \text{Ø}\ ]\ [_{\mathrm{X}}\ \mathrm{er}\ ]]\ [_{\mathrm{Y}}\ \mathrm{n}\ ]]$

As a matter of fact, it turns out that derivation 1 in scenario a→X is the *only* derivation that can give rise to partially superfluous extended exponence. Consider, e.g., a derivation that is minimally different from derivation 1 in that the order of the two insertion operations is reversed, and /b/ is inserted into Y before /a/ is inserted into X; cf. (19). As illustrated in (19d), premature insertion of /b/ on the Y cycle violates the Cyclic Principle (since insertion of /a/ on the X cycle is skipped); furthermore, final insertion of /a/ would then also violate the Strict Cycle Condition (so this is an environment where the two cyclicity constraints make the same predictions).

(19) *Scenario a→X, derivation 2: **
a. Initial structure:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\ }\ \mathrm{c}\ ]\ \mathrm{X}_{[\mathrm{f}_1]}\ ]\ \mathrm{Y}_{[\mathrm{f}_2]}\ ]$
b. Root lexicalization:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ \mathrm{X}_{[\mathrm{f}_1]}\ ]\ \mathrm{Y}_{[\mathrm{f}_2]}\ ]$
c. Feature copying on X cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ \mathrm{X}_{[\mathrm{f}_1],[\mathrm{f}_1]}\ ]\ \mathrm{Y}_{[\mathrm{f}_2]}\ ]$
d. Vocabulary insertion on Y cycle **Cyclic Principle*:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ \mathrm{X}_{[\mathrm{f}_1]}\ ]]\ [_{\mathrm{Y}}\ /\mathrm{b}/\ ]]$

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

Another derivation that is doomed to fail in the a→X scenario is given in (20). The first three steps are as before; however, in the fourth step in (20d), it is /b/ ↔ $[\text{f}_1,\text{f}_2]$ (rather than /a/ ↔ $[\text{f}_1]$) that is inserted into the X node. This, as such, is in accordance with both the Subset Principle and the cyclicity constraints: Insertion of /b/ affects both the X cycle and the Y cycle (the latter via deletion of $[\text{f}_2]$), but this is unproblematic. However, subsequent insertion of /a/ ↔ $[\text{f}_1]$ in (20d) will now be impossible because of the Subset Principle (Y only hosts an incompatible $[\text{f}_2]$ feature to begin with, and earlier insertion of /b/ has removed this feature in any event).

(20) *Scenario a→X, derivation 3: **

a. Initial structure:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\phantom{\alpha}}\ \text{c}\ ]\ \text{X}_{[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$

b. Root lexicalization:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$

c. Feature copying on X cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1],[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$

d. Vocabulary insertion on X cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}_{[\text{f}_1]}}\ \text{/b/}\ ]]\ \text{Y}\ ]$

e. Vocabulary insertion into Y on Y cycle **Subset Principle:*
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}_{[\text{f}_1]}}\ \text{/b/}\ ]]\ [_{\text{Y}}\ \text{/a/}\ ]]$

Further derivations based on an a→X scenario are also excluded. In particular, derivations in which, on a given cycle where it can apply, enrichment does not precede feature-removing vocabulary insertion can never give rise to extended exponence.[10]

Thus, as an interim conclusion regarding a→X scenarios, it can be noted that the more general exponent /a/ can occur before the more specific exponent /b/ as a result of one derivation (viz., derivation 1); other derivations of an /a/-/b/ sequence fail (cf. derivation 2), and the reverse /b/-/a/ order cannot

---

[10] It has indeed been suggested that operations that manipulate morpho-syntactic features in post-syntactic representations, like impoverishment and, under present assumptions, enrichment, are always ordered before vocabulary insertion in any cyclic domain, due to the nature of the material that they affect; cf. Arregi and Nevins (2012, ch. 6).

be generated in an a→X scenario for very basic reasons (derivation 3). It now remains to be shown that an /a/→Y scenario, where the feature [$f_1$] of the more general exponent /a/ is matched by the outer head Y (i.e., where /a/ realizes a hierarchically higher feature) cannot lead to a successful derivation under present assumptions – in particular, a /b/-/a/ order that would violate the Partially Superfluous Extended Exponence Generalization in (11) must not be generated.

As before, suppose that the enrichment rule in (17) is active, and that /a/ ↔ [$f_1$] and /b/ ↔ [$f_1$,$f_2$] are as before; the only difference is that the lower head X now has [$f_2$], and the higher head Y has [$f_1$]. As shown by derivation 4 in (21), if the more specific exponent /b/ is inserted on the X cycle in (21c), it will remove both [$f_2$] from X *and* [$f_1$] from the higher head Y.[11] When the derivation subsequently moves to the Y cycle, there are no features left for carrying out enrichment, and feature copying will not apply (cf. (21d)). As a consequence, the more general exponent cannot be inserted: There are no features left for morphological exponence of Y.

(21) *Scenario a→Y, derivation 4: **

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

d. Feature copying on Y cycle *cannot apply*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /b/\ ]]\ Y\ ]$

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

The only possible option for Y to be realized by /a/ would be to have feature copying *preceding* morphological realization of X by /b/, so that [$f_1$] on Y can be used to generate a second [$f_1$] before it is removed. However, as

---

[11] Note that this reasoning presupposes that an insertion operation affecting a lower head takes place in the cyclic domain defined by this head, even if, as a consequence of this insertion, a feature on a higher head is ultimately also discharged. This follows naturally if vocabulary insertion is viewed as a complex operation consisting of two separate suboperations, viz., (i) insertion under feature matching followed by (ii) feature discharge.

shown in derivation 5 in (22), this is impossible because of cyclicity. Feature copying takes place early in (22c); but since this operation affects $[f_1]$ on Y, this operation takes place on the Y cycle. An alternative third step of the derivation would have been to carry out vocabulary insertion of /b/; as we have seen, this would have applied on the X cycle. Thus, the derivation has skipped X in (22c), in violation of the Cyclic Principle, and the derivation crashes.

(22) *Scenario a→Y, derivation 5: **

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:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /b/\ ]]\ [_{Y}\ /a/\ ]]$

An interesting question arising at this point is whether subsequent vocabulary insertion of /b/ into X in (22d) violates the Strict Cycle Condition. Here the exaxt wording of the constraint becomes relevant. On the one hand, recall the premise that vocabulary insertion into X *applies to* the cyclic domain X, even if as a consequence eventually some feature beyond X (i.e., in Y) is affected (i.e., discharged); so, for the purposes of the Cyclic Principle in (14), such insertion is an operation on the X cycle. On the other hand, for the Strict Cycle Condition in (15), the question is whether a vocabulary insertion into X *affects* the cyclic domain X or the cyclic domain Y if it ultimately gives rise to feature discharge in Y. If this is the case, the Strict Cycle Condition will not be violated by the step in (22d); if insertion into X in the case at hand does not affect (in the technical sense of (15)) the cyclic domain of Y, the Strict Cycle Condition will be violated. Given that the first option is arguably the more plausible one, this means that we now have a further argument for keeping the two concepts of cyclicity apart: The Cyclic Principle can exclude some sequences of operations as counter-cyclic that may be compatible with the Strict Cycle Condition.

These considerations notwithstanding, it can be concluded that cyclicity plays a major role in deriving the Partially Superfluous Extended Exponence Generalization since it ensures that feature copying in a higher domain (which is required for the presence of the more general exponent in an outer position) cannot take place before vocabulary insertion in a lower domain, which may bleed it. As with the a→X scenario, there are further derivations to consider, but they are all ruled out for obvious reasons. For instance, any derivation in which /a/ is inserted before feature copying takes place will never give rise to extended exponence. More generally, then, the question of how the existence of partially superfluous extended exponence, and the Partially Superfluous Extended Exponence Generalization in (11), can be derived in Distributed Morphology has received an answer: The Cyclic Principle and the Subset Principle can only be both satisfied in derivation 1: If the more general (proper subset) exponent realizes a hierarchically lower feature, it must come first; if it realizes a hierarchically higher feature, there is no good output because feature copying is bled by cyclic vocabulary insertion.

### 3.4. Overlapping Extended Exponence

At this point, the question arises of how overlapping extended exponence can be accounted for under the present system of assumptions; as noted above, this type of extended exponence is much less of a challenge for many theories of moprhology. Indeed, overlapping extended exponence is also predicted to be possible in the approach under consideration – but, as we will see, there is a caveat.

Suppose that there are two exponents /a/ ↔ [$f_1$,$f_2$] and /b/ ↔ [$f_1$,$f_3$] that share a feature [$f_1$], as in (6); and that there is a functional head X initially bearing the features [$f_1$], [$f_2$], and a functional head Y that is equipped with the feature [$f_3$]. In addition, the enrichment rule (17) is active in the language, as before. Under these assumptions, the derivation in (23) gives rise to overlapping exponence. Importantly, feature copying applying to [$f_1$] on X must take place early, on the X cycle; see (23c). After that, /a/ is inserted into X, thereby discharging [$f_2$] and one copy of [$f_1$] in X; see (23d). Finally, /b/ is inserted into Y, which gives rise to a discharge of [$f_3$] in Y, and of the other copy of [$f_1$] in X; see (23e).

(23) *Overlapping extended exponence, derivation 1:* √

a. Initial structure:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_1],[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_3]}\ ]$
b. Root lexicalization:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1],[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_3]}\ ]$
c. Feature copying on X cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_1],[\text{f}_1],[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_3]}\ ]$
d. Vocabulary insertion on X cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}_{[\text{f}_1]}}\ /\text{a}/\ ]\ \text{Y}_{[\text{f}_3]}\ ]$
e. Vocabulary insertion on Y cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}}\ /\text{a}/\ ]\ [_{\text{Y}}\ /\text{b}/\ ]]$

Thus, overlapping extended exponence is predicted to be possible. However, as noted, there is a caveat. Overlapping extended exponence is in fact ceteris paribus predicted to be *impossible* if the shared feature is not on the lower head, as in (23), but on the higher head, as in the minimally different derivation 2 in (24). Here, the shared feature [$\text{f}_1$] is not on X, but on Y; see (24a). Since [$\text{f}_1$] is not present on the X cycle, feature copying cannot apply here, and the derivation inserts /a/ on the X cycle, which removes both [$\text{f}_2$] from X, which is unproblematic, and [$\text{f}_1$] from Y, which is fatal because now there is no [$\text{f}_1$] feature left that enrichment on the Y cycle could apply to; cf. (24c). Subsequent insertion of /b/ on the Y cycle will therefore have to violate the Subset Principle; see (24d).

(24) *Overlapping extended exponence, derivation 2:* *

a. Initial structure:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_3]}\ ]$
b. Root lexicalization:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_3]}\ ]$
c. Vocabulary insertion on X cycle:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}}\ /\text{a}/\ ]\ \text{Y}_{[\text{f}_3]}\ ]$
d. Vocabulary insertion on Y cycle **Subset Principle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ [_{\text{X}}\ /\text{a}/\ ]\ [_{\text{Y}}\ /\text{b}/\ ]]$

As we have seen with the analogous issue for partially superfluous extended exponence, any attempt at solving the problem in (24) by applying feature copying to [$\text{f}_1$] on Y earlier will invariably lead to a violation of the Cyclic Principle; see (25).

(25) *Overlapping extended exponence, derivation 3: **

a. Initial structure:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_3]}\ ]$

b. Root lexicalization:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_3]}\ ]$

c. Feature copying on Y cycle **Cyclic Principle*:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\alpha}\ \text{c}\ ]\ \text{X}_{[\text{f}_2]}\ ]\ \text{Y}_{[\text{f}_1],[\text{f}_1],[\text{f}_3]}\ ]$

Thus, there is an unresolvable problem with a derivation where the shared feature in overlapping exponence is on the higher head: Insertion of a specific exponent discharges the shared feature before it can be copied for the other exponent in accordance with the Cyclic Principle. At present, we take it to be an open question whether this prediction might be called into question by empirical evidence; pursuing this issue in detail, based on the available empirical evidence, is beyond the scope of the present paper.

### 3.5. Fully Superfluous Extended Exponence

Fully superfluous exended exponence is predicted to be impossible: Either there is a problem with the compatibility (i.e., the Subset Principle), or there is a cyclicity problem. Thus, suppose that there are two exponents /a/ ↔ [$\text{f}_1$,$\text{f}_2$] /b/ ↔ [$\text{f}_1$,$\text{f}_2$] with identical feature specifications, as in (8); by assumption, the features [$\text{f}_1$] and [$\text{f}_2$] are located on two separate heads X and Y (if they were to show up on only one head, extended exponence would trivially be excluded). Given these assumptions, a schematic derivation illustrating the compatibility problem problem is given in (26). Feature copying applies early here (cf. (26c)), but subsequent insertion of /a/ (or, for that matter, /b/) in X leads to discharge of [$\text{f}_2$] on Y, so that the remaining exponent can never be inserted in Y (cf. (26d)).

(26) *Fully Superfluous extended exponence, derivation 1: **

a. Initial structure:
$[_{\text{Y}}\ [_{\text{X}}\ [_{\text{c}}\ \sqrt{\ }\ \text{c}\ ]\ \text{X}_{[\text{f}_1]}\ ]\ \text{Y}_{[\text{f}_2]}\ ]$

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:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}\ /a/\ ]]\ Y\ ]$
e. Vocabulary insertion into Y on Y cycle **Subset Principle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}\ /a/\ ]]\ [_{Y}\ /b/\ ]]$

The cyclicity problem arising with the alternative derivation where feature copying on the Y cycle precedes vocabulary insertion on the X cycle is shown in (27).

(27) *Fully Superfluous extended exponence, derivation 2: **

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

### 3.6. Convergence

Arguably, from a more general point of view, the present approach to partially superfluous extended exponence based on Distributed Morphology captures the same underlying core idea as the approaches based on derivational versions of Optimality Theory (strata, f-seq, harmonic serialism) that were discussed in section 2: In the present approach, the more general exponent needs to find a matching feature in the syntactic head into which it is supposed to be inserted (because of the Subset Principle), and the interaction of copying and insertion governed by cyclicity ensures that there will not be such a feature if the more general exponent comes too late. So, all these approaches share the common core that the more general exponent can only show up if it shows up early; once the more specific exponent is part of the structure, it will block the more general exponent.[12]

## 4. Empirical Issues

### 4.1. Exceptions

There are exceptions to the Partially Superfluous Extended Exponence Generalization, i.e., cases where the more general (proper subset) exponent can or must show up further away from the stem than the more specific (superset) exponent. To name just a few examples: There is partially superfluous extended exponence of negation in past contexts in Swahili (see Stump (2001)), in forms

---

[12] An issue that the preceding approach has so far remained silent on is extended exponence involving roots and their categorizing heads; such phenomena are usually discussed under rubrics like *root suppletion* or *stem allomorphy*. Thus, suppose that the combination $\sqrt{\ }$-c of an abstract root morpheme $\sqrt{\ }$ and a categorizing head c (which, to simplify matters, we will treat as a primitive item here) is equipped with a feature [$f_1$], and the next higher functional head X has the feature [$f_2$]. Now, if the actual root vocabulary item $\alpha$ is characterized as [$f_1$,$f_2$] (such that $\alpha$ is expected to give rise to [$f_2$]-conditioned suppletion), then there will be a problem because $\alpha$-insertion will discharge [$f_2$] on X, and a vocabulary item realizing X will not be insertable anymore; furthermore, this consequence is independent of whether we are dealing with partially superfluous extended exponence (as in this scenario) or overlapping extended exponence (if the vocabulary item for X also has some other feature). (Thanks to Elango Kumaran for noticing this.)

There are at least two ways to address this issue under present assumptions. One option would be to assume that the $\sqrt{\ }$-c stem is special in that it does not in fact qualify as a cyclic domain, in contrast to what we have assumed so far. As a consequence, [$f_2$] on X can be copied before $\alpha$ is inserted in the root position. Another option would be to stipulate that the insertion of root vocabulary items is special in that it does not lead to feature discharge. The two options differ with respect to the predictions for *non-local stem allomorphy*, where $\alpha$ bears a feature [$f_3$] that is located on some yet higher head Y (either instead of [$f_2$], or in addition to [$f_2$]): In the first approach, more must be said to permit such non-local extended exponence; in the second approach, it can be derived without problems (as long as Y is still part of the same morphological phase). In view of the fact that non-local stem allomorphy appears to be a marked phenomenon in the world's languages, and requires additional assumptions (and, often, additional tools, like spanning, hyper-contextual realization rules or buffers) in all existing derivational approaches to morphology (see Merchant (2015), Moskal and Smith (2016), Weisser (2017), Kastner and Moskal (2018), and Božič (2019)), we will refrain from deciding this question.

like *ha-tu-ku-taka* (NEG-1.PL-NEG.PAST-want; ‘We did not want’), where the pure, general negative prefix *ha* shows up outside of the more specific negative past prefix *ku*; there is also partially superfluous extended exponence of third person in plural contexts in Ojibwe (see Oxford (2019)), in forms like *waapam-ikw-waa-t-pan* (see-INV-3.PL-3-PRET; ‘The other saw them’), where the more specific third person plural exponent *waa* is closer to the root (*waapam*) than the more general bare plural exponent *t*; and several more of such examples can be found in the literature. Clearly, if these apparent exceptions are taken at face value, this would imply that there is no interesting generalization to be made about partially superfluous extended exponence after all. What is more, from the perspective of grammatical theory, there would be no principled answer left to the question of why the phenomenon exists in the first place (since all available approaches that have something to say about this question derive the Partially Superfluous Extended Exponence Generalization).

In view of this state of affairs, it seems to us that the most promising strategy is to account for apparent exceptions to the Partially Superfluous Extended Exponence Generalization in a way that leaves the generalization (and its explanation) intact. Accordingly, we would like to contend that if the more general exponent shows up outside of the more specific exponent in a partially superfluous extended exponence scenario, this is either (i) due to exponent movement (either in the morphological component or in the phonological component), or can (ii) be shown to be compatible with the generalization after all, due to a reanalysis of the data.

As for exponent movement (i), the assumption is that the more general exponent is first inserted into the word in a position that is closer to the root than the position of the more specific exponent but subsequently moves to a position outside the domain of this latter marker. Note that this is in all relevant respects identical to what happens in syntactic derivations; for instance, an object must be base-generated closer to the verb than the subject (cf. *Mary often [reads books]*), but may, as a consequence of movement, eventually come to be placed outside of the domain of the subject at the end of the derivation (cf. *What does [Mary read]?*). In Müller (2020) and Gleim et al. (2021, 2022), arguments are presented for the existence of word-internal movement of morphological exponents that is triggered by alignment constraints; in line with this, an example like Swahili *ha-tu-ku-taka* is argued to involve morphological movement of *ha*, triggered by a morphological alignment constraint that requires left-alignment of negative items in a word (whereas a higher-ranked alignment constraint on all tense exponents forces *ku* to stay in situ). An alternative to morphological movement of exponents is phonological movement of exponents, which is triggered by purely phonological requirements (this is analogous to the concept of PF movement in syntactic derivations; cf., e.g., Chomsky (1995), Truckenbrodt (1995), Agbayani et al. (2015)).

As for reanalyses of the data (ii), we suggest that closer inspection of apparent exceptions to the Partially Superfluous Extended Exponence Generalization will often reveal that there is in fact no extended exponence to begin with (in the sense that a single morpho-syntactic feature justified for syntactic reasons is realized by two or more exponents in the morphology). There are various possibilities as to how such a configuration can come about. One possibility is that there are two independently motivated occurrences of the same feature in a given word from the start (i.e., as a consequence of what happens in the syntax), which are then separately targetted by morphological realization without requiring enrichment (see, e.g., Sells (2004) and Alexiadou et al. (2021) for relevant discussion, also with respect to larger grammatical units). Another possibility is that an independently motivated decomposition of seemingly primitive morpho-syntactic features provides more targets for morphological realization by exponents – if, say, a feature [±A] that, at first sight, seems to be realized by two exponents $\alpha$ and $\beta$, is to be decomposed into a combination [±b,±c], it may be the case that $\alpha$ realizes only [±b], and $\beta$ only [±c] (cf. Stiebels (2016) and Caha (2021) for analyses along these lines).

In what follows, we will discuss two relevant cases in a bit more detail, viz., phonologically triggered movement of more general morphological exponents in Huave (see Kim (2010)), and the distribution of $\phi$-features and case features on class markers in Itelmen (see Bobaljik (2000)).

### 4.2. Partially Superfluous Extended Exponence and Phonological Exponent Movement in Huave

As noted in Grofulović et al. (2021), there are patterns of partially superfluous extended exponence in San Francisco del Mar Huave that seem to contradict the order predicted by the Partially Superfluous Extended Exponence Generalization in (11). The inflected verb in (28) exhibits partially superfluous extended exponence: *s* (here realized as [ʃ]) is a general first person marker, and *n* is a more specific exponent realizing first person and subordination (SB). Since both exponents are prefixes of the transitive verb *a-ʰʧ*, the Partially Superfluous Extended Exponence Generalization is contradicted on the surface.

(28) *ʃ*-i-*n*-a-ʰʧ
1-FT-1SB-TV-give
'I will give'

However, there is evidence that the positions of *s* and *n* are motivated by phonological requirements, and do not necessarily represent the positions occupied by the exponents when vocabulary insertion takes place in the morphology. The crucial observation is that San Francisco del Mar Huave exhibits the phenomenon of so-called mobile affixation (see Kim (2010), Zukoff (2021)): Depending on phonological constraints, one and the same morphological exponent may show up in different positions in a word. The phenomenon of mobile affixation in San Francisco del Mar Huave is illustrated in (29) (cf. Kim (2010)).[13]

(29) a. *t*-a-ʰʧ-ju-s
CP-TV-give-1
'I gave'

b. paʰk-a-*t*-u-s
face.up-V-CP-ITR-1
'I lay face up'

In San Francisco del Mar Huave, the completive aspect (CP) exponent *t* is one of several exponents that are 'mobile' in the sense that they can show up either as a suffix (as in (29b)) or as a prefix (as in (29a)). The placement is regulated by the phonotactic constraints of the language. If the CP exponent *t* occurs with a verbal base starting with a consonant and ending in a vowel, it is realized as a suffix; however, if the verbal base starts with a vowel and ends in a consonant, the consonant cluster that would result if *t* were to be realized as a suffix, and the ensuing vowel epenthesis that would take place as a repair, are avoided by realizing the exponent as a prefix instead.

Returning to the problematic case of partially superfluous extended exponence in (28), it can be noted that the general first person marker *s* and the more specific person/subordination marker *n* are both mobile affixes; e.g., (30) illustrates that *n* shows up as a suffix under the right phonotactic conditions.

---

[13] The examples are rendered in IPA, based on the convention in Zukoff (2021) and the glossing rules given in Kim (2010).

(30) *ʃ*-i-ʧut-u-*n*
1-FT-sit-V-1SB
‘I will sit’

Approaches to mobile affixation (like Kim (2010) and Zukoff (2021)) typically do not postulate that movement of any type is actually involved; rather, the assumption normally is that the exponents in question are directly placed in a word according to the demands of syllable structure constraints, with morphology and phonology intermingled. However, as argued in Grofulović et al. (2021), the available evidence is fully compatible with an approach that respects modularity by separating vocabulary insertion in underlying morphology from the phonologically conditioned placement of an exponent in actual output forms (see also Kalin and Rolle (2021) for independent arguments for such an approach). On this view, the two exponents in a partially superfluous extended exponence relation in (28) and (30) are first inserted in the order required by the Partially Superfluous Extended Exponence Generalization, with the more general exponent coming first and the more specific one coming second; and subsequently, phonologically driven movement takes place, leading to an opaque surface representation.

### 4.3. Partially Superfluous Extended Exponence Reanalyzed in Itelmen

As a second case study, let us look at the distribution of $\phi$-features and case features on agreement exponents and inflection class markers in Itelmen, which Bobaljik (2000) takes to provide the core empirical evidence in support of his model of contextual allomorphy discussed in section 3.1 above. There are two patterns in Itelmen verb inflection that initially would seem to support the view that there is outwards-sensitivity to morpho-syntactic features on functional heads, but no inwards-sensitivity. From the present perspective, this implies that the more specific exponent is closer to the root than the more general exponent in a partially superfluous extended exponence scenario; consequently, the patterns in question pose potential challenges for the generalization in (11) and, more specifically, the present account of the generalization based on how the interaction of vocabulary insertion and feature copying (enrichment) is governed by cyclicity.

Following Bobaljik (2000), a simplified structure of Itelmen verbs containing the relevant functional morphemes in need of morphological realization is given in (31): AgrS is realized as a prefix, and Class and AgrO are realized by suffixal exponents.

(31) [$_{\text{AgrS}}$ AgrS [$_{\text{AgrO}}$ [$_{\text{Class}}$ V Class ] AgrO ]]

Two instances of (what at first sight looks like) partially superfluous extended exponence can be observed. First, as illustrated in (32a), an object (accusative) agreement exponent in AgrO (here: *čeʔn*) can realize features of both the subject (nominative) and the object (or, in Bobaljik's terms, an object agreement exponent can be conditioned by subject agreement features); in addition, a more general subject (nominative) exponent shows up in AgrS. Second, in (32b), the suffixal inflection class exponent *ki* is specified for $\phi$-features of both the nominative and the accusative argument but clearly shows up closer to the root than at least the AgrO exponent.

(32) *Extended exponence of nominative/accusative and case/class exponents in Itelmen*:

a. [$_{\text{AgrS}}$ *tʹ* [$_{\text{AgrO}}$ [$_{\text{Class}}$ [$_{\text{V}}$ əŋkzu-s ] Ø ]
1.SG.NOM help-PRES CL.I
*čeʔn* ]]
1.NOM./3.PL.ACC
'I'm helping them.'

b. [$_{\text{AgrS}}$ *t* [$_{\text{AgrO}}$ [$_{\text{Class}}$ [$_{\text{V}}$ t$\phi$-s ]
1.SG.NOM bring-PRES
*ki* ] *čeʔn* ]]
CL.II.1.SG.NOM/3.PL.ACC 1.NOM./3.PL.ACC
'I'm bringing them.'

Given (32a) and (32b), it looks as though Itelmen exhibits two patterns where a more specific exponent shows up in a position that is closer to the stem than the position occupied by the more general exponent in a partially superfluous extended exponence configuration.[14] Assuming this to be the case, a potential problem arises for the generalization in (11) and the present analysis.

Turning to the subject and object agreement exponents in (32a) first, it can be noted that the specific analysis that Bobaljik (2000), following Bobaljik and Wurmbrand (1997), suggests for examples like (32a) does not actually instantiate extended exponence in the sense adopted througout this paper, as *one* morpho-syntactic feature resulting in realization by *two* (or more) exponents. More specifically, here are Bobaljik and Wurmbrand's (1997) assumptions about the AgrO slot: First, in third person object environments, there are no person features in AgrO. Second, there is an EPP-like requirement in Itelmen to have person features in this position. Third, to satisfy this requirement, Itelmen employs a general Agree-like copying mechanism where *all* the features of AgrS are copied onto AgrO (i.e., not just the missing person information, but the whole feature bundle). Fourth, Agree-like copying of subject features also takes place in intransitive environments. As a consequence of these assumptions, the $\phi$-features and case features of subjects are available twice in the Itelmen verb in the relevant contexts, on AgrS and on AgrO; and assuming that the Agree-like copying operation is either syntactic or, at least, takes place very early in the post-syntactic component (a view which is supported by the fact that the operation is non-local and transfers feature bundles corresponding to entire categories), there will be no cyclicity restrictions of the type addressed in section 3 above. Thus, AgrS in (32) is realized by a subject agreement exponent, and AgrO in (32) is realized by a *portmanteau exponent*. Consequently, if this analysis is adopted, there is no problem with either the Partially Superfluous Extended Exponence Generalization or its derivation via cyclic application of feature copying and

---

[14] At least, this corresponds to Bobaljik's (2000) conclusion, based on the premise that the structure of inflected verbs in Itelmen looks as in (31), and that his assumptions about contextual allomorphy hold. Strictly speaking, however, the data in (32) only unequivocally show that more specific Class exponents are closer to the root than more general AgrO exponents. Bobaljik notes that independent evidence for the height of prefixes in an Itelmen verb is hard to come by; but Bobaljik and Wurmbrand (2001) provide a couple of arguments for the view that subject agreement prefixes are higher than object agreement prefixes in the language. At least for the sake of the argument, we will follow Bobaljik in assuming that not only are more specific Class exponents closer to the root than the AgrO exponents that they are "conditioned by", but more specific AgrO and Class exponents are also closer to the root than the AgrS prefix exponents they are "conditioned by".

vocabulary insertion: There is no extended exponence, just faithful realization of two separate feature sets that are independently present in the structure.

Turning to the interaction of the subject and object agreement exponents with the inflection class exponents in (32b) next, the situation is a bit more complex; but the overall conclusion will be the same: It is likely that there is no partially superfluous extended exponence involved.

For concreteness, there is evidence that sheds doubt on the existence of a separate functional morpheme hosting class exponents. Bobaljik (2000) observes that there are a number of class II markers: Next to *ki* in (32b), there is *k*, there is *čiŋ*, there is *xk*, etc.; the choice among these is mainly determined by the $\phi$-features of subject and object. These inflection class exponents always show up immediately adjacent to the AgrO exponent, not adjacent to V (tense exponents, e.g., intervene between V and the alleged inflection class exponent, as in (32b)), which may already be regarded as somewhat suspicious. Accordingly, Georg and Volodin (1999) analyze strings like *ki-čeʔn* as *primitive*, non-decomposable AgrO exponents without any internal fine structure: *kičeʔn*. Interestingly, exactly the same string *kičeʔn* also shows up in the other, unmarked inflection class I *in intransitive contexts*; and in this context, Bobaljik (2000) also assumes that *kičeʔn* is indeed a primitive, non-decomposable AgrO marker. Thus, in Bobaljik's (2000) system, *kičeʔn* is viewed as a concatenation of two morphological exponents in one environment, and as a single morphological exponent in another, closely related environment. This looks like a generalization is being missed.

In view of all this, we would like to suggest that the functional morpheme AgrO in Itelmen is subject to *fission*, in the sense of Noyer (1997) and Trommer (1999): Vocabulary insertion discharges features in a fissioned morpheme, but the remaining features can trigger a new vocabulary insertion operation affecting the same functional morpheme.[15] At this point, the question arises of how an analysis of strings like *kičeʔn* that is based on subanalysis of exponents, i.e., fission, can avoid potential problems for the present account. There are at least two possible answers.

First, given fission, the various AgrO exponents are not hierarchically distinct; they are all inserted into one and the same functional morpheme,

---

[15] This implies that the *disjunctive blocking* assumption from section 3.2 must be qualified for fissioned morphemes. Independent evidence for fission of AgrO heads in Itelmen comes from the distribution of *partial syncretism*; there are several such cases in the paradigm; cf., e.g., *če-n* (1.NOM./3.SG.ACC) vs. *če-ʔn* (1.NOM./3.PL.ACC).

and are therefore all part of one and the same morphological cycle: [$_{\text{AgrO}}$ *ki–če–ʔn* ], not *[[[*ki*]–*če*]–ʔn ]. Therefore, features can be copied early, before they are discharged by vocabulary insertion. On this view, there is extended exponence (of the partially superfluous or overlapping type), but it is entirely unproblematic: Given the absence of discriminating structure in a fissioned morpheme, the feature copying that is required for extended exponence is not required by cyclicity to come too late to feed exponence.

Second, a closer analysis of the morphological system reveals that the feature sets realized by the class marker and the object agreement exponent could in fact emerge as *complementary*: It looks as though it might be possible to maintain the view that the class exponent realizes inflection class and subject agreement features in AgrO, and the object agreement exponent realizes only object agreement features in AgrO. Under such an analysis, there would be no extended exponence in the system. For reasons of space and coherence, we will not try to advance a full-fledged analysis of the whole paradigm of verb inflection in Itelmen here, and decide for one of the two options; suffice it to say that on either view, the problem that the Itelmen data in (31) might initially pose for the Partially Superfluous Extended Exponence Generalization (and its derivation based on cyclicity) disappears.

## References

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.

Agbayani, Brian, Chris Golston and Toru Ishii (2015): 'Syntactic and Prosodic Scrambling in Japanese', *Natural Language and Linguistic Theory* **33**, 47–77.

Alexiadou, Artemis, Imke Driemel, Johannes Heinand Fabienne Martin and Yining Nie (2021): Multiple Exponence in Child Language. Ms., ZAS Berlin.

Alhoniemi, Alho (1993): *Grammatik des Tscheremissischen (Mari)*. Helmut Buske Verlag, Hamburg.

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

Bermúdez-Otero, Ricardo (2011): Cyclicity. *In:* M. van Oostendorp, C. Ewen, E. Hume and K. Rice, eds, *The Blackwell Companion to Phonology*. Vol. 4, Wiley-Blackwell, Malden, MA, pp. 2019–2048.

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 Susi Wurmbrand (1997): Preliminary Notes on Agreement in Itelmen. *In:* B. Bruening, Y. Kang and M. McGinnis, eds, *PF: Papers at the Interface*. MIT Workinpg Papers in Linguistics, MIT, Cambridge, Mass., pp. 395–423.

Bobaljik, Jonathan and Susi Wurmbrand (2001): Seven Prefix-Suffix Asymmetries in Itel′men. *In:* M. Andronis, C. Ball, H. Elston and S. Neuvel, eds, *Proceedings from the Panels of the Chicago Linguistic Society's Thirty-Seventh Meeting*. Chicago Linguistic Society, Chicago, pp. 205–219.

Božič, Jurij (2019): 'Constraining Long-Distance Allomorphy', *The Linguistic Review* **36**, 485–505.

Brown, Dunstan and Andrew Hippisley (2012): *Network Morphology*. Cambridge University Press, Cambridge.

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.

Caha, Pavel (2021): No Multiple Exponence. Ms., Masarykova Univerzita.

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 (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 (2015): Problems of Projection: Extensions. *In:* E. Di Domenico, C. Hamann and S. Matteini, eds, *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Benjamins, Amsterdam, pp. 1–16.

Dahlstrom, Amy (2000): Morphosyntactic Mismatches in Algonquian: Affixal Predicates and Discontinuous Verbs. *In:* A. Okrent and J. Boyle, eds, *Proceedings from the Panels of the Chicago Linguistic Society's Thirty-Sixth Meeting*. Chicago Linguistic Society, Chicago, pp. 63–87.

Don, Jan and Elma Blom (2006): A Constraint-Based Approach to Morphological Neutralization. In: *Linguistics in the Netherlands 2006*. Benjamins, Amsterdam, pp. 78–88.

Driemel, Imke (2018): Person Mismatch Agreement. *In:* S. Hucklebridge and M. Nelson, eds, *Proceedings of NELS 48*. GLSA, Amherst, Mass.

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

Embick, David and Rolf Noyer (2007): Distributed Morphology and the Syntax/Morphology Interface. In: *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press, Oxford, pp. 289–324.

Fenger, Paula and Philipp Weisser (2023): What Umlaut Tells Us about the Underlying Morphological Structure of Verbs in Sinhala. Ms., Universität Leipzig. To appear in Proceedings of (F)ASAL.

Georg, Stefan and Alexander Volodin (1999): *Die itelmenische Sprache*. Harrasowitz, Wiesbaden.

Gleim, Daniel, Gereon Müller, Mariia Privizentseva and Sören Tebay (2021): Non-Local Phonology by Morphological Movement. *In:* A. Farinella and A. Hill, eds, *Proceedings of NELS 51*. University of Massachusetts, Amherst.

Gleim, Daniel, Gereon Müller, Mariia Privizentseva and Sören Tebay (2022): 'Reflexes of Exponent Movement in Inflectional Morphology: A Study in Harmonic Serialism', *Natural Language and Linguistic Theory* **41**, 103–158.

Grimshaw, Jane (2001): Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. *In:* G. Legendre, J. Grimshaw and S. Vikner, eds, *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 205–240.

Grofulović, Jelena, Gereon Müller and Sam Zukoff (2021): Mobile Affixes Undergo Movement: An Argument from Extended Exponence. Ms., Universität Leipzig and USC.

Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. *In:* B. Bruening, Y. Kang and M. McGinnis, eds, *Papers at the Interface*. Vol. 30, MITWPL, pp. 425–449.

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.

Hanink, Emily (2018): 'Super Light-Headed Relatives, Missing Prepositions, and Span-Conditioned Allomorphy in German', *The Journal of Comparative Germanic Linguistics* **21**, 247–290.

Hein, Johannes (2008): Verbflexion im Warembori: Eine Analyse im Rahmen der Distribuierten Morphologie. *In:* F. Heck, G. Müller and J. Trommer, eds, *Varieties of Competition*. Vol. 87 of *Linguistische Arbeits Berichte*, Universität Leipzig, pp. 49–63.

Inkelas, Sharon, Teresa McFarland and Anne Pycha (2006): The Flip Side of Blocking: Multiple Exponence in an Optimization Framework. Ms., University of California, Berkeley.

Kalin, Laura and Nicholas Rolle (2021): Deconstructing Subcategorization: Conditions on Insertion versus Conditions on Position. Ms., Princeton University and ZAS Berlin. To appear in *Linguistic Inquiry*.

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

Kastner, Itamar and Beata Moskal, eds (2018): *Non-Local Contextual Allomorphy*. Vol. 34 of *Snippets*, Ledonline. Special Issue.

Kim, Yuni (2010): 'Phonological and Morphological Conditions on Affix Order in Huave', *Morphology* **20**, 133–163.

Kiparsky, Paul (1973): Abstractness, Opacity and Global Rules. *In:* O. Fujimura, ed., *Three Dimensions in Linguistic Theory*. TEC, Tokyo, pp. 57–86.

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

Marvin, Tatjana (2002): Topics in the Syntax and Stress of Words. PhD thesis, MIT, Cambridge, Mass.

Matthews, Peter (1972): *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge University Press, Cambridge.

Matthews, Peter (1974): *Morphology*. Cambridge University Press, Cambridge.

McCarthy, John (2016): The Theory and Practice of Harmonic Serialism. *In:* J. McCarthy and J. Pater, eds, *Harmonic Grammar & Harmonic Serialism*. Equinox, Sheffield, pp. 47–87.

McFarland, Teresa (2009): The Verbal Morphology of Totonaco de Filomeno Mata. PhD thesis, University of California, Berkeley.

Merchant, Jason (2015): 'How Much Context is Enough? Two Cases of Span-Conditioned Allomorphy', *Linguistic Inquiry* **46**, 273–303.

Moskal, Beata and Peter Smith (2016): 'Towards a Theory Without Adjacency: Hyper-Contextual VI-Rules', *Morphology* **26**, 295–312.

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

Müller, Gereon (2023): Challenges for Cyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig, pp. 1–32.

Norris, Mark (2014): A Theory of Nominal Concord. PhD thesis, UCSC.

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

Ortmann, Albert (2004): A Factorial Typology of Number Marking. *In:* G. Müller,

L. Gunkel and G. Zifonun, eds, *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 229–267.

Oxford, Will (2019): Fission in Algonquian and the Status of Morphological Templates. *In:* D. Reisinger and R. Y.-H. Lo, eds, *Proceedings of the Workshop on the Structure and Constituency of Languages of the Americas*. UBCWPL, Vancouver, BC.

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

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.

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.

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

Sells, Peter (2004): Syntactic Information and Its Morphological Expression. *In:* L. Sadler and A. Spencer, eds, *Projecting Morphology*. CSLI Publications, Stanford University, pp. 187–225.

Starke, Michal (2001): Move Dissolves Into Merge. PhD thesis, University of Geneva.

Stiebels, Barbara (2006): 'Agent Focus in Mayan Languages', *Natural Language and Linguistic Theory* **24**, 501–570.

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

Stiebels, Barbara (2016): Multiple Exponence in a Lexical-Incremental Framework. Ms., Universität Leipzig.

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

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 (2001): Distributed Optimality. PhD thesis, Universität Potsdam.

Trommer, Jochen (2011): Phonological Aspects of Western Nilotic Mutation Morphology. Habilitation thesis, Universität Leipzig.

Truckenbrodt, Hubert (1995): Extraposition from NP and Prosodic Structure. *In:* J. Beckman, ed., *NELS 25*. University of Massachusetts, GLSA, Amherst, pp. 503–517.

Weisser, Philipp (2017): Allomorphy Tracks Selection. Ms., Universität Leipzig.

Wiese, Bernd (2000): Warum Flexionsklassen? Über die deutsche Substantivdeklination. *In:* R. Thieroff, M. Tamrat, N. Fuhrhop and O. Teuber, eds, *Deutsche Grammatik in Theorie und Praxis*. Niemeyer, Tübingen, pp. 139–153.

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.

Zukoff, Sam (2021): A Parallel Approach to Mobile Affixation in Huave. *In:* R. Bennett, R. Bibbs, M. L. Brinkerhoff, M. Kaplan, S. Rich, N. Van Handel and M. Wax Cavallaro, eds, *Supplemental Proceedings of AMP 2020*. LSA, Washington, DC.

# Delayed Exponence in Murrinhpatha as an Instance of Myopia in Morphology

Marie-Luise Popp*

## Abstract

In this paper, I discuss and analyze an intricate morphological pattern in Murrinhpatha which involves reordering of the dual marker *ngintha* and an alternation in the form of its adjacent morpheme. I will argue that the phonological correlates of morphemes provide evidence for a cyclic structure of the word in Murrinhpatha. In combination with independently motivated morphological constraints and the featural specifications of the marker, I suggest an analysis couched in Stratal Optimality Theory, where the cyclic architecture of the word provides a straightforward explanation for the placement of the dual marker and the resulting switch in the form of the classifier stem without stipulating position classes as primitive entities of morphological theory. Furthermore, I argue that a cyclic structure neatly explains the simultaneous realization of the daucal (dual/paucal) classifier stem and *ngintha*, which looks like multiple exponence on the surface. My analysis suggests that the overexponence results from the blocking of *ngintha* in the first cycle and the selection of the featurally more specific daucal stem. However, *ngintha* is not strictly bounded to the first cycle, and its realization is delayed until the second cycle. Put shortly, the morphological grammar in the first morphophonological domain cannot anticipate that *ngintha* will be realized in a later stage of the derivation, thus creating an instance of *myopia* in morphology.

## 1. The Peculiar Placement of Number in Murrinhpatha

Murrinhpatha is a morphologically highly complex language, which is spoken in the Northern Territory of Australia. The relative ordering of bound morphemes within the verbal complex in Murrinhpatha is sketched in table 1. As shown in table 1, the left edge of the verbal complex is occupied by a morpheme traditionally labeled as *classifier stem* or *finite stem*. Classifier stems are typically treated as portmanteau forms that encode classifying

---

[*] For comments and discussion, I would like to thank Gereon Müller, Barbara Stiebels and Jochen Trommer, as well as the participants of the *Linearization in Morphology* seminar in July 2022.

*Cyclicity*, 199–238
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

semantics, subject person and number, as well as tense and mood information (Mansfield 2019, Nordlinger and Mansfield 2021). While information about subject person is realized as part of the classifier stem, object person is marked by affixes that attach right to the classifier stem in slot 2. Another crucial part of the verbal complex is the *lexical stem*, which is sometimes referred to as *coverb*. The lexical stem is an uninflected part of the predicate and is realized in slot 5. In addition, a couple of morphemes may be concatenated in positions after the lexical stem; however, only two of these morphemes are relevant for the purpose of this paper.[1] First, TAM markers are linearized after the lexical stem. Second, certain number markers may be realized in positions following the lexical stem. Note also that the relative order of the TAM markers and the number markers is flexible to some extent, while the relative order of morphemes in the domain spanning from the classifier stem until the lexical stem is fixed (Mansfield 2017). Table 1 further shows that subject number is realized in three different positions: first, it is part of the subject information encoded in the classifier stem. Second, additional morphemes realizing subject number are realized either in slot 2 and hence, in direct adjacency to the classifier stem, or in slot 8 at the right edge of the verb. In this paper, I will explain the distribution and positioning of the number markers in Murrinhpatha and how their position patterns with their phonological properties.

| **Slot 1** | **Slot 2** | **Slot 3** | **Slot 4** | **Slot 5** | **Slot 6** | **Slot 8** |
|---|---|---|---|---|---|---|
| $\boxed{\text{Classifier stem}}$ (portmanteau w. SUBJ and TAM) | SUBJ number<br>OBJ marker<br>OBL marker | REFL/ REC | incorporated body part/ APPL | lexical stem | TAM | number (SUBJ or OBJ) |

Table 1: Relative ordering of morphemes (Nordlinger and Mansfield 2021: 2)

Table 1 illustrates a crucial property of Murrinhpatha morphology: the verbal predicate is typically bipartite, comprising a classifier stem in slot 1 combined with a lexical stem in slot 5. Throughout this paper, classifier stems are boxed while lexical stems are underlined. Classifier stems form

---

[1] The original overview on the relative ordering of bound morphemes within the morphological word in Murrinhpatha in Nordlinger and Mansfield (2021) includes three more suffixal positions: slots 7 and 9 include incorporated adverbials, while slot 10 marks the position for serialised classifiers. Since none of these morphemes is relevant for the phenomenon under discussion nor for the examples in this paper, I decided to omit these slots in table 1 for reasons of clarity and space.

a closed class, consisting of 38 distinct subparadigms (Nordlinger 2015, Mansfield 2019). The majority of predicates require both a classifier stem and an uninflected lexical stem. While a few classifier stems can function as standalone verbs without a lexical stem, lexical stems can never appear in the verb without a classifier stem (Nordlinger and Mansfield 2021). The example in (1) illustrates the interaction of the bipartite predicate in Murrinhpatha. The predicate which roughly parallels the English predicate ‘to tear’ is formed by combining an uninflected lexical stem *rartal* with a specific form of the classifier stem subparadigm 14 ‘slash’ which matches the subject and tense information.[2]

(1) Classifier and lexical stems (Nordlinger and Mansfield 2021: 3)
pam-ngintha-nu-ma-rartal
3SG.slash.NFUT-DU-REFL-APPL-tear
‘The two (non-siblings) will tear it (the cloth) from each other.’

Nordlinger and Mansfield (2021) discuss a thrilling alternation of the classifier stem in relation to the position of the dual marker *ngintha*. A relevant minimal example illustrating this alternation is given in (2). In (2a), the predicate roughly matching the English predicate ‘to see’ consists of the uninflected lexical stem *ngkardu* and the 1SG form of the classifier stem paradigm ‘see’, which is illustrated in table 2. Since the subject of (2a) is 1DU, there is an additional dual marker *ngintha* which is realized to the right of the classifier stem. The 3SG object is unmarked. In (2b), in contrast, there is an overt object affix encoding the 2SG object. In this context, the dual marker *ngintha* appears at the right edge of the word. In addition, the classifier stem does not appear in its 1SG form *ba*, but rather in its dual form *nguba*.[3]

---

[2] Throughout this paper, I will make use of the following abbreviations: 1 = first person; 2 = second person; 3 = third person; APPL = applicative; CAUS = causative; CL = verb class; DC = daucal; DU = dual; FEM = feminine; FUT = future; IND = indicative; IRR = irrealis; MASC = masculine; NFUT = non-future; NPST = non-past; OBJ = object; OBL = oblique; PC = paucal; PFV = perfective; PL = plural; PST = past; REC = reciprocal; REFL = reflexive; SG = singular; SUBJ = subject; TAM = tense/aspect/mood

[3] A recurrent comment touches the question whether *ngu-* could be considered to be a prefix to the singular stem. However, the morphological similarity between the singular stem and the dual stem is a coincidence of the ‘see’ paradigm in table 2 and does not occur in other paradigms, which show exactly the same alternation.

(2) Allomorphy of the classifier stem (Nordlinger and Mansfield 2021: 8)

a. ba-ngintha-ngkardu-nu
see.1SG.SUBJ.IRR-DU-see-FUT
'We (dual non-sibling) will see him / her.'

b. nguba-nhi-ngkardu-nu-ngintha
see.1DC.SUBJ.IRR-2SG.OBJ-see-FUT–DU
'We (dual non-sibling) will see you.'

| | | NFUT | IRR | PST | PST.IRR |
|---|---|---|---|---|---|
| SG | 1 | bam | ba | be | be |
| | 2 | dam | da | de | de |
| | 3 | bam | ba | be | be |
| PL | 1 | ngubam | nguba | ngube | ngube |
| | 2 | nubam | nuba | nube | nube |
| | 3 | pubam/kubam | kuba/puba | pube | pube |
| DC | 1 | | nguba | ngube | ngube |
| | 2 | | nuba | nube | nube |
| | 3 | | kuba/puba | pube | pube |

Table 2: Paradigm of classifier stem *ba* 'to affect, see' (Mansfield 2019: 249)

In summary, the placement of the dual marker *ngintha* and the form of the classifier stem depend on whether an overt object marker is present. With a covert 3SG object, *ngintha* appears next to the classifier stem, which is in its singular form in this context. However, when an overt object marker is used, *ngintha* attaches to the right end of the word, while the classifier stem appears in its dual form. Thus, the pattern in (2b) looks like an instance of multiple exponence of dual and a discontinuous dependency between the classifier stem and the dual marker *ngintha*, two phenomena typically associated with templatic morphology (Nordlinger 2010). Nordlinger and Mansfield (2021) argue that these changes in form and position suggest the existence of position classes, where the dual marker and object affixes compete for the same position to the right of the classifier stem. However, in this paper, I will explain the relationship between the form of the classifier stem and the position of *ngintha* without relying on the concept of position classes as a fundamental component of morphological theory. Instead, I will examine the phonological features associated with the placement of *ngintha* in section 2 arguing that the phonological properties uncover a cyclic structure of the word in Murrinhpatha. In section 3, I investigate the distribution of number exponents in order to infer assumptions about the internal morphological structure of the number feature and hence, the featural specifications of the number exponents. In section 4, I will elaborate on the assumptions of the Stratal Optimality Theory framework (Kiparsky 2000, Bermúdez-Otero 2016) that I adopt in my analysis. Specifically, I assume that the placement of *ngintha* follows from the interaction of independently motivated morphological constraints rather than from a competition for a specific position class. In section 5, I demonstrate that the constraint interaction causes suppression of *ngintha* in the presence of an overt object marker. Consequently, a more specific form of the classifier stem is selected by the morphological grammar to optimize feature realization. In section 6.1, I will show how my analysis captures the distribution of number exponents. In my analysis, I assume that *ngintha* may attach at a later morphophonological domain to realize features of the input since it is stratally underspecified. However, this is a lexical property of *ngintha* rather than a general property of Murrinhpatha. In section, 6.2, I provide further evidence that the stratal unboundedness of *ngintha* is independent of its suppression at the first cycle. Overall, my paper provides a new view on patterns where morphemes display a different phonological behavior in the context of other exponents. In section 6.3, I discuss how my analysis can potentially be extended to more cases of delayed exponence.

## 2. Phonological Properties of Murrinhpatha Morphemes

In Murrinhpatha, the phonological behavior of a bound morpheme is determined by its position within the verbal complex. Put simply, we can predict the phonological processes that apply to a particular morpheme based on its position. Mansfield (2017) notes that the position of an affix affects the assignment of word stress and interacts with compensatory lengthening of monomoraic roots. Specifically, prosodic words in Murrinhpatha must consists of at least two morae. In (3a), the word is assumed to have an underlying form of */ke/*. Since short vowels are typically assumed to be monomoraic, */ke/* would violate the minimum quantity of having at least two morae. Therefore, the vowel of the syllable is lengthened to satisfy the bimoraicity condition. In

(3b), the word consists of a monosyllabic classifier stem and an object suffix. Like the noun root in (3a), the classifier stem is a monomoraic CV syllable. However, unlike (3a), the vowel of the classifier stem is not lengthened in (3b). This suggests that the presence of the object marker is taken into account for the bimoraicity requirement on prosodic words. Nevertheless, this generalization does not hold for all affixes. Example (3c) demonstrates that some affixes do not prevent compensatory lengthening. The vowel of the monosyllabic classifier stem */ti/* in (3c) is lengthened despite the presence of another moraic future affix. Mansfield (2017) concludes that the absence of compensatory lengthening indicates that a given affix belongs to the same phonological domain as the classifier stem, whereas compensatory lengthening of the root vowel in (3c) suggests that the future affix *nu* does not belong to the same phonological domain as the classifier stem.

(3) Minimum quantity and phonological levels (Mansfield 2017: 362)

a. kéː
   'nerite shell'

b. $\boxed{\text{ná}}$-nge
   say.2SG.IRR-3SG.FEM.OBJ
   'tell her'

c. $\boxed{\text{tíː}}$-nu
   sit.2SG.IRR-FUT
   'you will sit'

Mansfield (2017) further notes that this domain coincides with the domain of stress assignment. In short, word stress is assigned to the penultimate syllable of the domain relevant for the bimoraicity condition. That being said, it follows that monosyllabic affixes that prevent compensatory lengthening interact with word stress, whereas monosyllabic affixes whose presence does not prevent compensatory lengthening are irrelevant for word stress assignment. This is exemplified in (4), where the phonological domain relevant for bimoraicity and word stress assignment is indicated by square brackets and word stress is indicated by an acute accent.

(4) Word stress and phonological levels (Mansfield 2017: 362, 366, 368)

a. [páta] [wuɾiní-ŋa]-đa
good go.SG.PST-3SG.FEM.OBL-PST
'He was good to her.'

b. [pumam-nga-páta]-ngintha-pibim
use.hands.3PL.NFUT-1SG.OBL-make-DU-IMPFV
'the two of them are making it for me'

In (4a), the first word *pata* fulfills the bimoraicity condition and assigns word stress to its penultimate syllable. The second prosodic word of the sentence consists of a classifier stem, an oblique object marker, and a PST marker. As shown in the examples in (3), object and oblique object markers prevent compensatory lengthening (see (3b)), while TAM markers do not, as in (3c). Example (4a) strikingly shows that word stress falls on the penultimate syllable of the domain including the oblique object marker *ŋa*, but excluding the TAM marker *đa*. In (4b), the lexical stem *pata* receives word stress on its penultimate syllable, thus illustrating that the domain relevant for word stress spans from the classifier stem to the lexical stem and includes all affixes attaching between those two, while affixes attaching further right than the lexical stem are always outside the word stress domain. Table 3 integrates these insights and provides an overview of the morphemes within the verbal complex and their phonological domains.

| **Slot 1** | **Slot 2** | **Slot 3** | **Slot 4** | **Slot 5** | **Slot 6** | **Slot 8** |
|---|---|---|---|---|---|---|
| Classifier stem (portmanteau w. SUBJ and TAM) | SUBJ number<br>OBJ marker<br>OBL marker | REFL/ REC | incorporated body part/ APPL | lexical stem | TAM | number (SUBJ or OBJ) |

domain for stress assignment / minimum quantity condition

Table 3: The verbal complex and phonological domains

This conclusion makes interesting predictions for the dual marker *ngintha*. As shown in the previous section, *ngintha* appears *before* the lexical stem in the absence of an overt object marker, but *after* the lexical stem whenever an overt object marker is present. The examples in (5a) and (5b) illustrate that the placement of *ngintha* correlates with its phonological behavior. In example (5a), there is no overt object marker and *ngintha* receives word stress. In (5b), however, an overt oblique object marker is realized next to the classifier stem with the consequence that *ngintha* is realized after the lexical stem. In this case, word stress falls on the penultimate syllable of the coverb which clearly shows that *ngintha* is outside the word stress domain.

(5) Word stress and phonological levels (Mansfield 2017: 362, 366, 368)

a. [ piɾim -ngíntha]
stand.3SG.NFUT-DU
'the two of them are standing'

b. [ pumam -nga-páta]-ngintha-pibim
use.hands.3PL.NFUT-1SG.OBL-make-DU-IMPFV
'the two of them are making it for me'

In summary, Mansfield (2017) clearly shows that the behavior of affixes offers evidence for distinct phonological domains and that the placement of *ngintha* is closely related to its phonological properties. The presence of overt object markers does not simply cause a reordering of the dual marker *ngintha* but also affects its concatenation within a different phonological domain. This implies that the prosodic word in Murrinhpatha is layered, and that its cyclic structure is significant in explaining the behavior of *ngintha*. However, morphological theories that assume a flat, templatic structure of words, such as Nordlinger (2010), fail to account for this insight. In the following section, I will discuss how number information is scattered among different morphemes to find out more about the featural specifications of these affixes.

## 3. The Distribution of Number Exponents

Murrinhpatha exhibits another unique, morphological feature in which number information is dispersed among multiple morphemes located in different positions within the verbal complex. Specifically, information on subject number is conveyed through three different positions: first, it is part of the portmanteau classifier stems. Second, additional number affixes can attach to the right of the classifier stem, thus belonging to the domain relevant for word stress assignment (slot 2 in table 3). Third, number affixes can be found in positions after the lexical stem, and hence, outside of the word stress domain (slot 8 in table 3). I will refer to the former group of number markers as *inner number affixes* and to the latter group as *outer number affixes*.

I follow Mansfield (2017, 2019) in assuming that the distinction between the two groups is based entirely on their phonological behavior, with inner number affixes affecting word stress assignment and outer number affixes being invisible to it. Crucially, the number value of a morphological form results from combinations of these three types of exponents. The attested combinations are listed in figure 1 for IRR classifier stems and in figure 2 for NFUT classifier stems. As already mentioned in section 1, the leftmost position is always occupied by the classifier stem. Hence, it is the only exponent of subject number present in all number contexts.

In the case of IRR classifier stems, there are three different forms: singular, daucal and plural.[4] The singular form of the classifier stem is interpreted as singular when it appears without any other number exponent, but it can also be combined with the dual marker *ngintha* in the inner position to refer to exactly two entities that are not siblings. The plural form of IRR classifier stems does not occur with other number markers and is used to refer to plural entities. The daucal form of the classifier stem, which is used in both dual and paucal contexts, is combined with either the dual marker *ngintha* or the paucal marker *ngime* to refer to dual non-sibling entities and paucal entities, respectively.[5] If the daucal classifier stem appears without any additional number affixes, it is used to refer to dual sibling entities. It should also be noted that the number system morphologically represents sibling relationships, which indicates the significant cultural significance of classificatory siblinghood.

The illustration in 1 shows that each number value is realized by exactly one combination of number exponents. However, the alternation of the placement of *ngintha* in the presence of overt object makers yields two possible realizations for dual non-sibling contexts. In the absence of overt object markers, the singular classifier stem is combined with *ngintha* in the inner position. When overt object markers are present, however, this number value is realized by the daucal classifier stem and *ngintha* in the outer position. The

---

[4] The observant reader will notice that the DC form is morphologically indistinct from the PL form. This syncretism appears in other classifier stem paradigms, as well. However, there exist a number of subparadigms in which the two forms come in different shapes, thus justifying the distinction.

[5] Note that the difference between paucal and plural is partially about the quantity of the entities referred to, but probably also about recognizable reference. Specifically, the paucal is typically used when the reference can be recognized, while the plural is used to refer to non-specific referents (Blythe 2009, Mansfield 2019).

distribution of number exponents in figure 1 raises the question of whether SG, DU.SIBLING and PL contexts are realized only by features encoded in the classifier stem or in combination with phonological empty affixes. Concerning this question, I assume that the number value is realized by features on the classifier stem only and crucially, without features in phonologically empty affixes. The reason for this assumption is basically that these phonologically empty affixes are used in a variety of semantically distinct number contexts (singular, dual sibling, and plural) and can therefore not be assumed to form a natural class. Following this assumption, the featural specification of the classifier stem exponents can directly be inferred from the contexts in which they do not occur with other number exponents.

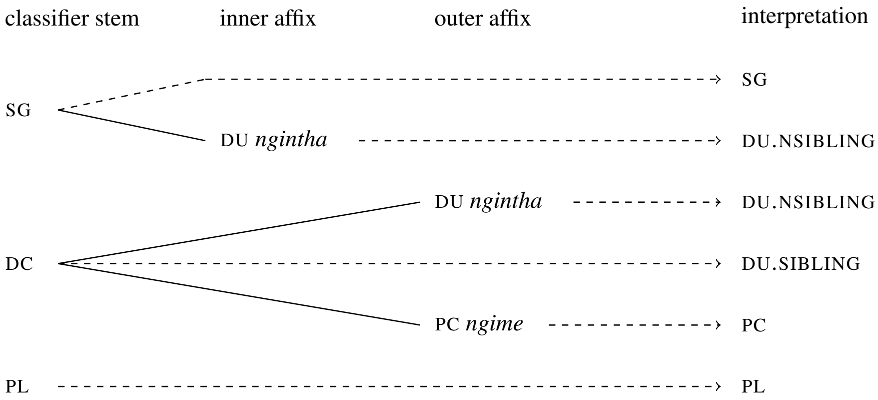


Figure 1: Distribution of SUBJ number in IRR stems (Mansfield 2019: 143)

Figure 2 illustrates the distribution of number exponents in combinations with NFUT classifier stems. Unlike IRR classifier stems, NFUT stems do not have morphologically distinct daucal forms. Instead, paucal and dual sibling contexts are expressed through the use of an inner number affix *ka* which combines with plural classifier stems. This suggests that the daucal is a specific form of a broader number category I will refer to as *non-singular*.

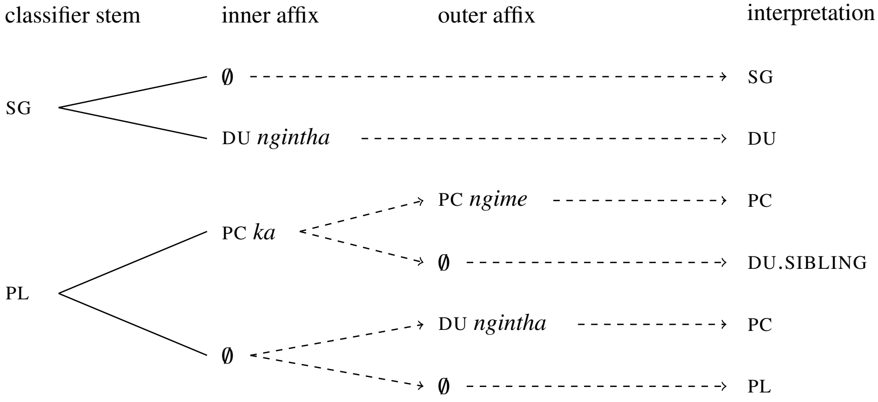


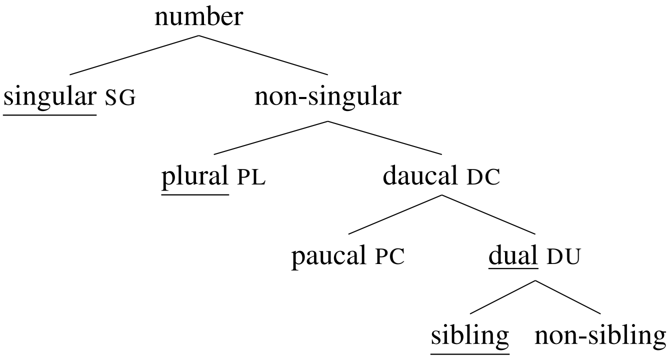
Figure 2: Distribution of SUBJ number in NFUT stems (Mansfield 2019: 143)

Drawing on our generalizations of the distribution of exponents, we can make inferences about the featural composition of morphological number and the specifications of the exponents. My conclusions about the complex number resolution patterns (illustrated in figure 1 and figure 2) suggest a feature geometry for morphological number as shown in (6). Specifically, the existence of only two distinct classifier forms in NFUT paradigms implies a primary division of number into singular and non-singular entities. When a PL classifier stem is used without additional number exponents, it refers to plural entities, indicating that the default interpretation of the non-singular category is plural. However, the non-singular category can also be divided into the daucal subcategory, which further splits into dual and paucal. Siblinghood is only reflected morphologically in dual contexts, indicating that it is a subcategory of dual. The fact that paucal IRR classifier stems refer to dual sibling entities in the absence of additional number exponents suggests that dual is the default interpretation of daucal, and sibling is the default interpretation of dual.

Technically, I propose that morphological number is represented by a set of privative features that are in a dependency relation to each other. Daughter nodes entail the presence of their mother nodes, following the logic of Harley and Ritter (2002). For example, the feature [non-sibling] entails the presence of [dual], [daucal], and [non-singular]. Put simply, [non-sibling] can only be realized in the presence of [dual], [daucal], and [non-singular]. Furthermore, it is technically excluded that a number value comprises two sister nodes.

In the absence of a daughter node, the default interpretation of the mother node is active. Specifically, a feature [non-singular] will be interpreted as [plural] in the absence of a [daucal] feature. In (6), the default interpretation of a mother node is indicated by underlining the respective daughter node. As a consequence, there are two different morphological possibilities for the default values. A default value can either be inferred if only its mother node is realized by a feature or its feature can be spelled out on an exponent. Due to this featural composition, the different number contexts differ in their morphological specificity with dual non-sibling being the most specific number context.

(6) Number specification in Murrinh-Patha



Based on the morphological structure of number in (6) and the distribution of the number exponents in the different contexts, I further infer the following featural specifications of the different exponents. Crucially, I assume that the singular classifier stem does not carry any number features. Rather the singular interpretation is inferred through the default interpretation of number. The plural classifier stem realizes only the feature [non-singular] since it can be combined with paucal markers in NFUT contexts. Crucially, the most specific number context – dual non-sibling – is realized by a SG classifier stem and *ngintha* only. Since I have already established that the SG classifier stem does not realize any number features, it follows automatically that *ngintha* realizes [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING]. The featural specifications of number exponents in IRR contexts are shown in Figure 3, which also demonstrates that each combination of exponents corresponds to the minimal featural representation of each number context. For instance, the paucal context requires three features: [non-singular] and [plural] are represented in combination in the DC classifier stem, while [paucal] is represented by the distinct outer number affix *ngime*.

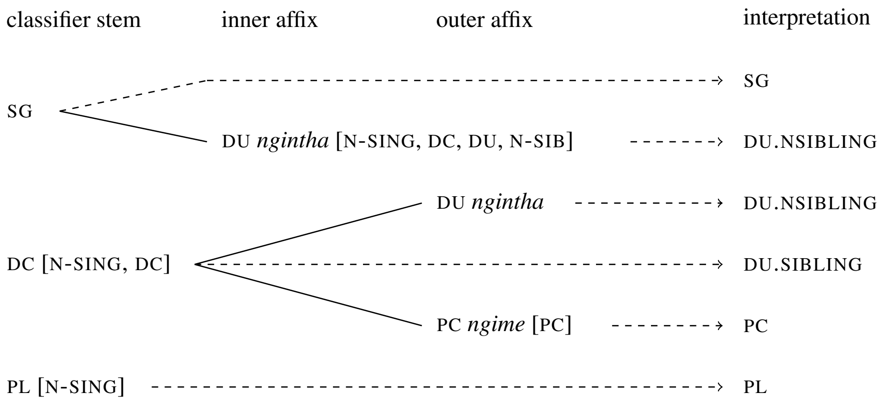


Figure 3: Featural specification of number exponents in IRR classifiers stems

Figure 3 further shows that combination of the DC classifier stem and *ngintha* as an outer number affix is exceptional, since the features [non-singular] and [daucal] are realized twice in this context. Hence, it is the only number context which is not minimally represented by morphological features. In the following two sections, I will connect the featural specifications of the number exponents to the observation that prosodic words in Murrinhpatha are cyclic in order to explain the exceptional phonological and morphological patterning of *ngintha*.

## 4. Background Assumptions

In section 2, I have demonstrated that the phonological correlates of morphemes serve as a window into the cyclic structure of the prosodic word in Murrinhpatha. Specifically, the prosodic domain relevant for word stress assignment spans from the classifier stem at the left edge of the word to the lexical stems, with all affixes following the lexical stem being invisible for stress assignment. In this paper, I implement the cyclic structure of the word by assuming that affixes are concatenated at different morphophonological strata, following the ideas of *Stratal Optimality Theory* (StratOT) (Kiparsky 2000, Bermúdez-Otero 2011). StratOT is a derivational version of *Standard Parallel Optimality Theory* (SPOT) (Prince and Smolensky 1993), and is based on assumptions similar to those posited by *Lexical Phonology and Morphology* (Kiparsky 1982*a*). Just as SPOT, StratOT pursues the idea that the grammar of Human language consists of a set of *violable, rankable and universal constraints*. The grammars of each individual language results from an individual ranking of these constraints. A core difference of StratOT is the division of labor into several different cyclic domains. A concrete suggestion with respect to the number of domains comes from Bermúdez-Otero (2011), who assumes three different morpho-phonological domains:[6]

1. the *stem-level*
2. the *word-level*
3. the *phrase-level*

An important assumption by StratOT is that morphological derivations are accompanied by cycles of phonological optimization such that the morphological component of the grammar and the phonological component of the grammar are interleaved. After each stratum, bracket erasure takes place, which renders morphological structure inaccessible to further cycles. Bracket erasure is a mechanism introduced by Pesetsky (1979) (referring to Chomsky and Halle 1968) and relates to the process of making morphological boundaries invisible to phonological or morphological rules at the end of a cyclic domain. Consequently, neither phonological nor morphological rules can make reference to these boundaries. In this work, I assume that only the morpheme boundaries are deleted, while the grammar still has access to the morphosyntactic information realized in a previous stratum. In other words, a morphologically complex word, e.g. a root plus its affixes,

---

[6] A recurrent question in StratalOT is how the grammar determines at which stratum an affix enters the morphological structure. As for Murrinhpatha, the phonological behaviour of the individual morphemes clearly reveals the stratum it belongs to. While it would be highly desirable if affixes belonging to the same stratum would also form a natural class with respect to their morphosyntactic properties, this is not a technical necessity. Rather, it is commonly assumed that it is specified in the lexical entry of each affix at which stratum it enters the optimizing derivation (Bermúdez-Otero 2011, 2016, 2019).

is treated as a morphologically simplex word after bracket erasure. Thus, access to morphological boundaries is only possible within a cycle. Put simply, StratalOT answers the non-trivial question of morphological sensitivity in phonology by restricting this access to morphological structure by phonology to smaller subdomains. The exact architecture of the cyclic model of the morpho-phonology interface I adopt is illustrated in figure 4.

| | |
|---|---|
| morphological optimization<br>phonological optimization | stem-level |
| *bracket erasure* | |
| morphological optimization<br>phonological optimization | word-level |
| *bracket erasure* | |
| morphological optimization<br>phonological optimization | phrase-level |
| *bracket erasure* | |

Figure 4: Assumed architecture of the morpho-phonology interface

In this paper, I assume that two strata suffice to explain the phenomenon under discussion. Specifically, I assume that the word stress domain corresponds to the *stem-level*, while affixes attaching outside the stress domain belong to the *word-level*. Example (7) illustrates how these assumptions relate to the exceptional placement of *ngintha*. In the absence of overt object markers, *ngintha* is concatenated at the stem-level, as in (7a). However, when an overt object marker is present, as in (7b), *ngintha* attaches at the word-level.

(7) Anomalous placement of *ngintha* (Nordlinger and Mansfield 2021: 8)

a. [[$\boxed{\text{ba}}$-ngintha-∅-$\underline{\text{ngkárdu}}$]$_{\text{stem}}$-nu]$_{\text{word}}$
see.1SG.SUBJ-DU-3SG.OBJ-see-FUT
'We (dual non-sibling) will see him / her.'

b. [[$\boxed{\text{nguba}}$-nhi-$\underline{\text{ngkárdu}}$]$_{\text{stem}}$-nu-ngintha]$_{\text{word}}$
see.1DC.SUBJ-2SG.OBJ-see-FUT–DU
'We (dual non-sibling) will see you'

Moreover, the dispersion of number information across different number exponents allows us to draw conclusions about the featural structure of morphological number, as well as the featural specifications of the exponents. Taking their phonological properties and their morphological position into account, we can now determine the featural specification as well as the stratum a morpheme belongs to. This information is summarized in table 4 for each affix relevant for the discussion. Following Harley and Ritter (2002), I assume that 1st and 2nd person are realized using privative person features, while the realization of 3rd person does not involve features and is inferred through default interpretation. The minimal pair in (7) involves two different classifier stem forms, both of which refer to 1st person subjects. As concluded above, singular classifier stems do not comprise any number feature, while the daucal stem carries the features [NON-SINGULAR] and [DAUCAL]. Hence, the featural specifications for the two classifier stems are [1, SUBJECT] for *ba* and [1, SUBJECT, NON-SINGULAR, DAUCAL] for *nguba*. I further assume that the 3rd person object in (7a) is realized by a covert object marker which has the feature [OBJECT], while the 2nd person object marker *nhi* comes with the specification [2, OBJECT]. The final stem-level affix is the number affix *ka*, which combines with NFUT classifier stems and carries the feature [DAUCAL]. Two different types of affixes belong to the word-level in Murrinhpatha. First, all TAM affixes attach at this level, like the [FUTURE] suffix *nu*. Second, some number affixes belong to this stratum, such as the [PAUCAL] suffix *ngime*. Note that the illustration in table 4 reveals that Murrinhpatha has no morphological possibility to realize the feature [PAUCAL] at stem-level. Rather, its realization is delayed until the word-level. In the previous section, I argued that the dual marker *ngintha* has to be specified for the features [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING], as it combines with the singular stem in the featurally most specific dual non-sibling context. In order to capture the observation that it occurs on both stem-level and word-level, I assume that *ngintha* is *underspecified* with respect to the stratum it belongs to, and may attach at any stratum, an analytical option previously made by Kiparsky (2015).[7]

---

[7] Note that this assumption is not problematic for the *Cyclic Principle* (see Chomsky 1965, Perlmutter and Soames 1979 and Müller 2023, this volume), given here in (i), which states that an operation has to be carried out as early as possible. In fact, I will show that *ngintha* has to be concatenated as early as possible, as long as the context for its realization is given. Hence, the realization of *ngintha* in a later cyclic domain does not pose a problem for the Cyclic Principle, since the context for the rule to apply is not given in the first domain.

(i) Cyclic Principle (Chomsky 1965, Perlmutter and Soames 1979)
When two operations can be carried out, where one applies to the cyclic domain $D_X$

| **Stratum** | **Category** | **Specification** | **Form** |
|---|---|---|---|
| Stem | [CL.STEM] | [1, SUBJECT] | *ba* |
| | [CL.STEM] | [1, SUBJECT, NON-SINGULAR, DAUCAL] | *nguba* |
| | [LX.STEM] | 'to see' | *ngkardu* |
| | [OBJ] | [2, OBJECT] | *nhi* |
| | [OBJ] | [OBJ] | ∅ |
| | | [DAUCAL] | *ka* |
| Word | [TAM] | [FUTURE] | *nu* |
| | | [PAUCAL] | *ngime* |
| unspecified | [SUBJ] | [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING] | *ngintha* |

Table 4: Murrinh-Patha affixes divided into strata

To illustrate how my analysis couched in StratOT derives the peculiar placement of *ngintha*, let me assume that the verb root comes with a list of contextual features that need to be realized by morphological exponents in an optimal way. This list is then checked against the available affixes at each stratum. To ensure that the morphological grammar on a given stratum concatenates only the affixes that are lexically affiliated with it, I assume that the GEN function accesses the lexical entries of the morphemes, in which the stratal specification is stored as a diacritic. Thus, GEN restricts possible output forms to those containing only morphemes with the correct stratal specification. In this paper, I remain agnostic about the origin of these features. Since the core of my analysis rests on the interaction of violable constraints, my analysis is compatible with presyntactic morphological theories based on Optimality Theory (Prince and Smolensky 1993), e.g. Müller (2020) or postsyntactic theories combining OT and Distributed Morphology, like Trommer (2001, 2003), Rolle (2020). To derive the patterns in (7), let us assume that the verbal complex comes with the input features in (8), since it concatenates a classifier stem, a lexical stem, an object marker, and a TAM exponent. I follow the notation introduced by Müller (2020) in using the • symbol to mark features that neeed to be expressed in a morphological word.

(8) Input feature set: V, [•CL.STEM•], [•LX.STEM•], [•TAM•], [•OBJ•]

---

and the other applies to the cyclic domain $D_{X\text{-}1}$ included in $D_X$, then the latter is applied first.

These input features are the same for both (7a) and (7b), yet the sentences differ with respect to the features of the arguments that need to be realized. Hence, there are also input feature sets belonging to the arguments of the sentence. The feature sets for (7a) are listed in (9a), while the feature sets of the arguments in (7b) are listed in (9b)

(9) a. SUBJ: [SUBJECT, 1, NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING]
OBJ: [OBJECT] for (7a)
b. SUBJ: [SUBJECT, 1, NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING]
OBJ: [2, OBJECT] for (7b)

Previous work by Trommer (2003, 2008), Crysmann and Bonami (2016) and Müller (2020) has highlighted that the mapping between input features and output morphological forms is regulated by rules on morphological well-formedness. In this paper, I follow Trommer (2003, 2008) and Müller (2020) by implementing these morphological rules as violable constraints in Optimality Theory. An exhaustive list of constraints is given in (10). M(AX)(F) constraints are crucial, since they ensure that each feature of the input F is realized by an exponent in the output. M(AX)(ARG)$_{\text{SUBJ}}$ and M(AX)(ARG)$_{\text{OBJ}}$ are specific versions of M(AX) relating to the argument input feature sets. All M(AX) receive a violation mark for each feature in the input which is not realized by an exponent in the output.

In addition, there are constraints regulating the relative position of certain categories within a morphological word. To this end, Trommer (2003, 2008) observes that person information is typically aligned to the left edge of the word, while number exponents tend to be realized at the right edge of the word. These crosslinguistic tendencies are captured by two constraints which are violated whenever another exponent intervenes between the left edge of a word and an exponent of [Person] (L $\Leftarrow$ PERS(ON)) or the right edge of the word and an exponent realizing [Number] (NUM(BER) $\Rightarrow$ R), respectively. In addition, the markedness constraint *M(ULTIPLE) E(XPONENCE)$_{\text{F}}$ is violated if a feature of the input is realized more than once, thus preventing multiple exponence. Finally, the constraint COH(ERENCE) ensures that features belonging to the same feature set, i.e. the argument feature sets, are realized in adjacency to each other. In this respect, it is irrelevant if the features of the shared feature set are expressed by one and the same exponent or by two different, adjacent exponents. It will only be violated if another exponent which is not part of the shared feature set intervenes.

(10) a. L ⇐ PERS(ON): (Trommer 2003)
Assign * for each exponent between exponents of [Person] and the left edge of the word.
b. M(AX)(F): (Trommer 2008, Müller 2020)
Assign * for each feature [F] of the input if it is not realized on an exponent in the output.
c. M(AX)(ARG)$_{\text{SUBJ}}$:
Assign * for each feature [F] of the subject argument if it is not realized on an exponent in the output.
d. M(AX)(ARG)$_{\text{OBJ}}$:
Assign * for each feature [F] of the object argument if it is not realized on an exponent in the output.
e. *M(ULTIPLE) E(XPONENCE)$_{\text{F}}$:
Assign * for each feature F which is realised by more than one exponent.
f. COH(ERENCE): adapted from Trommer (2008), Müller (2020)
Assign * for each exponent that intervenes between two exponents realizing features from the same feature set in the input.
g. NUM(BER) ⇒ R: (Trommer 2003)
Assign * for each exponent between exponents of [Number] and the right edge of the word.

In contrast to SPOT, the ranking of constraints is only fixed within a stratum. Between the strata, re-ranking may apply. This assumption is based on the observation that certain phonological rules apply only to certain subdomains, suggesting that the ranking of the constraints may differ from one stratum to the other. In the following, I will show how the anomalous positioning of *ngintha* follows from the constraint-driven interaction of the different exponents. Put shortly, my analysis is couched in StratOT and implements the following generalizations:

1. Both objects markers and inner number markers are subject to morphological rules that require them to be a realized in adjacency to the

classifier stem. First, L ⇐ PERS(ON) ensures that object exponents carrying [Person] information are realized at the left edge of the word. Second, COH(ERENCE) requires exponents realizing features from the same feature set in adjacency to each other. Hence, both affixes preferably occupy the position to the direct right of the classifier stem which always occupies the leftmost position in the word.

2. In the presence of both overt object markers and inner number affixes, preference is given to the former.
3. Since *ngintha* cannot be concatenated in its designated position, highly ranked placement constraints suppress its realization in the stem-level.
4. In order to realize as many input features as possible, a featurally more specific form of the classifier stem is selected to minimize violations of M(AX)(ARG)$_{\text{SUBJ}}$, thus explaining the different form of the classifier stem.
5. Since *ngintha* is not strictly bounded to the stem-level, its realization is delayed until the word-level.

## 5. A StratalOT Analysis of Murrinhpatha

Having set the technical preliminaries in the previous section, let me now explain in detail how the peculiar placement of *ngintha* and its phonological correlates can be derived from the interaction of well-established morphological constraints. In this endeavor, let us first consider example (11), repeated from (7a), where *ngintha* attaches to the right of the classifier stem in its singular form.

(11) [[ba-ngintha-Ø-ngkárdu]$_{\text{stem}}$-nu]$_{\text{word}}$
see.1SG.SUBJ-DU-3SG.OBJ-see-FUT
'We (du. n-sib.) will see him/her.' (Nordlinger and Mansfield 2021: 8)

The relevant tableau is given in (12). The input to this derivation is the root $\sqrt{see}$, a set of contextual features, as well as the feature sets for the subject and the object argument. As noted earlier in this paper, classifier stems are always portmanteau morphemes carrying subject features. To this end, I assume that the root is an abstract pointer $\sqrt{see}$ to the respective classifier stem paradigm.

That is, it refers to a set of inflected forms of one and the same classifier stem paradigm, but does not choose a specific form of that paradigm. Note that this assumption is unproblematic in StratOT since the root is not a cyclic domain and does not undergo phonological optimization. The contextual features for (7a) are [•CL.STEM•], [•LX.STEM•], [•TAM•] and [•OBJ•], hence giving rise to the constraints MAX(CL.STEM), MAX(LX.STEM), MAX(OBJ) and MAX(TAM). Since all exponents realizing TAM are concatenated at word-level, MAX(TAM) is omitted from the tableau in (12), since it cannot be satisfied at stem-level. However, MAX(CL.STEM), MAX(LX.STEM) and MAX(OBJ) are high-ranked and ensure that a classifier stem, a lexical stem and an object marker are concatenated. As an example, candidate b. is ruled out since it does not comprise a lexical stem, thus fatally violating MAX(LX.STEM). The remaining constraints make sure that the argument feature sets are realized in an optimal way. Recall that the subject is a 1DU NON-SIBLING argument, thus requiring the features [SUBJECT, 1, NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING], while the 3rd person object only requires [OBJECT]. The output form of candidate a. splits the features of the subject onto two different morphemes: the 1st person singular form classifier stem form *ba* realizes [1] and [SUBJECT], whereas *ngintha* spells out the remaining number features [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING]. The candidates c. and d., both of which lack the dual marker *ngintha*, cannot become optimal, since they fatally violate $\text{MAX(ARG)}_{\text{SUBJ}}$, which ensures that the subject feature set is exhaustively realized. In candidate a., each feature is realized exactly once, thus avoiding violations of *MULTIPLE EXPONENCE. Candidate e. with the 1st daucal classifier stem, however, is ruled out since the two features [NON-SINGULAR] and [DAUCAL] are realized twice. Moreover, candidate a. does not violate COHERENCE, since the two exponents realizing features of the subject feature set are adjacent and not interrupted by different exponents. Most crucially, the object marker does not violate L⇐PERS although it is not at the left edge of the word, since it does not include any person features and is therefore not subject to this constraint. Note that candidate f., in which *ngintha* attaches as an outer affix, is ruled out as it violates COHERENCE due to two intervening morphemes. Put shortly, candidate a. does not violate any constraint on morphological well-formedness and becomes optimal.

(12) Morphological optimization at stem-level, (11)

| | |
|---|---|
| *ba* | [CL.STEM], [1, SUBJ], stem-level |
| *nguba* | [CL.STEM], [1, SUBJ, N-SING, DC], stem-level |
| *ngkardu* | [LX.STEM], ’to see’, stem-level |
| ∅ | [OBJ], [OBJ], stem-level |
| *nu* | [TAM], [FUT], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| $\sqrt{\textit{see}}$, [●CL.STEM●], [●LX.STEM●], [●TAM●], [●OBJ●]<br>SUBJ: [SUBJ, 1, N-SING, DC, DU, N-SIB]<br>OBJ: [OBJECT] | | | M(CL.STEM) | M(LX.STEM) | M(OBJ) | *ME | M(ARG)$_{\text{OBJ}}$ | L⇐PERS | COH | M(ARG)$_{\text{SUBJ}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| a. | ☞ | ba$_{[1,\ \text{SUBJ}]}$-ngintha$_{[\text{N-SING, DC, DU, N-SIB}]}$-∅$_{[\text{OBJ}]}$-ngkardu | | | | | | | | |
| b. | | ba$_{[1,\ \text{SUBJ}]}$-ngintha$_{[\text{N-SING, DC, DU, N-SIB}]}$-∅$_{[\text{OBJ}]}$ | | ∗! | | | | | | |
| c. | | ba$_{[1,\ \text{SUBJ}]}$-∅$_{[\text{OBJ}]}$-ngkardu | | | | | | | | ∗!∗∗∗ |
| d. | | nguba$_{[1,\ \text{SUBJ, N-SING, DC}]}$-∅$_{[\text{OBJ}]}$-ngkardu | | | | | | | | ∗!∗∗ |
| e. | | nguba$_{[1,\ \text{SUBJ, N-SING, DC}]}$-∅$_{[\text{OBJ}]}$-ngintha$_{[\text{N-SING, DC, DU, N-SIB}]}$ | | | | ∗!∗ | | | | |
| f. | | ba$_{[1,\ \text{SUBJ}]}$-∅$_{[\text{OBJ}]}$-ngkardu-ngintha$_{[\text{N-SING, DC, DU, N-SIB}]}$ | | | | | | | ∗!∗ | |

The output of the morphological optimization at stem-level is *ba-ngintha-ngkardu*, which is then taken to the phonological component of the stem-level for further phonological optimization. Note that the output form contains exactly those affixes with are relevant for word stress assignment. Concretely, it contains the classifier stem, inner affixes and the lexical verb, but crucially, no external affixes. Within the phonological component of the stem-level, stress assignment and compensatory lengthening apply. After this computation, bracket erasure takes place and deletes morpheme boundaries. The next step of the derivation takes place in the morphological component at word-level. At this step of the derivation, the grammar has access to the output of the stem-level *banginthangkardu*, the remaining contextual feature [●TAM●], as well as word-level and underspecified affixes. The morphological derivation at word-level is illustrated in (14). Most contextual features have already been satisfied at the previous stratum, except for [●TAM●], which can only be satisfied at word-level, since all TAM affixes are word-level affixes. In order to anchor the input at the left edge of the word, I use the high-ranked ALIGNMENT constraint L⇐V which ensures that all affixes attached at word-level will end up in a suffixal position. The concrete definition of L⇐V is given in (13).

(13) L ⇐V Assign * for each exponent between the base and the left edge of the word.

Since bracket erasure has taken place, the input *banginthangkardu* is treated as a morphologically simplex exponent of the features [SUBJECT, 1, NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING] and [OBJECT] as word-level. Hence, the constraint NUM⇒R is violated once by candidate b. as the TAM exponent *nu* intervenes between *banginthangkardu* and the right edge of the word. Nonetheless, candidate b. becomes optimal since candidate a. does not include any TAM marker and violates the high-ranked MAX(TAM), while candidate c. violates the general suffixing constraint L ⇐V. After this step of morphological optimization, the optimal candidate *banginthangkardu-nu* enters the phonological component of the word-level for further optimization.

(14) Morphological optimization at word-level, (11)

| | |
|---|---|
| *ba* | [CL.STEM], [1, SUBJ], stem-level |
| *nguba* | [CL.STEM], [1, SUBJ, N-SING, DC], stem-level |
| *ngkardu* | [LX.STEM], 'to see', stem-level |
| ∅ | [OBJ], [OBJ], stem-level |
| *nu* | [TAM], [FUT], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| *banginthangkardu*, [●TAM●]<br>SUBJ: [~~SUBJ, 1, N-SING, DC, DU, N-SIB~~]<br>OBJ: [~~OBJECT~~] | | M(TAM) | L⇐V | NUM⇒R | *ME | *COH |
|---|---|---|---|---|---|---|
| a. | banginthangkardu[SUBJ, 1, N-SING, DC, DU, N-SIB, OBJ] | *! | | | | |
| b. ☞ | banginthangkardu[SUBJ, 1, N-SING, DC, DU, N-SIB, OBJ]-nu | | | * | | |
| c. | nu-banginthangkardu[SUBJ, 1, N-SING, DC, DU, N-SIB, OBJ] | | *! | | | |

Let us now turn to example (7b), repeated here in (15), where *ngintha* is concatenated externally and the classifier stem appears in its daucal form.

(15) [[$\boxed{\text{nguba}}$-nhi-$\underline{\text{ngkárdu}}$]$_{\text{stem}}$-nu-ngintha]$_{\text{word}}$
see.1DC.SUBJ-2SG.OBJ-see-FUT–DU
‘We (du. n-sibl.) will see you’ (Nordlinger and Mansfield 2021: 8)

Recall that Nordlinger and Mansfield (2021) argue that the pattern in (15) suggests the existence of *position classes* as primitive entities of morphological theory. Since *ngintha* is blocked in the position after the classifier stem in (15) in the presence of an overt object marker, Nordlinger and Mansfield (2021) assume that both *ngintha* and the object markers compete for the same position class. Moreover, the different shape of the classifier stem in (15) is taken to be evidence for position-conditioned allomorphy where a different allomorph of the classifier stem is chosen in the presence of an object marker. Put shortly, Nordlinger and Mansfield (2021) suggest that position classes exist as abstract elements in the morphological grammar, because there are rules that refer to them. In what follows, I will demonstrate that the model forwarded in this paper derives the pattern in (15) without assuming position classes. Instead, I argue that the placement of *ngintha* follows from the interaction of well-established morphological constraints and the cyclic structure of the word. The tableau illustrating this derivation is provided in (16).

(16) Morphological optimization at stem-level, (15)

| | |
|---|---|
| *ba* | [CL.STEM], [1, SUBJ], stem-level |
| *nguba* | [CL.STEM], [1, SUBJ, N-SING, DC], stem-level |
| *ngkardu* | [LX.STEM], ’to see’, stem-level |
| *nhi* | [OBJ], [2, OBJ], stem-level |
| *nu* | [TAM], [FUT], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| | | $\sqrt{\textit{see}}$, [•CL.STEM•], [•CL.STEM•], [•TAM•], [•OBJ•]<br>SUBJ: [SUBJ, 1, N-SING, DC, DU, N-SIB]<br>OBJ: [OBJ, 2] | M(CL.STEM) | M(LX.STEM) | M(OBJ) | *ME | M(ARG)$_{\text{OBJ}}$ | L⇐PERS | COH | M(ARG)$_{\text{SUBJ}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| a. | | ba$_{\text{[1, SUBJ]}}$-nhi$_{\text{[2, OBJ]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-ngkardu | | | | | | | ∗! | |
| b. | | ba$_{\text{[1, SUBJ]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-nhi$_{\text{[2, OBJ]}}$-ngkardu | | | | | | ∗! | | |
| c. | | ba$_{\text{[1, SUBJ]}}$-nhi$_{\text{[2, OBJ]}}$-ngkardu | | | | | | | | ∗∗∗∗! |
| d. | | ba$_{\text{[1, SUBJ]}}$-nhi$_{\text{[2, OBJ]}}$-ngkardu-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$ | | | | | | | ∗!∗ | |
| e. | ☞ | nguba$_{\text{[1, SUBJ, N-SING, DC]}}$-nhi$_{\text{[2, OBJ]}}$-ngkardu | | | | | | | | ∗∗ |
| f. | | nguba$_{\text{[1, SUBJ, N-SING, DC]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-nhi$_{\text{[2, obj]}}$-ngkardu | | | | ∗!∗ | | ∗∗ | | |
| g. | | ba$_{\text{[1, SUBJ]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-ngkardu | | | ∗! | | | | | |
| h. | | ba$_{\text{[1, SUBJ]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-$\emptyset_{\text{[OBJ]}}$-ngkardu | | | | | ∗! | | | |

In contrast to example (11), there is an overt object marker *nhi* in (15), which comes with the featural specification [2, OBJECT]. Thus, the constraint L⇐PERS becomes active, thus shifting the marker to the right of the finite stem.[8] In the previous derivation in (12), the constraint remained inactive since the covert object marker does not spell out person features. In the context of *nhi*, however, L⇐PERS now causes a competition between the object marker and *ngintha* for the position to the right of the classifier, thus following the empirical intuition by Nordlinger and Mansfield (2021). In my analysis, however, the competition arises from morphotactic constraints on positioning preferences rather than from position classes. Specifically, candidate b. replicates the order of affixes that became optimal in (12), yet fatally violates L⇐PERS since the overt object marker *nhi* carries person features. However, shifting the dual marker *ngintha* to the right of the object marker, as in candidates a. or d.,

---

[8] Since both the classifier stem and the object marker carry person features, an additional constraint would be needed to determine which affix will end up in the left-most position. This could be achieved with a high-ranked L⇐V, as in (14), which generates structures in which the classifier stem is always to the left.

causes fatal violations of COHERENCE. Not realizing an object marker at all in candidate g. or choosing a different object marker in candidate h. in order to avoid violations of L⇐PERS or COHERENCE is not possible, either, due to the high-ranked constraint MAX(OBJ) and MAX(ARG)$_{\text{OBJ}}$. Since *ngintha* cannot be realized in its preferred position, the grammar chooses to not concatenate the marker at stem-level. Since *ngintha* realized the input features [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING], non-realization of the markers yields four violations of the constraint MAX(ARG)$_{\text{SUBJ}}$, thus ruling out candidate c. However, the grammar still has the option to choose the more specific classifier stem *nguba*, which is specified for [1, SUBJECT, NON-SINGULAR, DAUCAL], in contrast to *ba*. In (12), the choice of *nguba* was blocked since simultaneous realization of *nguba* and *ngintha* creates a violation of *ME. In the derivation in (16), choosing *nguba* becomes now the preferred option since non-realization of *ngintha* prevents a violation of *ME and creates only two violations of MAX(ARG)$_{\text{SUBJ}}$. Thus, candidate (e), which includes *nguba*, but excludes *ngintha*, becomes optimal.

The optimal output form *nguba-nhi-ngkardu* is taken to the phonological component of stem-level, where the evaluation of the minimum quantity condition and stress assignment apply. After this step, computation at stem-level is complete, bracket erasure takes place and the output is shifted to word-level, illustrated in (17). In contrast to the derivation in (14), no exponent is realizing the input features [DUAL, NON-SIBLING] yet, which caused two violations of M(ARG)$_{\text{SUBJ}}$ at stem-level. As a consequence, the grammar will try to find a matching exponent and a TAM exponent. Since *ngintha* is unbounded with respect to the stratum it attaches to, it is concatenated now at word-level and will therefore be realized outside the word stress domain. Since Murrinhpatha does not only have the underspecified *ngintha* number exponent, but also a word-level only number marker *ngime*, I believe that the grammar at this level still requires access to the input feature structure to find the matching exponent. Thus, the constraints M(ARG)$_{\text{SUBJ}}$ and *ME are still active, however, the relative ranking of these constraints has changed. At word-level, *ME is ranked below M(ARG)$_{\text{SUBJ}}$. As a consequence, the grammar will favor candidates in which all input features are realized. The high-ranked MAX constraints require that both a number and a TAM exponent are concatenated at this step, thus ruling out candidate a. in (17). Again, there is a constraint L⇐V ensuring that all affixes added at this level are suffixes, therefore excluding candidate d. At this point of the derivation, NUM⇒R (Trommer 2001, 2003, 2008) becomes active and regulates the relative ranking of TAM and *ngintha*. Candidate b., which surfaces in (2b) is therefore successfully predicted to become the optimal candidate.[9]

(17) Morphological optimization at word-level, (15)

| | |
|---|---|
| *ba* | [CL.STEM], [1, SUBJ], stem-level |
| *nguba* | [CL.STEM], [1, SUBJ, N-SINGR, DC], stem-level |
| *ngkardu* | [LX.STEM], 'to see', stem-level |
| *nhi* | [OBJ], [2, OBJ], stem-level |
| *nu* | [TAM], [FUT], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| *ngubanhingkardu*, [●TAM●]<br>SUBJ: [~~SUBJ, 1, N-SING, DC~~, DU, N-SIB]<br>OBJ: [~~OBJECT, 2~~] | | | M(TAM) | L⇐V | NUM⇒R | *ME | *COH |
|---|---|---|---|---|---|---|---|
| a. | | ngubanhingkardu$_{\text{[SUBJ, 1, N-SING, DC, OBJ, 2]}}$ | *! | * | | | |
| b. | ☞ | ngubanhingkardu$_{\text{[SUBJ, 1, N-SING, DC, OBJ, 2]}}$-nu-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$ | | | ** | ** | *** |
| c. | | ngubanhingkardu$_{\text{[SUBJ, 1, N-SING, DC, OBJ, 2]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-nu | | | ***! | ** | |
| d. | | nu-ngubanhingkardu$_{\text{[SUBJ, 1, N-SING, DC, OBJ, 2]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$ | | *! | * | ** | |

In the analysis suggested in this paper, the anomalous placement of *ngintha* is an instance of *myopia* in grammar. On the surface, the pattern in (15) seems like overexponence of the features [NON-SINGULAR] and [DAUCAL]. However, the phonological properties of the word reveal that the apparent overexponence results from cyclicity in the style of Kiparsky (1982*a*,*b*) (see also the discussion about different versions of cyclicity in Müller 2023, this volume). First, *ngintha* is suppressed in the presence of an overt object marker. Due to the non-realization of *ngintha* at stem-level, the grammar selects a featurally more specific classifier stem. Second, *ngintha* is underspecified with respect to the stratum at which it attaches, and is therefore realized at word-level. Crucially, the grammar at stem-level cannot anticipate that *ngintha* will be realized in a later step. Hence, the stem-level grammar chooses the optimal option for its domain although this results in overexponence at a later domain.

---

[9] It is worth mentioning that the relative order of the TAM exponents and the number exponents are word-level are rather flexible. Thus, it remains unclear whether the relative order should be regulated by morphotactic constraints or whether the order is subject to free variation.

In this paper, I follow Nordlinger and Mansfield (2021) in assuming that there is in fact a competition between overt object markers and *ngintha* for the position to the right of the classifier stem. However, the theoretical device triggering the competition are constraints that are based on crosslinguistic preferences of the realization of person and number markers rather than position classes. My analysis is superior in two more aspects. First, it heavily relies on cyclicity, thus naturally explaining the phonological behavior of *ngintha* in both positions. Second, while Nordlinger and Mansfield (2021) exploit position classes to explain *that* the classifier stem alternates, the analysis forwarded in this paper can also explain *why* it shifts to the paucal form. It is worth mentioning that affixation itself is only limited by *MULTIPLE EXPONENCE and other constraints on morphological well-formedness. As long as these constraints are obeyed, affixation may in principle apply without any restriction on the maximum number of affixes. In this respect, this work differs from a position-class analysis in the style of Nordlinger and Mansfield (2021), but also from other morphological analyses of affixation, such as Wunderlich and Fabri (1995), Wunderlich (1997), Ortmann (1999), Aissen (2003), Don and Blom (2006), Müller (2020).

In the remainder of this paper, I will first elaborate on how the interaction of morphological constraints can neatly explain the distribution of object number exponents in section 6.1. Section 6.2 emphasizes that the anomalous placement of *ngintha* is an interplay of suppression, reranking, and stratal underspecification, and hence, a lexical property of *ngintha*. Moreover, the placement of *ngintha* and its phonological correlates are connected to cyclicity, universal morphological constraints, and stratal underspecification. Since these properties can be assumed to exist in other languages, as well, the analysis suggests that we should find more patterns of delayed realization in other languages than Murrinhpatha. To this end, I discuss Umlaut in Sinhala in section 6.3.

## 6. Discussion

### 6.1. An Extension to Object Number

In the previous section, we have seen that the realization of *ngintha* is delayed since it cannot be realized in its preferred position to the right of the classifier stem. Specifically, the intervention of an object marker causes a fatal violation of COHERENCE, which ensures that exponents belonging to the same argument appear in adjacency. These assumptions predict that *ngintha* should be allowed to appear after the object marker when it spells out features of the object argument, since this would not cause a violation of COHERENCE. The examples in (18), however, illustrate that this prediction is not borne out. In both subexamples, the features of the object are realized by means of three separate markers. In (18a), there is an inner, pronominal affix *ngan*, a daucal marker *ngku* and an outer paucal, feminine affix *ngime*. We already encountered the paucal exponent *ngime* when discussing the distribution of subject number exponents in figures 1 and 2 and concluded that it always appears as an outer affix. Hence, nothing contradicts the assumption that *ngime* is a word-level affix, thus explaining that it appears as an outer affix after the lexical stem in (18a). However, this assumption cannot be extended to *ngintha* in (18b). For this example, we would expect *ngintha* to appear after the object pronominal *ngan*, since *ngintha* is stratally unbounded and does not violate COHERENCE when it marks object features. Put shortly, the placement of *ngintha* as an outer affix in (18b) seems unexpected and contradicts the analysis suggested in the previous section.

(18) Distribution of object number participants (Mansfield 2019: 150f)

a. [[$\boxed{\text{pan}}$-ngan-ngku-$\underline{\text{bat}}$]$_{\text{stem}}$-ngime]$_{\text{word}}$
slash.3SG.NFUT-1PL.OBJ-DC.OBJ-hit-PC.FEM
'she hit us (paucal, female)'

b. [[$\boxed{\text{pan}}$-ngan-ngku-$\underline{\text{bat}}$]$_{\text{stem}}$-ngintha]$_{\text{word}}$
slash.3SG.NFUT-1PL.OBJ-DC.OBJ-hit-DU
'she hit us (dual, female)'

Let us delve deeper into this pattern and determine the featural specifications of the number exponents by examining the distribution of object number exponents, which is given in figure 5. The leftmost column refers to the possible forms of the pronominal affix, which is the 1PL form *ngan* in (18a) and (18b). In contrast to the classifier stem forms in figures 1 and 2, the singular never combines with other number exponents. As a consequence, I assume that the singular object pronominal is specified for [SINGULAR] whereas singular classifier stems are unspecified. Thus, Murrinhpatha exploits two different realization strategies for the singular category: it is inferred by default in the singular classifier stems but realized by the features [SINGULAR] in the object pronominals. Without any additional number exponents, the PL forms refer to plural entities and can therefore be assumed to be specified for [NON-SINGULAR]. Example (18a) demonstrates that the plural pronominal may combine with an additional daucal marker *ngku*. In the absence of additional outer number exponents, the combination of a plural pronominal and daucal *ngku* refers to dual, sibling referents. Thus, I infer that plural pronominals are only specified for [NON-SINGULAR] whereas *ngku* is specified for [DAUCAL]. For *ngime* and *ngintha*, we have already established the featural specifications [PAUCAL] and [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING], respectively. Recall that *ngime* and *ngintha* differ in their stratal affiliation. While *ngintha* attaches at both phonological domains depending on the morphological context, *ngime* ever only attaches at word-level. Hence, we have to assume that it is a word-level affix.

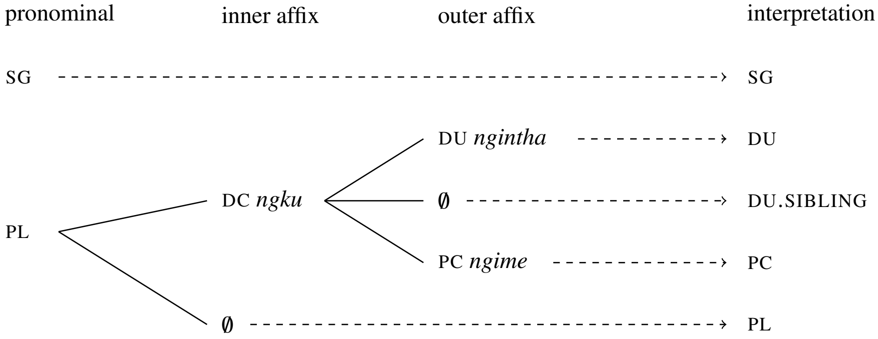


Figure 5: Distribution of OBJ number exponents (Mansfield 2019: 143)

That being said, we can now list the featural specifications and stratal affiliations of the exponents in (18) in 5. Crucially, the plural pronominal is specified for [NON-SINGULAR], while the singular pronominal is [SINGULAR].

| **Stratum** | **Category** | **Specification** | **Form** |
|---|---|---|---|
| Stem | [CL.STEM] | [SUBJECT] | *pan* |
| | [LX.STEM] | ‘to hit’ | *bat* |
| | [OBJ] | [2, OBJECT] | *nhi* |
| | [OBJ] | [1, OBJECT] | *ngi* |
| | [OBJ] | [OBJ] | ∅ |
| | [OBJ] | [1, OBJECT, NON-SINGULAR] | *ngan* |
| | | [DAUCAL] | *ngku* |
| Word | | [PAUCAL] | *ngime* |
| unspecified | [SUBJ] | [NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING] | *ngintha* |

Table 5: Murrinh-Patha affixes divided into strata

In the following, I will show that the featural specifications of the number exponents explain why *ngintha* is realized as an outer affix despite referring to the object argument in (18b). The input to the derivation in (19) is the contextual features [•CL.STEM•], [•LX.STEM•] and [•OBJ•], as well as the feature sets of the arguments. Since the subject is 3SG, the subject argument set only requires the feature [SUBJECT], which will automatically be realized by concatenating a classifier stem. The object argument is 1DU, hence requiring the features [1, OBJECT, NON-SINGULAR, DAUCAL, DUAL, NON-SIBLING]. Note that there is no contextual feature [•TAM•] and therefore no constraint MAX(TAM), since the syntactic context does not require it. The tableau in (19) allows the following observation: since the object pronominal is already specified for [NON-SINGULAR], simultaneous realization of *ngintha* will always result in a violation of *MULTIPLE EXPONENCE.

(19) Morphological optimization at stem-level, (18b)

| | |
|---|---|
| *pan* | [CL.STEM], [SUBJ], stem-level |
| *bat* | [LX.STEM], 'hit', stem-level |
| *ngan* | [OBJ], [1, OBJ, N-SING], stem-level |
| ∅ | [OBJ], [OBJ], stem-level |
| *ngku* | [DC], stem-level |
| *ngime* | [PC], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| $\sqrt{\textit{slash}}$, [●CL.STEM●], [●LX.STEM●], [●OBJ●]<br>SUBJ: [SUBJECT]<br>OBJ: [1, OBJ, N-SING, DC, DU, N-SIBL] | | | M(CL.STEM) | M(LX.STEM) | M(OBJ) | *ME | M(ARG)$_{\text{OBJ}}$ | L⇐PERS | COH | M(ARG)$_{\text{SUBJ}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| a. | ☞ | pan-ngan$_{\text{[1, OBJ, N-SING]}}$-ngku$_{\text{[DC]}}$-bat | | | | | ∗∗ | | | |
| b. | | pan-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-bat | | | ∗! | | ∗∗ | | | |
| c. | | pan-ngan$_{\text{[1, OBJ, N-SING]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-bat | | | | ∗! | | | | |
| d. | | pan-ngku$_{\text{[DC]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-bat | | | ∗! | ∗ | ∗∗ | ∗ | | |
| e. | | pan-ngan$_{\text{[1, OBJ, N-SING]}}$-bat | | | | | ∗∗∗! | | | |
| f. | | pan-ngan$_{\text{[1, OBJ, N-SING]}}$-ngku$_{\text{[DC]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-bat | | | | ∗!∗ | | | | |

Since *MULTIPLE EXPONENCE is higher ranked than M(ARG)$_{\text{OBJ}}$, these violations are fatal for candidates c., d. and f., all of which contain an object pronominal and *ngintha*. Note also that switching to the singular pronominal is not possible, since the [SINGULAR] feature on *ngi* in 5 contradicts the required [NON-SINGULAR] feature of the object.[10] Deleting the object pronominal altogether, however, creates a fatal violation of M(OBJ) in candidate b. The only remaining option for the stem-level grammar is to not realize *ngintha* at stem-level. This causes three violations of M(ARG)$_{\text{OBJ}}$ in candidate e. The grammar has the option to minimize the violations of M(ARG)$_{\text{OBJ}}$ by concatenating the daucal marker *ngku* in candidate a, which becomes optimal.

From this point of the derivation, the computation proceeds as already described in section 5. The optimal candidate of the derivation in (19),

---

[10] In Murrinhpatha, it seems that exponents with non-matching features never surface. Recent work by Privizentseva (2021), however, has shown that conflicting features do not necessarily cause the derivation to crash. To this end, it can either be assumed that surface forms with conflicting exponents are ruled out due to high-ranked constraints on morphological matching, or excluded from the generated set of output forms by a restriction on GEN.

*pan-ngan-ngku-bat* passes the phonological computation at stem-level, after which bracket erasure takes place. Afterwards, *pannganngkubat* enters the morphological derivation at word-level, which is illustrated in (20). Recall that the word-level includes re-ranking of *MULTIPLE EXPONENCE and M(ARG$_{\text{Obj}}$). Consequently, the optimal output candidate of the derivation in (20) is candidate b., in which *ngintha* serves to realize the remaining features [DUAL, NON-SIBLING] of the object feature set despite violating *MULTIPLE EXPONENCE, while candidate a. which avoids a violation of *MULTIPLE EXPONENCE by not concatenating another number exponent is ruled out since it fatally violates M(ARG$_{\text{Obj}}$).

(20) Morphological optimization at word-level, (18b)

| | |
|---|---|
| *pan* | [CL.STEM], [SUBJ], stem-level |
| *bat* | [LX.STEM], ’hit’, stem-level |
| *ngan* | [OBJ], [1, OBJ, N-SING], stem-level |
| ∅ | [OBJ], [OBJ], stem-level |
| *ngku* | [DC], stem-level |
| *ngime* | [PC], word-level |
| *ngintha* | [N-SING, DC, DU, N-SIB], unspecified |

| *pannganngkubat*<br>SUBJ: [~~SUBJECT~~]<br>OBJ: [~~1, OBJ, N-SING, DC~~, DU, N-SIB] | | | L⇐V | M(ARG$_{\text{Obj}}$) | NUM⇒R | *ME | *COH |
|---|---|---|---|---|---|---|---|
| a. | | pannganngkubat$_{\text{[SUBJ, OBJ, 1, N-SING, DC]}}$ | | *!* | | | |
| b. | ☞ | pannganngkubat$_{\text{[SUBJ, OBJ, 1, N-SING, DC]}}$-ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$ | | | * | ** | |
| c. | | ngintha$_{\text{[N-SING, DC, DU, N-SIB]}}$-pannganngkubat$_{\text{[SUBJ, OBJ, 1, N-SING, DC]}}$ | *! | | * | ** | |

Put shortly, the analysis forwarded in this paper can also capture the observation that *ngintha* appears as an outer affix when it refers to the object argument. However, the delayed realization of *ngintha* results from a violation of *MULTIPLE EXPONENCE rather than from a violation of COHERENCE.

## 6.2. Morphological Blocking of Stem-Level Affixes

In the analysis I forward in 5, I assume that the grammar at stem-level determines the non-realization of *ngintha* in the context of overt objects. Since *ngintha* is stratally unbounded, it has the chance to be realized at a later level. A core assumption of StratOT is that the stratal affiliation is a lexical property of each affix. In simpler terms, it is a lexical coincidence that *ngintha* can be realized later, which is entirely independent of its suppression at stem-level. This assumption further predicts that stem-level affixes with similar featural specifications would be blocked in the context of overt object markers. Example (21) illustrates that this prediction is in fact borne out. In both subexamples, the subject is 3PC. Recall from the distribution of number exponents in 2 that this context is realized by a combination of the PL classifier stem and an additional daucal affix *ka* in inner position in NFUT contexts. This is exactly the combination that surfaces in example (21a), which does not contain overt object markers. In (21b), however, the presence of an overt object marker *nga* blocks the realization of *ka*, yet the subject is 3PC. In contrast to *ngintha*, however, *ka* is a stem-level affix only and can therefore not be realized at word-level. As a result, the feature [DAUCAL] remains unrealized.

(21) -*ka* as a stem-level affix only (Mansfield 2017)

a. [[ $\boxed{\text{Pumám}}$ -ka]$_{\text{stem}}$-ngime]$_{\text{word}}$.
say.3PL.NFUT-DC.SUBJ-PC.FEM
‘They (paucal) said’

b. dɻáf [[ $\boxed{\text{pumám}}$ -nga]$_{\text{stem}}$-neme]$_{\text{word}}$
draft do.3PL.NFUT-1SG.OBJ-PC.MASC
‘They (paucal) drafted me.’

### 6.3. Another Instance of Delayed Realization: Umlaut in Sinhala

Due to the differential phonological behaviour of *ngintha* in the two possible positions, I treat the placement of *ngintha* as delayed realization due to morphological blocking. Given that the morphological constraints, cyclicity, and stratal underspecification are expected to exist in other languages, as well, my analysis predicts more patterns of delayed concatenation. Specifically, we should find languages in which one and the same affix displays different phonological properties depending on the morphological context of the affix. Such a pattern is found in Sinhala, as exemplified in (22). In this language, certain affixes like the perfective suffix *la* trigger umlaut of the root. In (22a), the root with the underlying form *ad* ‘to pull’ surfaces as *æ* in the context of the perfective suffix *la*. Similarly, the underlying ‘root *bal* ‘to look’ becomes *bæl* in the context of the past suffix *u* in (22b). When a causative suffix intervenes, as in (22c) and (22d), the umlaut-triggering past suffix behaves differently than the perfective suffix. While the past suffix triggers umlaut across the causative in (22d), umlaut is blocked in the context of the causative in (22c). In short, it cannot be assumed that umlaut only applies in strictly local configurations, since it does apply across intervening affixes in (22d). A possible explanation for the blocking of umlaut in (22c) is delayed realization. In similarity to delayed concatenation of *ngintha*, we could assume that the causative blocks concatenation of the perfective marker in the cyclic domain responsible for umlaut. Parallel to *ngintha*, the perfective marker *la* is stratally unbounded and attaches at a later, cyclic domain.

(22) Umlaut in Sinhala (Fenger and Weisser 2022: 5,7)

a. æ-ə-la tie-nə-wa
pull-CL2-PFV be-NPST-IND
'have pulled'

b. bæl-∅-u-wa
look-CL1-PST-IND
'looked'

c. ad-ə-wə-la tie-nə-wa
pull-CL2-CAUS-PFV be-NPST-IND
'have made someone pull'

d. bæl-ə-wə-u-wa
look-CL1-CAUS-PST-IND
'made someone look'

## 7. Conclusion

In this paper, I have discussed and explained the peculiar placement of the dual marker *ngintha* in the morphologically highly complex language Murrinhpatha, in which the presence of overt object markers affects the position of the dual marker *ngintha* and the form of the classifier stem. Specifically, *ngintha* appears to the right of the classifier stem in the absence of overt object markers in (23a) but at the right edge of the word when object markers are overtly realized in (23b). Furthermore, Murrinhpatha uses the singular form of the classifier stem when adjacent to the dual marker in (23a), but the daucal form when followed by the object marker in (23b).

(23) Placement of *ngintha* (Nordlinger and Mansfield 2021: 8)

a. ba-ngintha-ngkardu-nu
see.1SG.SUBJ.IRR-DU-see-FUT
'We (dual non-sibling) will see him / her.'

b. nguba-nhi-ngkardu-nu-ngintha
see.1DC.SUBJ.IRR-2SG.OBJ-see-FUT–DU
'We (dual non-sibling) will see you.'

Nordlinger and Mansfield (2021) have argued that these two phenomena provide evidence for the existence of position classes in morphological theories. Specifically, the authors analyze the alternation of the classifier stem as an instance of *position-dependent allomorphy*, where the form of the classifier stem depends on the morphological content of the following position class. Moreover, Nordlinger and Mansfield (2021) assume that the replacement of *ngintha* follows from its competition with the object marker for the position class to the right of the classifier stem.

In this paper, I tackle this view and illustrate that both phenomena follow from the interaction of universal and violable morphological constraints, the featural specifications of the exponents, and the cyclic structure of the word in Murrinhpatha. To this end, the phonological behaviour of affixes in different positions was discussed in section 2 with the conclusion that the word in Murrinhpatha is separated into two different morphophonological layers.

Section 3 examines the distributions and combinations of the different number exponents in Murrinhpatha, which allowed us to infer the morphological structure of number and the featural specifications of the number exponents.

Section 4 capitalizes on the StratalOT framework adopted in the analysis. StratalOT neatly captures the cyclic structure of the word and the interaction of violable constraints. Crucially, these universal, morphological constraints are based on typological tendencies of the realization of phi features (Trommer 2001). In sum, my assumptions build upon independent evidence, whereas position classes have to be stipulated as primitive entities of morphological theory.

In section 5, I explain how the interaction of constraints and the featural specifications of the exponents explain both the placement of *ngintha* and the alternation of the classifier stem form. Specifically, the position of *ngintha* results from a competition between different morphological constraints, where both object markers and inner number markers are required to attach to the right of the classifier stem. First, L ⇐ PERS(ON) ensures that object exponents carrying [Person] information are realized at the left edge of the word. Second, COH(ERENCE) requires exponents realizing features from the same feature set in adjacency to each other. Since L ⇐ PERS(ON) outranks COH(ERENCE), object markers win the competition and appear to the right of the classifier stem in (23b). Since *ngintha* can no longer be realized in its designated position, it is suppressed at the first morphophonological cycle altogether. As a consequence, a featurally more specific form of the classifier stem is selected to realize as many input features as possible. Thus, the analysis forwarded in this paper does not only explain that the form of the classifier stem changes but also why it changes to the daucal marker. Since *ngintha* is not strictly bounded to the stem-level, its realization is delayed until the word-level.

In the remainder of this paper, I illustrate how my analysis can be extended to object number in section 6.1. Put shortly, the extraordinary placement of *ngintha* follows from suppression at stem-level, the stratal unboundedness of *ngintha*, and constraint reranking, which allows the grammar to delay its realization. Section 6.2 highlights that these factors are independent of each other. Evidence for this claim comes from the paucal marker *ka*, which is suppressed in the very same morphosyntactic context but cannot be concatenated later, thus resulting in deletion of the exponent. This paper opens an entirely new view on patterns where morphemes display a different phonological behavior in the context of other exponents. Section 6.3 illustrates how this generalization can potentially be extended to more cases of delayed exponence.

In sum, I have analyzed a complex morphological pattern by means of a StratOT analysis which rests on independently motivated assumptions and is therefore beneficial to analyses using position classes. Moreover, this paper has shown that studying the phonological properties of affixes provides a window into the morphological structure of the word, which allows us to answer recalcitrant morphological problems.

## References

Aissen, Judith (2003): 'Differential object marking: Iconicity vs. economy', *Natural Language & Linguistic Theory* pp. 435–483.

Bermúdez-Otero, Ricardo (2011): The architecture of grammar and the division of labour in exponence. *In:* J. Trommer, ed., *The morphology and phonology of exponence*. Vol. 41, Oxford University Press, Oxford, pp. 8–83.

Bermúdez-Otero, Ricardo (2016): Stratal phonology. In: *The Routledge handbook of phonological theory*. Routledge, London, pp. 100–134.

Bermúdez-Otero, Ricardo (2019): 'Nonanalytic listing and the theory of the stem level'. Lecture given at Leipzig University.

Blythe, Joe (2009): Doing referring in Murriny Patha conversation. PhD thesis, University of Sydney.

Chomsky, Noam (1965): 'Aspects of the Theory of Syntax'.

Chomsky, Noam and Morris Halle (1968): 'The sound pattern of English.'.

Crysmann, Berthold and Olivier Bonami (2016): 'Variable morphotactics in information-based morphology', *Journal of Linguistics* **52**(2), 311–374.

Don, Jan and Elma Blom (2006): A Constraint-Based Approach to Morphological Neutralization. In: *Linguistics in the Netherlands 2006*. Benjamins, Amsterdam, pp. 78–88.

Fenger, Paula and Philipp Weisser (2022): 'Matching domains: The syntax, morphology, and phonology of the verb in Sinhala'. Manuscript available at https://ling.auf.net/lingbuzz/006932.

Harley, Heidi and Elizabeth Ritter (2002): 'Person and Number in Pronouns: A Feature-Geometric Analysis', *Language* **78**(3), 482–526.

Kiparsky, Paul (1982*a*): Lexical phonology and morphology. *In:* T. L. S. of Korea, ed., *Linguistics in the morning calm*. Hanshin, Seoul, pp. 3–91.

Kiparsky, Paul (1982*b*): Word-formation and the lexicon. *In:* F. Ingemann, ed., *1982 Mid-America Linguistics Conference Papers*. Department of Linguistics, University of Kansas, Lawrence, pp. 1–29.

Kiparsky, Paul (2000): 'Opacity and cyclicity', *The Linguistic Review* **17**(2-4), 351–366.

Kiparsky, Paul (2015): Stratal OT: a synopsis and FAQs. *In:* Y. E. Hsiao and L.-H. Wee, eds, *Capturing phonological shades within and across languages*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 2–44.

Mansfield, John (2017): 'Prosodic words in cyclic derivation: the strange case of Murrinhpatha compound verbs', *Morphology* **27**(3), 359–382.

Mansfield, John (2019): *Murrinhpatha morphology and phonology*. De Gruyter Mouton, Boston/Berlin.

Müller, Gereon (2020): *Inflectional Morphology in Harmonic Serialism*. Advances in Optimality Theory, Equinox, Sheffield.

Müller, Gereon (2023): Challenges for Cyclicity. *In:* M. Privizentseva, F. Andermann and G. Müller, eds, *Linguistische Arbeits Berichte 95*. Institut für Linguistik, Leipzig, pp. 3–72.

Nordlinger, Rachel (2010): 'Verbal morphology in Murrinh-Patha: Evidence for templates', *Morphology* **20**(2), 321–341.

Nordlinger, Rachel (2015): Inflection in Murrinh-Patha. *In:* M. Baerman, ed., *The Oxford Handbook of Inflection*. Oxford University Press, Oxford, pp. 491–519.

Nordlinger, Rachel and John Mansfield (2021): 'Positional dependency in Murrinhpatha: expanding the typology of non-canonical morphotactics', *Linguistics Vanguard* **7**(1). https://doi.org/10.1515/lingvan-2020-0079.

Ortmann, Albert (1999): 'Affix repetition and non-redundancy in inflectional morphology', *Zeitschrift für Sprachwissenschaft* **18**(1), 76–120.

Perlmutter, David M and Scott Soames (1979): *Syntactic argumentation and the structure of English*. Vol. 231, Univ of California Press, Berkeley.

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

Prince, Alan and Paul Smolensky (1993): 'Optimality Theory: Constraint Interaction in Generative Grammar'. Tech. rep. RuCCS-TR-2 at Rutgers University.

Privizentseva, Mariia (2021): 'Mixed agreement in Russian: Gender, declension, and morphological ineffability'. Draft available at https://ling.auf.net/lingbuzz/006086.

Rolle, Nicholas (2020): 'In support of an OT-DM model: Evidence from clitic distribution in Degema serial verb constructions', *Natural Language & Linguistic Theory* **38**, 201–259.

Trommer, Jochen (2001): Distributed optimality. PhD thesis, Universität Potsdam, Potsdam.

Trommer, Jochen (2003): The interaction of morphology and syntax in affix order. *In:* G. Booij and J. van Marle, eds, *Yearbook of Morphology 2002*. Springer, Dordrecht, pp. 283–324.

Trommer, Jochen (2008): 'Coherence in affix order', *Zeitschrift für Sprachwissenschaft* **27**(1), 99–140.

Wunderlich, Dieter (1997): A minimalist model of inflectional morphology. *In:* C. Wilder, M. Bierwisch and H.-M. Gärtner, eds, *The role of economy principles in linguistic theory*. Akademie Verlag, Berlin, pp. 267–298.

Wunderlich, Dieter and Ray Fabri (1995): 'Minimalist morphology: An approach to inflection', *Zeitschrift für Sprachwissenschaft* **14**(2), 236–294.

# The Strict Cycle Condition in Stratal OT

Jochen Trommer

## Abstract

In this paper, I argue that the phonological phenomena traditionally attributed to the Strict Cycle Condition (henceforth SCC, Kean 1974, Mascaró 1976) are better understood as the result of ranked and violable constraints in Stratal Optimality Theory.[1]

## 1. Introduction

A simple example of a potential SCC effect is hiatus resolution in Emai where a word-final vowel is deleted if it is followed by a vowel-initial word (e.g. /kɔ/ 'plant' +/ema/ 'yam' → [kema], Casali 1997:513). This could be captured by a Phrase-Level rule as in (1):

(1) V → Ø /__V

Word-internal hiatus is not repaired in the same way. Thus the noun [oa] 'house' apparently surfaces as such in isolation (not as *[o]) and in contexts where it triggers vowel deletion across words (e.g., /ɔli/ 'the' +/oa/ 'house' → [ɔloa], Casali 1997:512). This asymmetry could be captured in a stratal model of phonology with three stratal domains (Stem Level, Word Level, and Phrase Level) and no stratum-internal cycles by a condition as in (2):

(2) *Strict Cycle Condition:* A phonological process in a given stratal domain *S* applies if and only if its focus and context match material not exclusively contained in a single stratal domain embedded in *S*.

Under the standard assumptions that internal brackets of a stratum are deleted at the point when computation enters a subsequent stratum, at the Phrase Level

---

[1] I will not address here SCC effects where other phonological processes in a given cycle license the application of an SCC-bound alternation, as predicted by the version of the SCC advocated in Mascaró (1976). See Gleim (2023) on different versions of the SCC proposed in the literature.

*Cyclicity*, 239–270
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

the only embedded domains visible for SCC are the Word Level boundaries, as in (3):

(3) $[[\text{ɔli}]_{\text{Word Level}}\ [\text{oa}]_{\text{Word Level}}]_{\text{Phrase Level}}$

Thus (3) correctly predicts that rule (1) is not applied to the string **oa** since both the focus and the context V of the rule are contained in a Word Level domain. On the other hand, (1) applies to the string **io** where **i** (matching the focus-V) is contained in one embedded domain and **o** (matching the context-V) in a different embedded domain (see section 5.2 for an alternative analysis).

The SCC can also be applied to word-internal processes. Thus it is well-known that front vowels in Finnish trigger spirantization ('assibilation') on preceding coronal stops as in (4a). However this happens only across morpheme boundaries (4b):

(4) *Assibilation in Finnish* (Kiparsky 1993)

| | | | | |
|---|---|---|---|---|
| a. | /halut-i/ | → | [halusi] | 'want-PAST' |
| | /halut-a/ | → | [haluta] | 'want-INF' |
| b. | /koti/ | → | [koti], *[kosi] | 'home' |

This asymmetry could be captured by (2) under the assumption that bare roots are Stem-Level domains in a SOT architecture whereas Assibilation applies at the Word Level:

(5) a. $[[\text{halut}]_{\text{Stem Level}}\ \text{i}]_{\text{Word Level}}$ b. $[[\text{koti}]_{\text{Stem Level}}\ ]_{\text{Word Level}}$

The alternative to a general unviolable SCC I will advocate here following van Oostendorp (2008) is that apparent SCC effects follow from specific constraints sensitive to morphological colors.

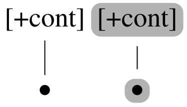
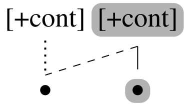
van Oostendorp's crucial observation is that most cases of SCC effect involve feature spreading, and as pointed out by Wolf (2008:329) this holds also for Finnish assibilation which might be interpreted as spreading of [+continuant] from a high front vowel to a left-adjacent [t].

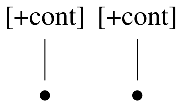
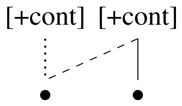
High-ranked ALTERNATION as defined in (6) would block this process morpheme-internally (e.g. in [koti]) since the [+cont] feature of [i] and the root node of [t] have the same morphological color, and spreading would mean that an epenthetic association line links them.

(6) ALTERNATION: If an association line links two elements of color $\alpha$, the line should also have color $\alpha$. (van Oostendorp, 2007:16)

Ranked above the relevant constraint triggering spreading (here: SHARE), this predicts assibilation across a morpheme boundary (7a), but not inside a morpheme (7b) (see below on the encoding of morphemes by color/background shading):

(7) *Finnish Assibilation*

| *Input:* a. | ALT | SHARE |
|---|---|---|
| a.  | | *! |
| ☞ b.  | | |

| *Input:* a. | ALT | SHARE |
|---|---|---|
| ☞ a.  | | * |
| b.  | *! | |

There are both conceptual and empirical reasons to assume that SCC effects are due to specific violable constraints and not to a general inviolable convention. Conceptually there doesn't seem to be a natural way to implement the idea that a process must involve new material in OT. This is because OT lacks a reified notion of process (which are largely equivalent in rule-based phonology where a rule typically captures a process). Consider again the case of Emai. Vowel deletion is an operation of GEN, but vowel deletion by itself (i.e., viewed independently from its trigger) doesn't happen in any reasonable sense across morpheme or word boundaries. It is always the vowel of a single morpheme (and word) which is deleted. One might consider requiring that a general SCC convention should restrict the application not of processes, but of markedness constraints triggering them. However, under the standard assumption that vowel deletion under hiatus is due to the constraint ONSET, this also would not work out for Emai since the relevant ONSET at a word/morpheme boundary as in [ebe ɔna] would also be restricted to a single word/morpheme (the syllable containing [ɔ]).

Empirically, there are many obvious violations of the SCC. An especially well-documented case is American English flapping which changes intervocalic coronal stops after stressed vowels into flaps:

(8) *American English flapping (Kenstowicz 1994:195)*

a. á[ɾ]om cf. a[t]óm-ic c. whát[ɾ] is wrong? cf. whá[t]
b. méé[ɾ]ting cf. méét[t]

Flapping happens across word boundaries, as shown by (8c), hence must be phrasal, but also happens inside single words/morphemes (8a,b), which would violate the SCC if it is a Phrase Level process.

Another case is Arabic vowel insertion which breaks up consonant clusters. It is clearly a phrasal process since it can be bled by a following vowel-initial word, but it applies in single isolated words:

(9) *Arabic vowel insertion: /fihm/ 'understanding' (Kiparsky 2000:352)*

| | |
|---|---|
| a. fíhm il-wálad | 'the boy's understanding' |
| b. fíhim | 'understanding' |
| c. fíhimna | 'our understanding' |

A tonal example is Jita where an underlying H-tone shifts to a following syllable (10c). Again, this must be a phrasal process because it applies across word boundaries (11a,b) and is barred from applying to phrase-final positions (10b). At the same time it can apply in single words as in (10c):

(10) *Jita tone shift (Downing 2014:103)*

| | | |
|---|---|---|
| a. /oku-ljá/ | [oku-ljá] | 'to eat' |
| b. /oku-βóna/ | [oku-βóna] | 'to see' |
| c. /oku-βón-an-a/ | [oku-βon-án-na] | 'to see each other' |

(11) *Jita tone shift (Downing 2014:103)*

a. [oku-βóna iiɲoɲi] 'to see a bird'
(cf. [oku-βóna] 'to see'; iiɲoɲi 'bird'
b. [oku-lja múnó] 'to eat a lot'
(cf. [oku-ljá] 't eat'; munó 'a lot'

Now it is well-known and virtually universally acknowledged in the phonological literature that the SCC cannot be an unconditional restriction on all phonological processes. The general research strategy in Lexical Phonology has been to define a class of processes that are universally subject to the SCC and a complement class that is not. (12) lists the most important hypotheses pursued in this tradition:

(12) *Potential criteria for SCC compliance*

| *SCC-compliant* | *SCC-non-compliant* |
|---|---|
| a. lexical phonology | postlexical phonology |
| b. neutralizing/phonemic | allophonic/non-phonemic |
| c. structure-changing | structure-building |
| d. cyclic | non-cyclic |

Kiparsky (1993) provides a general refutation of most of these claims, and we have already seen some other counterexamples. Thus Emai is a bona fide case of a phrasal phonological process which is SCC-compliant counter to (12a). Vowel epenthesis in Arabic is neutralizing and non-allophonic violating (12b). The Finnish coalescence process discussed by Kiparsky (1993) is clearly structure-changing and cyclic but not subject to the SCC (violating (12c) and (12d)). Note also that in the version of Stratal OT adopted here there is no stratum-internal cyclicity, hence there is no distinction between cyclic and non-cyclic processes.

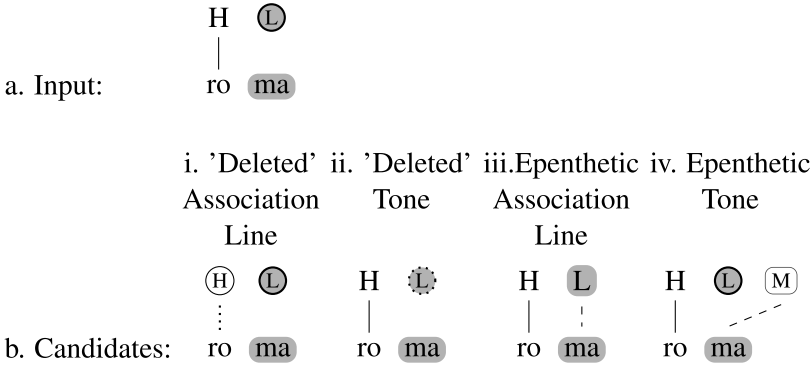
## 2. Theoretical Background

I will adopt a substantially restricted implementation of Optimality Theory – Colored Containment Theory – and a minimal version of Stratal Optimality Theory which only comprises three strata: Stem, Word, and Phrase Level (following Bermúdez-Otero 2018, pace Kiparsky 1982, Rubach 2011).

Colored Containment Theory (Revithiadou 2007, van Oostendorp 2008, Trommer 2011, Paschen 2018) is a conservative extension of the original implementation of OT in Prince and Smolensky (1993) with a more limited set of possible structural changes than Correspondence Theory – restricting them basically to insertion and marking for non-pronunciation – and principled modularity restrictions on the phonology-morphology interface. Crucially, the theory limits access of phonology to morphosyntactic information to 'colors', an encoding of morphemic affiliation, especially useful for autosegmental representations, which I illustrate with a toy example in (13). Thus in the structure in (13a), color (realized here graphically by background shading) identifies the floating Low-tone as part of the same morpheme as the syllable

[ma] (and distinct from [ro] and its H-tone) with a different color, even though they do not form a coherent phonological object, a fact which would be difficult to capture by morpheme boundaries. The second crucial function of color is to distinguish underlying (= morphological = colored) and epenthetic (= non-underlying = colorless) material. Thus the notation used here for the lack of morphological colors, dashing of association lines as in (13b-iii,iv) and boxes for tones (13iv), directly encodes their status as epenthetic material in output representations. The Containment Requirement of the theory states that input structure can never be literally deleted in possible outputs. The representation of deletion is instead by diacritically marking parts of the input as phonetically invisible, graphically indicated by dotting of association lines (13b-i), and circles for floating tones (13b-ii). Crucially, there is no candidate where tones (or segments) would be literally removed from possible output representations. Thus inputs and then changes performed by GEN are fully reconstructable from outputs, obviating input-out comparisons and indices as in Correspondence Theory. Hence, (13b) illustrates *all* possible tonal changes to the input candidate in (13a). Besides full deletion of a tone, changing a tone (say from High to Low), splitting a tone, or tone metathesis are in principle excluded.

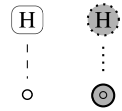
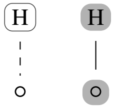
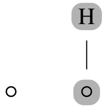
(13) *Autosegmental representations in Colored Containment Theory*



Containment allows optimality-theoretic markedness constraints to still access structure which is floating or marked as phonetically invisible.

This is illustrated in (14) with a constraint which plays an important role in the analysis of dissimilation effects in section 4.1, the ban on adjacent identical melody tones, a version of the well-known Obligatory Contour Principle (Leben 1973, Myers 1997). The general version is abbreviated as OCP τ and its phonetic clone as $\underline{\text{OCP}}$ τ. The input representation (14c) violates neither constraint, whereas (14b), where a melody-H is inserted to associate to the unspecified tonal root node violates both. The crucial contrast between the two versions of the constraint emerges in (14a) where the second tone and its root node are 'deleted' (rendered phonetically invisible). Since the second melody-H is invisible for the phonetic constraint version $\underline{\text{OCP}}$ τ, this is not violated here. In contrast, the generalized constraint clone OCP τ evaluates the full phonological representation and is hence violated by (14a), just like by the visible tone in (14b).

(14) *Phonetic and general constraints: The OCP for melody tones*

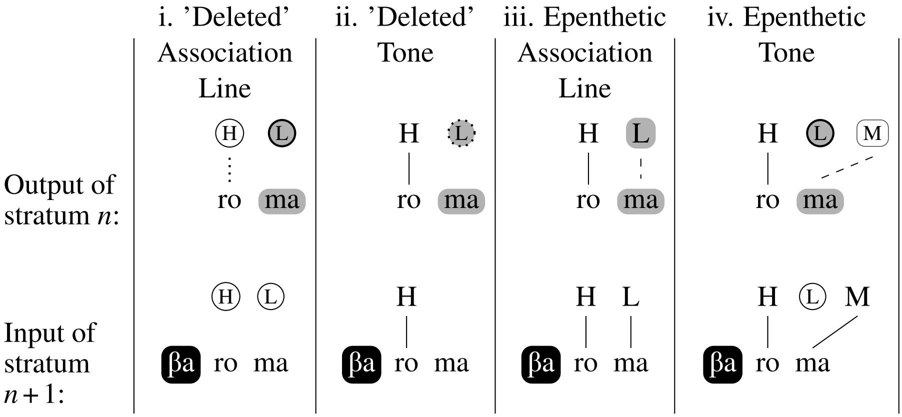
| *Input:* c. | $\underline{\text{OCP}}$ τ | OCP τ |
|---|---|---|
| a.  | | * |
| b.  | * | * |
| c.  | | |

Generalized and phonetic constraint versions will also play a crucial role for Klamath in section 5.3 below.

In contrast to early work in Colored Containment which assumes a single parallel evaluation cycle (van Oostendorp 2005, Revithiadou 2007), I integrate the theory here into **Stratal Optimality Theory** (Kiparsky 2000, Bermúdez-Otero 2018). This raises the question what happens at the transition from one stratum to the next. I assume that there are two natural, but significant processes illustrated with the toy examples from (14b) in (15), assuming that the next stratum adds a new morpheme, [βa]. *Cleanup* removes all material which is marked as phonetically invisible such as the association line in (i) and the Low-tone in (ii) from the representation. This means that

Containment holds for the optimality-theoretic evaluation at a single stratum, but not globally across strata. *Monochromization* assigns a uniform color to all material which is the result of an evaluation at a previous stratum. Thus the two morphemes and the epenthetic Mid tone (and its association line), which can all be differentiated in their morphological status in (15iv) at the output of Stratum *n* all acquire the same color (i..e, behave representationally as a single morpheme) as the input of the next stratum $n+1$ in contrast to [βa] which didn't participate in the earlier evaluation cycle. Monochromization is thus the Colored Containment equivalent of Bracket Erasure in Lexical Phonology and other versions of Stratal OT (Pesetsky 1979, Kiparsky 1982).

(15) *Clean-up and Monochromization ("Bracket Erasure") between strata*



## 3. Reranking of ALTERNATION between Strata: Catalan

The following cases (as others in latter chapters) show that ALTERNATION must be ranked differently in different strata.

### 3.1. Catalan

In Catalan, unstressed high vowels after another vowel become glides. This happens across words and word-internally (16) but is blocked in specific cases word-internally (17):[2]

---

[2] The analysis here is based on the assumption that the descriptive generalizations made in Mascaró (1976) are correct. See Gleim (2023) for critical empirical discussion.

(16) *Catalan diphthongization (Kenstowicz 1994:206 after Mascaró 1976)*

a.[i,u] → [j,w] /V__ (in unstressed syllable

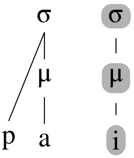
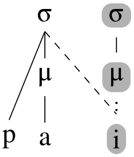
| | | | |
|---|---|---|---|
| b.sál i pá | ‘salt and bread’ | pá j sál | ‘bread and salt’ |
| fɛ́r̃-u | ‘iron’ | dé-w | ‘god’ |
| fɛ́rrous | ‘iron’ | əlʒəbrá-jk | ‘algebraic’ |
| | | rej | ‘king’ |

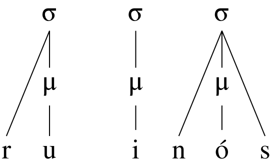
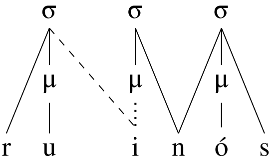
(17) *No word-internal diphthongization (Kenstowicz 1994:206*

| | | | |
|---|---|---|---|
| a. rəím | ‘grape’ | rəim-ɛ́t | diminutive |
| b. ruín-ə | ‘ruin’ | ruin-ós | ‘ruinous’ |

Again the blocking follows straightforwardly from ALTERNATION high-ranked at the Phrase Level:

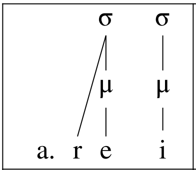
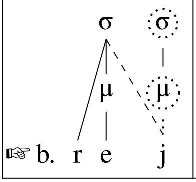
(18) *Catalan Phrase Level evaluations*

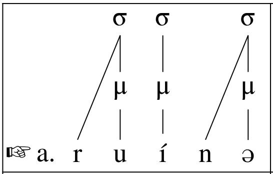
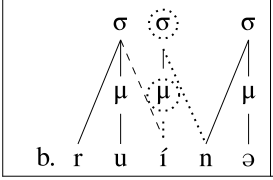
| *Input:* a. | ALT | ONS |
|---|---|---|
| a.  | | *! |
| ☞ b.  | | |

| *Input:* a. | ALT | ONS |
|---|---|---|
| ☞ a.  | | * |
| b.  | *! | |

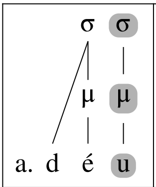
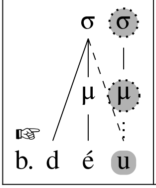
However, diphthongization may happen word-internally and even root-internally (as in [de-w] and [rej]). The decisive factor here is not ALTERNATION, but stress. An underlyingly stressed high vowel cannot glide (in (20ii) stress is represented graphically as an autosegmental grid mark to make explicit its underlying presence):

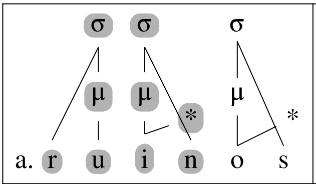
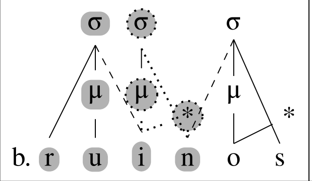
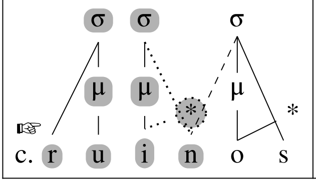
(19) *Catalan Word Level evaluations*

| *Input:* a. | Max V́ | Ons | Alt |
|---|---|---|---|
| a.  | | *! | |
| ☞ b.  | | | * |

| *Input:* a. | Max V́ | Ons | Alt |
|---|---|---|---|
| ☞ a.  | | * | |
| b.  | *! | | * |

(20) *Catalan Word Level evaluations*

| *Input:* a. | Alt | Max V́ | Ons |
|---|---|---|---|
| a.  | | | *! |
| ☞ b.  | | | |

| *Input:* a. | Alt | Max V́ | Ons |
|---|---|---|---|
| a.  | | | **! |
| b.  | | *! | |
| ☞ c.  | | | * |

## 3.2. Kuria

Mora-counting tone in Kuria also nicely illustrates both SCC effects and their violability. The discussion here is based on the more detailed Stratal-OT analysis of these data in Trommer (2023).

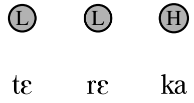
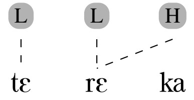
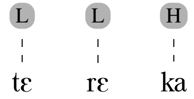
Kuria expresses the Remote Future of verbs by imposing a High tone on the 3rd mora of the stem, as shown in (21) (this tone then spreads further rightwards postlexically):

(21) *Remote Future (μ3) (MMP:254) ‘we will . . . ’*

| | | |
|---|---|---|
| a. 3μ-*Stems*: | n-to-re-[tɛrɛk-á̲] | ‘brew’ |
| b. 4 μ-*Stems* | n-to-re-[teremé̲k-a] | ‘be calm’ |
| | n-to-re-[karaáŋg-a] | ‘fry’ |
| c. 5 μ-*Stems* | n-to-re-[koondó̲kór-a] | ‘uncover’ |
| | n-to-re-[kiriɣí̲ít-a] | ‘scrub’ |
| d. 6 μ-*Stems* | n-to-re-[hootó̲ótér-a] | ‘reassure’ |
| | FOC-1PL-TAM-[√-FV] | |

Following Trommer (2023) I capture this by the morphological melody ⓁⓁⒽ, which associates left-to-right and one-by one to a tonally underspecified verb stem. This is shown in (22)

(22) *Kuria Stem Level (Remote Future/μ3)*

| *Input:* a. | $^{*}{}_{\mathrm{L}}\underline{\mu}_{\mathrm{H}}$ | $\tau \triangleright \mu$ |
|---|---|---|
| a.  | | **!* |
| b.  | *! | |
| ☞ d.  | | |

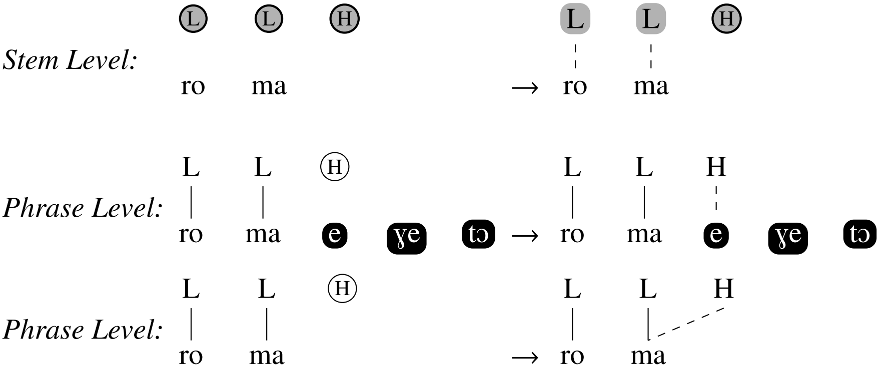
If the verb stem is shorter by one mora than the tone melody, there are different repairs depending on the phrasal context. For verbs which are followed by a direct object, the tone melody is realized on the combination of verb+object:

(23) *Remote Future µ3 on following object (MMP:259)*

a. n-to-re-[rj-a] eɣé̱tɔ́ɔ́kɛ
FOC-1PL-TAM-[eat-FV] banana
'we will eat a banana'

b. n-to-re-[rom-a] é̱ɣétɔ́ɔ́kɛ
FOC-1PL-TAM-[bite-FV] banana
'we will bite a banana'

(24)



On the other hand, if the verb is in phrase-final position, the H tone is realized on the verb itself:

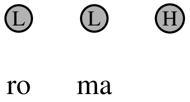
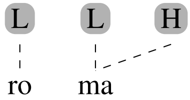
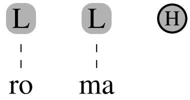
(25) *Remote Future (µ3) (MMP:254) 'we will . . . '*

a. 1µ-*Stems*: n-to-re-[rj-a] 'eat'
n-to-re-[h-a] 'give'

b. 2µ-*Stems*: n-to-re-[rom-ǎ̱] 'bite'
n-to-re-[βun-ǎ̱] 'break'

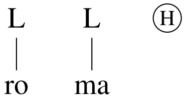
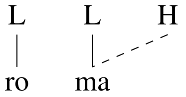
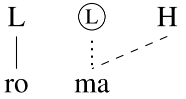
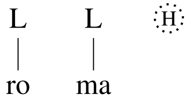
The following tableaux show how this is derived in SOT. At the Stem Level, the final H remains floating to avoid a final contour:

(26) *Kuria Stem Level (Remote Future/μ3)*

| *Input:* a. | ${}^{*}_{\mathrm{L}}\underline{\mu}_{\mathrm{H}}$ | τ ⊳ μ |
|---|---|---|
| a.  |  | **!* |
| b.  | *! |  |
| ☞ d.  |  | * |

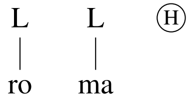
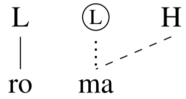
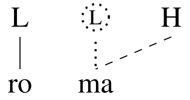
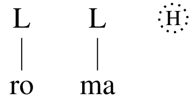
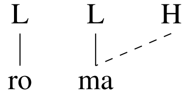
At the Word Level, undominated ALTERNATION blocks association of the floating H to the tautomorphemic word-final mora:

(27) *Kuria Word Level (Remote Future/μ3)*

| *Input:* a. | ALT | *$\textcircled{\tau}$…] | ${}^{*}_{\mathrm{L}}\underline{\mu}_{\mathrm{H}}$ | ${}^{*}_{\mathrm{L}}\underline{\mu}_{\mathrm{L}}$ | τ ⊳ μ | MAX τ |
|---|---|---|---|---|---|---|
| ☞ a.  |  |  |  |  | * |  |
| b.  | *! |  |  | * |  |  |
| c.  | *! | * |  |  |  |  |
| d.  |  |  |  |  | * | *! |

In contrast at the Phrase Level, ALT is ranked below τ ⊳ μ, and the H finally associates:

(28) *Phrase Level (Remote Future/μ3)*

| *Input:* a. | $*\textcircled{\tau}\ldots]$ | $\overset{*}{{}_{\mathrm{L}}\underline{\mu}_{\mathrm{L}}}$ | MAX τ | $\tau \rhd \mu$ | $\overset{*}{{}_{\mathrm{L}}\underline{\mu}_{\mathrm{H}}}$ |
|---|---|---|---|---|---|
| a.  | | | | *! | |
| b.  | *! | | | | |
| c.  | | | *! | | |
| d.  | | | *! | * | |
| ☞ e.  | | | | | * |

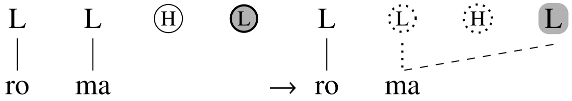
Thus again ALTERNATION must be ranked crucially differently at the Word and Phrase Level.

That the blocking of H-tone association at the Word Level is due to ALT and not to general faithfulness constraints protecting already associated tones is evident from the fact that affixal Word Level tones in fact associate to bases with already associated tones and overwrite them. Thus the L-tone suffix expressing negation in the Negative Remote Future replaces the associated H-tone on the last mora in (29):

(29) *Kuria (Negative) Remote Future (µ3) (Mwita 2008:198)*

| | *Remote Future* | *Negative Remote Future$_1$* | |
|---|---|---|---|
| | ‘we will ... ’ | ‘they will not ... then’ | |
| a. 1µ-*Stems*: | n-to-re-[rj-a] | βa-ta-re-[rj-à] | ‘eat’ |
| b. 2µ-*Stems*: | n-to-re-[rom-ǎ̲] | βa-ta-re-[rom-à] | ‘bite’ |
| c. 3µ-*Stems*: | n-to-re-[tɛrɛk-á̲] | βa-ta-re-[tɛrɛk-à] | ‘brew’ |
| | FOC-1PL-TAM-[√-FV] | 3PL-NEG-TAM-[√-FV] | |

This is illustrated for (29b) in (30):

(30) 

Similarly, in the Inceptive and Immediate Past, Word Level H tone prefixes overwrite associated L stem tones (see Trommer 2023 for details).

## 4. Other Constraints on Color as SCC-Triggers

### 4.1. Tone Spreading and Dissimilation in Bari

Bari has two interacting processes, H-tone spreading, where a word-final High tone spreads to a following word-initial L-toned syllable and replaces its tone (31), and H-dissimilation where a word-initial H is replaced by L after a word-final High (32):

(31) *H-Spreading (Yokwe, 1986:208)*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. ríp | ‘sawed’ | + | dù.pà | ‘cradle’ | → | ríp dú.pà |
| b. nín | ‘twisted’ | + | gwàkà | ‘forked stick’ | → | nín gwá.kà |

(32) *H-Dissimilation (Yokwe, 1986:207)*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. dók | ‘fetched’ | + | kó.pò | ‘cup’ | → | dók kò.pò |
| b. gwó | ‘kicked’ | + | gú.rè | ‘dove’ | → | gwó gù.rè |

Derivationally speaking, H-deletion may feed H-Spreading, as shown by the examples in (33):

(33) *H-Dissimilation feeds H-Spreading (Yokwe, 1986:206)*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. dép | ‘held’ | + | ké.ré | ‘gourd’ | → | dép ké.rè |
| b. kúr | ‘dug’ | + | kí.dí | ‘well’ | → | kúr kí.dì |

Thus (33a) can be understood as first applying dissimilation (/dép ké.ré/ → dép kè.rè) and subsequently spreading (dép kè.rè → [dép ké.rè]). However with slightly different inputs H-Dissimilation applies without following H-Spreading:

(34) *H-Dissimilation counterfeeds H-Spreading . . .*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. dók | ‘fetched’ | + | kó.pò | ‘cup’ | → | dók kò.pò |
| b. gwó | ‘kicked’ | + | gú.rè | ‘dove’ | → | gwó gù.rè |

The crucial difference between (34) and (33) seems to be that in the cases in (34), H-spreading would restore the input pronunciation (/dók kó.pò/ → [dók kò.pò] → *[dók kó.pò]).

We are now in a position to turn to the relevance of Bari tone for SCC effects. Crucially, H-spreading does not happen word-internally:

(35) *No word-internal H-spreading* (Yokwe, 1986:126+129+130)

| | | | |
|---|---|---|---|
| a. dúlúr | ‘castor oil plant’ | d. bírìsì | ‘mat’ |
| b. bángìʔ | ‘marijuana’ | e. básàlà | ‘onion’ |
| c. wúrì | ‘cork’ | f. ng’únùmì | ‘whiskers’ |

Similarly, spread H-tone does not trigger word-internal dissimilation:

(36) *H-Spreading counterfeeds H-Dissimilation*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. tór | ‘tied’ | + | bòn.gó | ‘dress’ | → | tór bón.gó |
| b. pák | ‘scared’ | + | dì.rán | ‘birds’ | → | pák dí.rán |

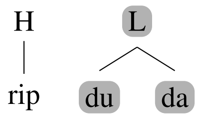
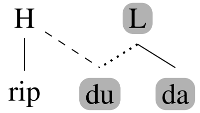
While we could attribute, the blocking of word-internal H-spreading to ALTERNATION, this wouldn’t account for the blocking of dissimilation. Fortunately, there seems to be a single unifying principle behind both data. Specifically I propose that Bari tone shows the effect of the constraint in (37a):

(37)

| | |
|---|---|
| a. $[\sigma_H$ | Assign $*$ to every initial syllable in a prosodic word<br>which is not the right edge<br>of a morphological or phonetic H-tone span |
| b. OCP | Assign $*$ to every pair of adjacent H-tones in I |
| c. $*\textcircled{\tau}$ | Assign $*$ to every floating epenthetic tone |
| d. Faith \| | Assign $*$ to every epenthetic or deleted association line |

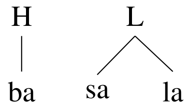
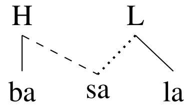
$[\sigma_H$ triggers spreading across a word boundary:

(38) *H-Spreading*

| **Input:** = a. | OCP | $[\sigma_H$ | $*\textcircled{\tau}$ | Faith \| |
|---|---|---|---|---|
| a.  | | *! | | |
| ☞ b.  | | | | ** |

Word-internally H-spreading of initial Hs is blocked simply because $[\sigma_H$ is already satisfied, and additional spreading would involve unnecessary violations of Faith |:

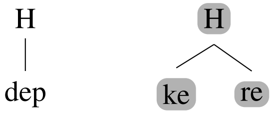
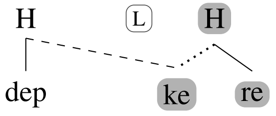
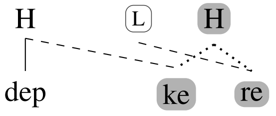
(39) *Blocking of word-internal H-Spreading*

| **Input:** = a. | OCP | $[\sigma_H$ | $*\textcircled{\tau}$ | Faith \| |
|---|---|---|---|---|
| ☞ a.  | | | | |
| b.  | | | | *!* |

The tableau in (40) illustrates dissimilation in conjunction with concomitant

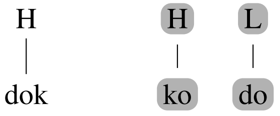
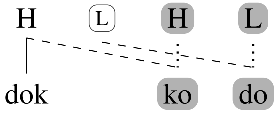
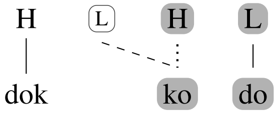
H-spreading. A L-tone is inserted to satisfy the OCP. Since it cannot remain floating by virtue of *$\textcircled{\tau}$, the final syllable becomes L:

(40) *H-Spreading + L-Epenthesis*

| **Input:** = a. | OCP | [$\sigma_H$ | *$\textcircled{\tau}$ | Faith \| |
|---|---|---|---|---|
| a.  | *! | *! | | |
| b.  | | | *! | ** |
| ☞ c.  | | | | **** |

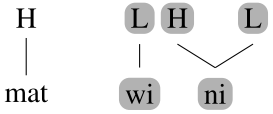
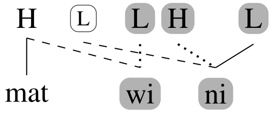
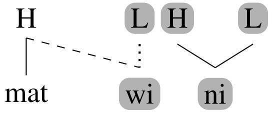
For H-initial Hs, spreading is again blocked because [$\sigma_H$ is already satisfied by the underlying H-tone:

(41) *Duke-of-York Blocking: L-Epenthesis without H-Spreading*

| **Input:** = a. | OCP | [$\sigma_H$ | *$\textcircled{\tau}$ | Faith \| |
|---|---|---|---|---|
| a.  | *! | | * | |
| b.  | | | | ***!* |
| ☞ c.  | | | | ** |

The crucial SCC effect is derived in (42). H spreads to the following word. That the OCP does not trigger dissimilation follows directly from Containment, as laid out in section 2. OCP (in contrast to its clone $\underline{\text{OCP}}$) is a generalized markedness constraint. Hence the unpronounced underlying initial L-tone of [winî] is visible to it, thus it is not violated by the sequence HLH, and deassociating the underlying H is simply blocked by FAITH |:

(42) *SCC for Dissimilation: Blocking of L-Epenthesis by Underlying L*

| **Input:** = a. | OCP | [σ$_H$ | *ⓣ | FAITH \| |
|---|---|---|---|---|
| a.  | | *! | | |
| b.  | | | | ***!* |
| ☞ c.  | | | | ** |

4.2. CRISP EDGE: Kashaya Laryngeal Dissimilation

Kashaya Laryngeal Dissimilation (Buckley 1994) is a SCC effect which obviously cannot be captured by ALTERNATION since it does not involve spreading. Crucially, word-initial aspirated stops lose their aspiration if they are followed by another aspirated stop or by [h]:

(43) *Kashaya Laryngeal Dissimilation (Buckley 1994:83)*

| | | | |
|---|---|---|---|
| a. /p$^h$i-$^h$mi-w/ | → | [pihmíw] | 'see in detail' |
| b. /p$^h$u-$^h$c$^h$a-w/ | → | [puhc$^h$áw] | 'blow over' |
| c. /p$^h$a-hol-ʔ/ | → | [pahól'] | 'look for an unseen object with end of stick' |

However, Laryngeal Dissimilation only applies across morpheme boundaries, i.e., it is not triggered by other root internal laryngeal features:

(44) *Kashaya: No root-internal dissimilation (Buckley 1994:85)*

| | | | |
|---|---|---|---|
| a. ṭ$^h$ahqa- | 'play' | c. ṭ$^h$eq$^h$á-le | 'elderberry' |
| b. q$^h$oh´j | 'eighty' | d. k$^h$om$^h$ca | 'eight' |

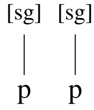
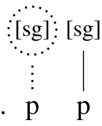
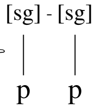
I assume that dissimilation involves the feature [spread glottis] ([sg]) characterizing aspirate stops and [h], and is triggered by OCP [sg]. Assuming that [sg] is the only feature on its autosegmental tier, this OCP constraint can only be repaired by deleting (making phonetically invisible) one of the involved features or by fusion, implemented here as orthogonal association. Following Trommer (2018) I assume that fused nodes count as single objects for markedness constraints, hence [sg] - - - [sg] avoids violating OCP [sg]. The SCC effect can now be captured by a constraint type which is in a sense the logical complement of ALTERNATION, CRISP EDGE (Ito and Mester 1999, Kaplan 2018)

(45) CRISPEDGE□ (CE□) Assign ∗ to every nodes of different color which are connected by an epenthetic association line

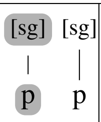
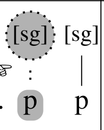
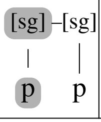
Now root-internally fusion of two adjacent [sg] features is preferred over deletion due to ranking MAX [sg] above DEP – (which penalizes insertion of a horizontal association line) as shown by (46i). However across a morpheme boundary fusion would violate CE-□, and deletion applies instead (46ii):

(46) *Kashaya Laryngeal Dissimilation*

(i) *Root-internal fusion*

| *Input:* a. | OCP [sg] | CE □ | MAX [sg] | DEP – |
|---|---|---|---|---|
| a.  | *! | | | |
| b.  | | | *! | |
| ☞ c.  | | | | * |

(ii) *Intermorphemic deletion*

| *Input:* a. | OCP [sg] | CE □ | MAX [sg] | DEP – |
|---|---|---|---|---|
| a.  | *! | | | |
| ☞ b.  | | | * | |
| c.  | | *! | | * |

### 4.3. Turkish Velar Deletion

In Turkish, specific (vowel-initial) suffixes such as the accusative and person number inflection trigger deletion of a final velar obstruent (47), whereas others such as the Future suffix -[eʤek] do not (48):

(47) *Turkish affixes triggering velar deletion (Orgun 1996:106)*

| | | | | | |
|---|---|---|---|---|---|
| a. | bɑdʒɑk | bɑdʒɑ-ɑ | bɑdʒɑ-ɯ | bɑdʒɑ-ɯm | bɑdʒɑ-ɯn |
| | ‘leg’ | ‘leg-DAT’ | ‘leg-ACC’ | ‘leg-1SG.POSS’ | ‘leg-2SG.POSS’ |
| b. | sɑlɑk | sɑlɑ-ɯm | sɑlɑ-ɯz | | |
| | ‘stupid’ | ‘stupid-1SG.SUBJ’ | ‘stupid-1PL.SUBJ’ | | |

(48) *Turkish affixes not triggering velar deletion (Orgun 1996:106)*

| | | | |
|---|---|---|---|
| gʲedʒikʲ | gʲedʒikʲ-edʒekʲ | gʲedʒikʲ-ebilʲ-ir | gʲedʒikʲ-indʒe |
| ‘be late’ | ‘be.late-FUT’ | ‘be.late-ABIL.IMPF’ | ‘be.late-ADV’ |

Strict cycle effects emerge with two kinds of data. First, velar deletion does not apply to intervocalic velars in underived roots:

(49) *No velar deletion in underived roots (Inkelas and Orgun 1995:768)*

| | | | |
|---|---|---|---|
| a. gaga | ‘beak’ | c. oku | ‘read’ |
| b. sigara | ‘cigarette’ | d. sokak | ‘street’ |

*Second,* if an affix triggering velar deletion embeds a non-trigger affix, only the outer affix itself causes deletion of the preceding velar, the non-trigger affix still fails to trigger the process.

(50) *Turkish deletion triggers outside of non-triggers (Orgun 1996:106)*

| | | | |
|---|---|---|---|
| a. | gʲedʒikʲ-edʒe-im | b. | birikʲ-edʒe-i |
| | ‘be.late-FUT-1SG.SUBJ’ | | ‘accumulate-FUT-3SG.SUBJ’ |
| c. | bɯrɑk-ɑdʒɑ-ɯn | d. | gʲerekʲ-edʒe-imiz |
| | ‘let.go-FUT-2SG.SUBJ’ | | ‘be.necessary-FUT-1PL.SUBJ’ |

I assume that velar deletion itself is a Phrase-Level process which deletes intervocalic velars (51a) except in case this is the onset of a syllable

(51) *Velar deletion (Phrase Level)*

| *Input:* a. | Max $_\sigma$[C | *VkV | Max C |
|---|---|---|---|
| a. [sɑ][lɑk]-[ɯm] | | *! | |
| ☞ b. [sɑ][lɑk]-[ɯm] | | | * |

| *Input:* a. | Max $_\sigma$[C | *VkV | Max C |
|---|---|---|---|
| ☞ a. [o][ku] | | * | |
| b. [o][ku] | *! | | * |

The crucial contribution of the lexical phonology is to ensure the word-internal syllabification of velars in a way such that this leads to the correct outputs in every case. Before we consider the apparent SCC effects, let us turn to an important purely phonological restriction on velar deletion which patterns nicely with the syllable-based approach pursued here. Velar deletion is also blocked after long vowels:

(52) *No velar deletion after long vowels (Inkelas 2009:394)*

| | *Nominative* | *Dative* | |
|---|---|---|---|
| a. /infilaːk/ | [in.fi.lak] | [in.fi.laː.ka] | ‘explosion’ |
| a. /meraːk/ | [me.rak] | [me.raː.ka] | ‘curiosity’ |

This can be captured by the constraint $*\sigma_{3\mu}$ which penalizes trimoraic syllables and is independently motivated by the fact that Turkish has regular vowel shortening in closed syllables (der Hulst and Weijer 1991):

(53) *Syllabification of velars after short and long vowels (Word Level)*

| *Input:* a. | PRS | $*\sigma_{3\mu}$ | $^{*}_{\sigma}$[k | ONS |
|---|---|---|---|---|
| a. sɑlɑk-ɯm | *! | | | |
| b. [sɑ][lɑ]-[kɯm] | | | *! | |
| ☞ c. [sɑ][lɑk]-[ɯm] | | | | * |

(→ Phrase Level [sɑ][lɑːk]-[ɯm])

| *Input:* a. | PRS | $*\sigma_{3\mu}$ | $^{*}_{\sigma}$[k | ONS |
|---|---|---|---|---|
| a. meraːk-a | *! | | | |
| ☞ b. [me][raː][k-a] | | | * | |
| c. [me][raːk]-[a] | | *! | | * |

(→ Phrase Level [me][raː][k-a])

The crucial constraint for the SCC-effect is □ONS:[3]

(54) □ONS: Assign $*$ to every consonant which is not in the onset of a tautomorphemic right-adjacent vowel

(55) *Syllabification of velars after short and long vowels (Word Level)*

| *Input:* a. | PRS | $*\sigma_{3\mu}$ | □ONS | $^{*}_{\sigma}$[k | ONS |
|---|---|---|---|---|---|
| a. oku | *! | | | | |
| b. [ok][u] | | | *! | | * |
| ☞ c. [o][ku] | | | | * | |

(→ Phrase Level [o][ku])

Let us finally address the fact that velar deletion is blocked for velars which are the final consonants of verb roots (der Hulst and Weijer 1991:35). I assume that this is the reflex of a categorial verbal theme marker consisting of a

---

[3] For reasons of space, I do not analyze the fact here that velar deletion is also blocked inside morphemes in roots which are broken up by an epenthetic vowel due to phonotactic reasons (e.g., /aks/ → [akis] 'reflection', Inkelas and Orgun 1995:776) . This could be captured either by an additional constraint requiring that colorless (epenthetic) vowels have onsets, or by generalizing □ONS such that it requires syllabification of a consonant with an (underlyingly) following tautomorphemic segment.

floating mora which attaches to the final syllable of the word.[4] This effectively forces [k] to form an onset, despite lower-ranked $^{*}_{\sigma}$[k:

(56) *Syllabification of velars after verb root (Word Level)*

| *Input:* a. | Prs | $*\sigma_{3\mu}$ | ☐Ons | $^{*}_{\sigma}$[k | Ons |
|---|---|---|---|---|---|
| a. gedʒik$^{\mu}$-en | *! | | | | |
| b. [ge][dʒik$^{\mu}$][-en] | | *! | | | * |
| ☞ c. [ge][dʒi$^{\mu}$][k-en] | | | | * | |

(→ Phrase Level [ge][dʒi$^{\mu}$][k-en])

### 4.4. Turkish Bisyllabicity

A further apparent SCC effect for some speakers of Turkish cited by Inkelas and Orgun (1995, 1998) is a disyllabic minimality requirement restricted to derived forms. Thus the name of the musical note C [doː] is grammatical in isolation, but not in combination with a non-syllabic possessive affix (*[doː-m] 'my C', cf. [sol$^{j}$] 'musical note G' and [sol$^{j}$-ym] 'my G', which are both grammatical, Orgun 1996:19). Inkelas and Orgun (1995) further argue that this minimality requirement is tied to specific phonological strata since it can be repaired by 'inner' affixes such as possessive suffixes, but not by outer affixes such as case suffixes (compare, e.g., [sol$^{j}$-ym] 'my G' and *[doː-m-u] 'my C (Acc.)', Orgun 1996:19).

(57) *Strata in Turkish according to Inkelas and Orgun (1998:368)*

| Level: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Morphology | root | passive<br>aspect<br>relative<br>negative | plural<br>possessive | case | tense<br>agreement<br>interrogative |
| Phonology | [μμ] | | | | |
| | | [σσ] | [σσ] | | |
| | | | velar deletion | velar deletion | velar deletion |

I will not attempt a full analysis of this pattern here since it hinges on the

---

[4] This syllable is catalectic in the sense of Kiparsky (1991) See Zimmermann and Trommer (2014) for extensive independent evidence for catalectic moras in morphophonology.

orthogonal controversial question of how to implement absolute ungrammaticality in OT. However in principle, the disyllabic minimality requirement could be simply in Colored Containment SOT by a version of Downing's Morpheme-Syllable Correlation (Downing 2006:ch.3)

(58) MORPHEME-SYLLABLE CORRELATION (MSC): A form containing at least two colors should also contain at least two syllables

Consider finally how this could be embedded into a model which allows only for two word-internal strata, as assumed here, in contrast to the 5 strata posited by Inkelas and Orgun (1998) shown in the diagram in (57). I assume that Turkish has a Stem Level comprising levels 1+2+3 of Inkelas & Orgun, and a Word Level comprising complete word forms. The Stem Level under this definition is not only the domain for the bisyllabicity constraint under discussion, but also corresponds also to the site which can undergo suspended affixation (Inkelas and Orgun 1995:765). As already indicated by the diagram in (57), Inkelas and Orgun (1995, 1998) do not provide any specific evidence for distinguishing stratum 4 and stratum 5, and their only reason for distinguishing stratum 2 and 3 is velar velar deletion, which as already shown in section 4.3 can be captured purely representationally (the fact that level-2 affixes in Turkish don't allow suspended affixation is also an argument against positing a separate stratum 2). The only evidence Inkelas and Orgun (1995) provide for stratum 1, is a process where certain speakers the long vowel of monosyllabic roots such as [doː] 'C' is shortened in a construction Inkelas and Orgun call 'root compounding' (e.g. [do-dijez] 'C-sharp', Inkelas and Orgun 1995 :773). Inkelas and Orgun attribute this to a bimoraic minimality requirement at stratum 1 (comprising root compounding) such that monosyllabic roots lengthen in all contexts apart from root compounds. but it could also be captured by assuming that the underlying linking element of these compounds triggers shortening of preceding roots. Again this could be implemented by positing a floating mora as linker which leads to shortening in a similar way as described for shortening morphologies in Zimmermann and Trommer (2014). Independent evidence for the assumption that this is morphological shortening comes from the fact that Turkish has roots which are not subject to the alleged bimoraicity constraint for all speakers (e.g., [je] 'eat' and [su] 'water' Inkelas and Orgun 1995:786).

## 5. Beyond Colors: Representational Sources of SCC-Effects

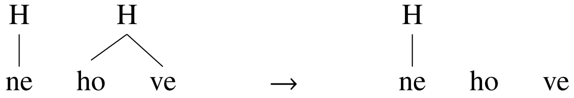
### 5.1. Fusion and the OCP across Strata

in Shona High tone (= H-tone) dissimilation ('Meeussen's Rule'), the second of two adjacent H-tones is deleted (e.g. [né] + [mbwá] → [né mbwa] 'with a dog', inside a phonological word at the Phrase Level , but crucially this effect is restricted to appear across word boundaries. Thus a H-toned verb such as /hóvé/ 'fish' is not changed to [hóve] at the Phrase Level. The standard account for this and similar facts in Autosegmental Phonology is based on the assumption that the two cases differ representationally because the OCP has applied word-internally as a morpheme structure constraint (Hyman 2014) or as an OT-constraint at the Word Level in a Stratal OT approach. Thus [hóve] does not have two H-tones, but just one which is doubly linked escaping the applicability of the OCP:

(59) a.

Independent evidence for this representation comes from the fact that in words such as [hóve] all adjacent H-tone syllables become L if preceded by a H-toned word (e.g. [né] + [hóvé] → [né hove] 'with a fish'.[5]

(60) a.



### 5.2. Syllabification as an SCC-Trigger: Emai

If generalized to syllable structure, the representational-stratal account of SCC effects in Shona tone in section 5.1 can also account for the Emai data discussed at the very beginning of this paper. Assume that the relevant markedness constraint *Hiatus is only violated by adjacent vowels in different syllables. At the Word Level, *Hiatus violations of two consecutive vowels are repaired phonetically vacuously by grouping them into the same syllable (oa → [oa]$_{\sigma}$).

---

[5] Note that the Shona case is different from laryngeal dissimilation in Kashaya, where dissimilation/deletion and fusion are possible inside the same stratum.

At the Phrase Level, the generalized markedness constraint *[VV]$_\sigma$ blocks creation of new diphthongal syllables ([kɔ]$_\sigma$ [e]$_\sigma$[ma]$_\sigma$ *→ [kɔe]$_\sigma$[ma], but diphthongs inherited from the Word Level are maintained (since *[VV]$_\sigma$ is a generalized markedness constraints modifying it would not avoid the constraint violation triggered by such a syllable since it would still be visible to the constraint due to Containment). At the same time, Phrase Level *HIATUS violations are repaired by deletion ([kɔ]$_\sigma$ [e]$_\sigma$[ma]$_\sigma$ → [ke]$_\sigma$[ma]$_\sigma$), but isolated [oa]$_\sigma$ is maintained because by assumption it does not violate *HIATUS in the first place.

### 5.3. Klamath

As Catalan (see section 3.1), vowel reduction and deletion in Klamath has been used as one of the major arguments for the SCC based on the classical paper by Kean (1974). Here I will show that the relevant effects can also be captured making use of the basic representational mechanism of Containment Theory. See Gleim (2023) on a critical evaluation of Kean's original analysis.

The phonological process involved is schwa epenthesis in consonant clusters containing sonorants (e.g., /tgalm/ → [tgaləm] 'west', Kean 1974:188). This interacts with another general process in Klamath the initial vowel of vowel-initial stems and formatives is obligatorily deleted (e.g. /lolal-op-ga/ → [lolal-p-ga], 'are lying down (pl.)', p.184).The classification of these elements as vowel-initial is somewhat abstract since the putative vowels are virtually always deleted. The major evidence for their presence is the behavior of vowel-initial stems in combination with reduplicative prefixes such as [pV]- 'pull' (e.g., toːka 'hair falls out' ~ [po-toːka] 'pull hair out', p. 180; [katsga] 'tooth falls out' ~ [pa-kətsga] 'pull someone's tooth out', p.181). Before a vowel-initial root these prefixes copy the initial root vowel (which is simultaneously deleted,, as in all other contexts): /-aciːk-/ 'wrings out' → pa-ciːka 'pulls and twists'; /-odg-/ 'out of container' → [po-tga]'pulls out of container' (p.182). (61) shows one of the three examples for the alleged SCC effects cited by Keane (the other two examples are almost completely parallel wrt the underlying phonological features and alternations involved). Phonological rules (cycles) apply after the addition of every new formative. No epenthesis happens in cycle 1 since ə-insertion is restricted to consonant clusters with three consonants (or two-consonant clusters at the end of a word). The vowel of [elg] is not deleted since the deletion rule in Kean's analysis only applies after a morpheme boundary (since the relevant vowel-initial formatives never surface without a further prefixal element, this still means that their initial vowel is deleted across the board). In the second cycle a prefix is added and the root-initial [e] is deleted. This in turn results in a 4-consonant cluster involving sonorants which triggers ə-insertion. In cycle 3 the initial vowel of [otən] is deleted and its final [n] assimilates to the following [l]. The SCC-effect diagnosed by Keane now consists in the fact that the 3-consonant cluster [n-l-g] in the input of Cycle 3 doesn't trigger an additional instance of ə-epenthesis:

(61) *[ntiwtəllga] 'falls against something' (Kean 1974:188)

| | | | | |
|---|---|---|---|---|
| | | [elg-a] | Cycle 1 | |
| | | [elg-a] | | |
| | [otn- | [elg-a]] | Cycle 2 | (ə-epenthesis + V-deletion) |
| | [otən- | [lg-a]] | | |
| [ntiw- | [otən- | [lg-a]] | Cycle 3 | (V-deletion + n/l-assimilation) |
| [ntiw- | [təl- | [lg-a]] | | |
| *[ntiw- | [tən- | [əlg-a]] | | (additional ə-epenthesis) |

My reanalysis of this datum is based on the observation that additional insertion of a vowel here would essentially consist in a Duke-of-York derivation (similarly to the Bari case discussed in section 4.1). Insertion of a schwa in this position would add a vowel exactly at the point where root-initial [e] was deleted in the first cycle. In Containment, this can be captured by the constraint *VV (Assign ∗ to every pair of adjacent vowels). Since this is a generalized markedness constraint sensitive to input and output structure alike, it crucially excludes the Duke-of-York/SCC candidate (62b). (in contrast, *[V and *Cluster are phonetic markedness constraints, hence they can in principle be satisfied by vowel deletion). I refrain from formally defining the constraint *Cluster here since the conditions triggering vowel epenthesis in clusters in Klamath are complex. For the sake of the tableau in (62), I assume that it is violated by every distinct phonetic three-consonant sequence containing at least one sonorant. Thus there is one violation in the winning candidate (62d) (incurred by [l-l-g]) and two in (62c) (incurred by w-t-n and t-n-l).[6]

---

[6] With Zoll (2002), I assume that suffixes, roots (and prefixal intensive reduplication) form the

(62) *Blocking of Duke of York epenthesis in Klamath*

| *Input:* a. | *[V | *VV | *Cluster | Max V | Dep V |
|---|---|---|---|---|---|
| a. [ntiw- [otn- [elga]] | *!* | | | | |
| b. [ntiw- [otən-ə[elga]] | | *! | | ** | ** |
| c. [ntiw- [otn- [elga]] | | | **! | ** | * |
| ☞ d. [ntiw- [otəl- [elga]] | | | * | ** | * |

## References

Bermúdez-Otero, Ricardo (2018): Stratal phonology. *In:* S. Hannahs and A. R. K. Bosch, eds, *The Routledge Handbook of Phonological Theory*. Routledge, Abingdon, pp. 100–134.

Buckley, Eugene (1994): *Theoretical aspects of Kashaya phonology and morphology*. CSLI Publications, Stanford, California.

Casali, Roderic F. (1997): ‘Vowel Elision in Hiatus Contexts: Which Vowel Goes?’, *Language* **73**(3), 493–533.

der Hulst, Harry Van and Jeroen Van De Weijer (1991): Topics in Turkish phonology.. *In:* H. E. Boeschoten and L. T. Verhoeven, eds, *Turkish linguistics today*. Brill, Leiden, pp. 11–59.

Downing, Laura (2014): ‘Melodic verb tone patterns in Jita’, *Africana Linguistica* **20**(1), 101–119.

Downing, Laura J. (2006): *Canonical Forms in Prosodic Morphology*. Oxford University Press, Oxford.

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.

Hyman, Larry M. (2014): ‘How autosegmental is phonology?’, *The Linguistic Review* **31**(2), 363 –400.

Inkelas, Sharon (2009): ‘Another Look at Velar Deletion in Turkish, with Special Attention to the Derived Environment Condition’, *UC Berkeley PhonLab Annual Report,* **5**(5), 387–403. Permalink https://escholarship.org/uc/item/32s446d4.

Inkelas, Sharon and Cemil Orhan Orgun (1995): ‘Level ordering and economy in the lexical phonology of Turkish’, *Language* **71**, 763–793.

Inkelas, Sharon and Cemil Orhan Orgun (1998): Level (non)ordering in recursive morphology: evidence from Turkish. *In:* S. G. Lapointe, D. Brentari and P. Farrell,

---

Stem Level stratum in Klamath, whereas all other prefixes are added at the Word Level. Since the derivation in (i) happens at the Word Level, the suffix has acquired the same morphological color as the root, and is hence not incurring a violation of *[V.

eds, *Morphology and its Relation to Phonology and Syntax*. CSLI Publications, Stanford, CA, pp. 360–392.

Ito, Junko Itô and Armin Mester (1999): Realignment. *In:* R. K. Kager, H. van der Hulst and W. Zonneveld, eds, *The Prosody-Morphology Interface*. Cambridge University Press, Cambridge, pp. 188–217.

Kaplan, Aaron (2018): Asymmetric Crisp Edge. *In:* R. Bennett, A. Angeles, B. Adrian, D. Buckley, N. Kalivoda, S. Kawahara, G. McGuire and J. Padgett, eds, *Hana-bana: A Festschrift for Junko Ito and Armin Mester*. .

Kean, Mary-Louise (1974): 'The Strict Cycle in Phonology', *Linguistic Inquiry* **5**(2179-203).

Kenstowicz, Michael (1994): *Phonology in generative grammar*. Blackwell, Cambridge MA.

Kiparsky, Paul (1982): Lexical phonology and morphology. In: *Linguistic Society of Korea: Linguistics in the morning calm*. Hanshin Publishing Company, Seoul, pp. 3–92.

Kiparsky, Paul (1991): Catalexis. Unpublished manuscript Stanford University and Wissenschaftskolleg zu Berlin.

Kiparsky, Paul (1993): Blocking in nonderived environments. In: *Studies in Lexical Phonology*. Vol. 4 of *Phonetics and Phonology*, Academic Press, San Diego, pp. 277–313.

Kiparsky, Paul (2000): 'Opacity and cyclicity', *The Linguistic Review* **17**, 351–67.

Leben, William (1973): Suprasegmental Phonology. PhD thesis, MIT.

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

Mwita, Leonard Chacha (2008): Verbal tone in Kuria. PhD thesis, UCLA.

Myers, Scott (1997): 'OCP effects in Optimality Theory', *NLLT* **15**(4), 847–892.

Orgun, Cemil Orhan (1996): Sign-based Morphology and Phonology with Special Attention to Optimality Theory. PhD thesis, UC Berkeley.

Paschen, Ludger (2018): The Interaction of Reduplication and Segmental Mutation: A Phonological Account. PhD thesis, Leipzig University.

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

Prince, Alan and Paul Smolensky (1993): 'Optimality Theory: Constraint Interaction in Generative Grammars'. Technical reports of the Rutgers University Center of Cognitive Science.

Revithiadou, Anthi (2007): Colored Turbid Accents and Containment: A Case Study from Lexical Stress. *In:* S. Blaho, P. Bye and M. Krämer, eds, *Freedom of Analysis?*. Mouton De Gruyter, Berlin & New York, pp. 149–174.

Rubach, Jerzy (2011): 'Syllabic repairs in Macedonian', *Lingua* **121**(2), 237–268.

Trommer, Jochen (2011): Phonological Aspects of Western Nilotic Mutation Morphology. Habilitation Thesis, University of Leipzig.

Trommer, Jochen (2018): An Infection Analysis of Across-the-Board Lowering in Jumjum. Talk at OCP 15. London, January 14 2018.

Trommer, Jochen (2023): 'The Stratal Structure of Kuria Morphological Tone', *Phonology (to appear)* .

van Oostendorp, Marc (2005): The Theory of Faithfulness. Ms., Meertens Institute, Amsterdam.

van Oostendorp, Marc (2007): Derived Environment Effects and Consistency of Exponence. *In:* S. Blaho, P. Bye and M. Krämer, eds, *Freedom of Analysis?*. Berlin: Mouton de Gruyter, pp. 123–148.

van Oostendorp, Marc (2008): 'Incomplete Devoicing in Formal Phonology', *Lingua* **118**(9), 1362–1374.

Wolf, Matthew (2008): Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. PhD thesis, University of Massachusetts Amherst.

Yokwe, Eluzai M. (1986): The tonal grammar of Bari. PhD thesis, University of Illinois at Urbana-Champaign.

Zimmermann, Eva and Jochen Trommer (2014): 'Generalized Mora Affixation and Quantity-Manipulating Morphology', *Phonology* **31**, 463–510.

Zoll, Cheryl (2002): 'Vowel Reduction and Reduplication in Klamath', *Linguistic Inquiry* **33**(3), 520–527.

# Part III
# Removing Obstacles

# Anti-local Agree and Cyclicity

Rosa Fritzsche[*]

## Abstract

This paper argues that the difference of potential agreement controllers in Lak (Nakh-Daghestanian, Russia) Biabsolutive Constructions (lexical verb agreement controlled by the internal argument and auxiliary agreement by the external argument) and Ergative Constructions (agreement controlled only by the internal argument) stems from *Generalized Anti-Locality* outlawing strictly local Agree relations. I propose that in Ergative Constructions, the external argument is too close to the probe, while in Biabsolutive Constructions it moves sufficiently far away from the probe to control (delayed) upward agreement. However, Generalized Anti-Locality restricting the search space of upward Agree leads to problems with cyclicity. I argue that these can be dealt with by Reciprocal Subcategorization applying before φ-Agree. Moreover, Strict Cyclicity will prevent downward auxiliary agreement in biabsolutive constructions, obviating the need for a language-specific directional bias of Agree.

## 1. Introduction

Standardly, Agree (Chomsky 2000, 2001) is assumed to be a strictly local operation. A widely discussed challenge for this assumption comes from apparent cases of long-distance agreement suggesting, at first glance, that Agree *does not have to be* strictly local (Polinsky and Potsdam 2001, Bhatt 2005, Bobaljik and Wurmbrand 2005). Anti-local agreement phenomena challenge this standard approach to Agree in suggesting that Agree *must not be* strictly local. If on the right track, this line of reasoning suggests that anti-locality (Grohmann 2003, Abels 2003) emerges as a general constraint on syntax.

---

[*] I would like to thank the audience of the Workshop on (Strict) Cyclicity at Universität Leipzig on June 14, 2022 and the participants of the Leipzig Syntax/Morphology Colloquium (where I presented related work on several occiasions) for valuable feedback, questions and discussion. Special thanks are due to Fabian Heck, Gereon Müller, Mariia Privizentseva, Nina Radkevich and Sören E. Tebay.

*Cyclicity*, 273–296
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

One such anti-local agreement phenomenon is the Biabsolutive construction in Lak (Nakh-Daghestanian, Russia) in (1b), where the absolutive external argument controls agreement on the auxiliary (Kazenin 1998, Forker 2012, Gagliardi et al. 2014). This deviates from the baseline Ergative construction in (1a), where the absolutive internal argument controls agreement on the auxiliary, as well as on the lexical verb.

(1) *Ergative and biabsolutive constructions* (Gagliardi et al. 2014: 144):

a. Aˤli-l q:ata b-ullaj b-ur.
Ali.I-ERG house.III.ABS III-do.PROG III-AUX
'Ali is building a house.'

b. Aˤli q:ata b-ullaj ∅-ur.
Ali.I.ABS house.III.ABS III-do.PROG I-AUX
'Ali is in the state of building a house.'

I propose that in Lak biabsolutives the external argument moves (motivated by agent topicalization) to a position from which it can control upward agreement on the auxiliary without violating anti-locality.

However, as anti-locality restricts the search space of a probe, Agree either has to reach deep into already-built structure (in the case of downward Agree) or it has to be delayed until sufficiently enough structure is built-up (in the case of upward Agree). This can lead to problems with cyclicity (e.g. PIC and Earliness), resulting in the possibility of no licit agreement step. As we will see, anti-locality and the PIC have the capacity to make the searchable space too narrow for Agree, while the Earliness Principle can outlaw anti-local upward agreement.

The paper is structured as follows: Section 2 provides the background to an anti-locality constraint on Agree. Section 3 investigates the syntax of Lak biabsolutives and contains the proposal, as well as analysis. Finally, Section 4 deals with the problems for cyclicity arising as a consequence of the analysis. The paper concludes in Section 5.

## 2. Background: Anti-locality and Agree

Anti-locality (Grohmann 2003, Abels 2003) deals with too close relations in syntax in that syntactic dependencies must span a certain distance in order to be grammatical.[1] Anti-locality is well-established for movement-related phenomena such as *that*-trace effect and *tough*-constructions (Brillman and Hisch 2016); agent-focus in Kaqchikel (Erlewine 2016); subject and adjunct condition (Bošković 2016); raising-to-ergative in Nez Perce (Deal 2019); non-iterable symmetry in A-movement (Branan 2022); and extractions from subjects (Zyman 2021).

On the other hand, the standard assumption that Agree (Chomsky 2000, 2001) is strictly local is challenged by apparent cases of long-distance agreement (cf. Polinsky and Potsdam 2001, Bhatt 2005, Bobaljik and Wurmbrand 2005). Now, given that movement is constrained by both locality (i.e., minimality and cyclicity/PIC) as well as anti-locality, and agreement is also subject to the same notions of locality, a logical assumption would be that anti-locality also holds for agreement.

Parallel to what has been suggested for movement by e.g. Erlewine (2016), Bošković (2016), Deal (2019), Branan (2022), I propose that agreement dependencies that do not cross a full projection are outlawed by anti-locality. Thus, Generalized Anti-Locality (2) prohibits agreement between a head and its specifier or the specifier of its complement.[2]

(2) *Generalized Anti-Locality*:
*[ … $\alpha$ … $\beta$ … ] (where $\alpha$ and $\beta$ are participants in an Agree relation) unless there is a $\Gamma$ such that

a. $\Gamma$ is in the non-edge domain of a phrase XP.
b. $\alpha$ c-commands XP.
c. $\beta$ is reflexively included in $\Gamma$.

I take that the edge of a phrase XP consists of all specifiers of X and adjuncts to XP (Chomsky 2000, 2001). The non-edge domain of a phrase XP is everything excluding the edge (i.e., X and its complement).[3]

For now, I will not commit to the direction of Agree (3), but assume that

---

[1] The term for the phenomena at hand goes back to Grohmann (2003) with theoretical predecessors in Bošković (1994) and Saito and Murasugi (1999); see Grohmann (2011) for an overview.

[2] The version of anti-locality in (2) is adapted from similar formulations in Müller (2020) and Lee (2020).

[3] Identifying $\Gamma$ to be either the head or the complement of the c-commanded phrase would allow for head movement while still ruling out spec-to-spec movement as in e.g. Erlewine (2016).

Agree (Chomsky 2000, 2001) is bidirectional – it can be downward or upward (cf. Baker 2008, Georgi 2014, Carstens 2016, Himmelreich 2016; see also Preminger and Polinsky 2015, Bjorkman and Zeijlstra 2019 for a debate on the direction of Agree) and that specific languages have a directional bias.

(3) *Agree*:
A probe $\alpha_{[*F*]}$ triggers Agree and copies the feature-values of a goal $\beta_{[F]}$ iff (a) to (c) hold. Call $\alpha$, $\beta$ participants of the Agree relation.
a. $\alpha$ is unvalued and seeks the value of (a feature on) $\beta$
b. $\alpha$ and $\beta$ are in a c-command relation
c. $\beta$ is the closest available goal to $\alpha$

Generalized Anti-Locality, thus, restricts the search space of a probe and expects a certain distance between the probe and a goal in order for a goal to be available. In the next section, I investigate one case of anti-local agreement in Lak, where I argue that auxiliary agreement can be controlled by external arguments only if they are sufficiently far away from the φ-probe in T.

## 3. Lak Biabsolutives

Lak (Nakh-Daghestanian, Russia) transitive Ergative constructions containing an ergative external argument and an absolutive internal argument can alternate with Biabsolutive constructions containing two absolutive arguments (Kazenin 1998, Forker 2012, Gagliardi et al. 2014, Ganenkov 2016, Radkevich 2017; see also Chumakina and Bond 2016, Polinsky 2016, Ganenkov 2019 for Biabsolutive constructions in other Nakh-Daghestanian languages).

In Ergative constructions, both the lexical verb (*bullaj*) and the auxiliary (*bur*) agree with the internal argument (*q:ata*) in gender/class (4a). In Biabsolutive constructions, on the other hand, agreement on the auxiliary (*ur*) is controlled by the external argument in the absolutive (*A*ˤ*li*), while the internal argument still controls agreement on the lexical verb (4b).

(4) *Ergative and biabsolutive constructions* (Gagliardi et al. 2014: 144):

| | | | |
|---|---|---|---|
| a. Aˤli-l | q:ata | b-ullaj | b-ur. |
| Ali.I-ERG | house.III.ABS | III-do.PROG | III-AUX |

‘Ali is building a house.’

b. Aˤli q:ata b-ullaj ∅-ur.
Ali.I.ABS house.III.ABS III-do.PROG I-AUX
‘Ali is in the state of building a house.’

I argue that the difference of potential agreement controllers in Lak biabsolutives compared to ergative constructions stems from the anti-locality constraint on Agree: Arguments that are too close to the probe are not available as goals for Agree. In the ergative construction, the in-situ external argument is too local to the probe on T to control agreement on the auxiliary. In the biabsolutive construction, however, the external argument moves to a position from which it can control upward agreement on the auxiliary without violating anti-locality. This movement step is motivated by agent topicalization.

### 3.1. The Syntax of Biabsolutives

I assume a structure of (periphrastic) biabsolutives as in (5), where auxiliaries are located in T and not in Aux (contra Gagliardi et al. 2014, Ganenkov 2016). T bears a φ-probe [$*\varphi*$] that agrees with the external argument in Biabsolutive constructions. TP is dominated by a topic phrase (TopP), which hosts the absolutive external argument in its specifier (after movement).

(5) $[_{\text{TopP}}\ \text{DP}_{\text{Ext}}\ [\text{Top}]\ [_{\text{TP}}\ [\text{T}_{*\varphi*}]\ [_{\text{vP}}\ t_{DP}\ [\text{v}_{*\varphi*}]\ [_{\text{VP}}\ \text{V}\ \text{DP}_{\text{Int}}]]]]$

Based on observations concerning deverbal nouns (so-called *masdars*), there is reason to assume biabsolutives always involve a layer higher than vP. In Lak, masdars are formed using either suffix *-awu* (6) or *-šiwu* (7). According to Gagliardi et al. (2014) and Radkevich (2017) *-awu*-masdars are vP nominalisations and express only Aktionsart, while *-šiwu*-masdars involve a TP layer as they also express tense, aspect and mood. Crucially for the argument, biabsolutive-based *-awu*-masdars (vP) are ungrammatical (6b): the external argument cannot show up with absolutive case marking. On the other hand, *-šiwu*-masdars can be formed on the basis of biabsolutives (7b) with the external argument in absolutive controlling agreement on the auxiliary.

(6) *-awu masdars* (Radkevich 2017):

a. Aˤli-l q:ata b-ullal-awu
Ali.I.SG-ERG house.IIISG.ABS III.SG-do.PROG-MSDR
‘Ali’s building of the house.’

b. *Aˤli q:ata b-ullal-awu
Ali.I.SG.ABS house.IIISG.ABS III.SG-do.PROG-MSDR
'Ali's building of the house.'

(7) *-šiwu masdars* (Gagliardi et al. 2014: 155):
a. ?Aˤli-l q:ata b-ullaj b-aq:a-šiwu
Ali-ERG house.III.ABS III-do.PROG III-AUX.NEG-MSDR
b. Aˤli q:ata b-ullaj 0-aq:a-šiwu
Ali.ABS house.III.ABS III-do.PROG I-AUX.NEG-MSDR
'Ali's not building of the house.'

Data from person agreement and biabsolutives with a synthetic verb form (i.e., without a auxiliary) suggest that the probe relevant for agreement between the external argument and the auxiliary is located on T. As illustrated in (8), person agreement patterns alongside gender agreement in analytic (i.e., periphrastic) biabsolutive constructions.[4] As we would expect, the internal

---

[4] In the unmarked case, person agreement in Lak is controlled by an absolutive argument (i). Only finite clauses exhibit person agreement (Radkevich 2017; the same holds true for other Nakh-Daghestianian languages, such as Mehweb; see Ganenkov 2019).

(i) *Lak person agreement in perfective* (Radkevich 2017):
a. Na ina 0-uwhunu 0-ur-a.
I.I.SG you.SG.I.ABS I.SG-I.SG.catch.PRF.GER I.SG-AUX-1/2SG
'I caught you.'
b. Na ga 0-uwhunu 0-ur-0.
I.I.SG he.I.SG.ABS I.SG-I.SG.catch.PRF.GER I.SG-AUX-3
'I caught him.'

Note also that Lak personal pronouns exhibit case syncretism in 1st and 2nd person (+PARTICIPANT): the same form of the pronoun is used in absolutive and ergative contexts (*na* ↔ 1SG.ERG/ABS, *ina* ↔ 2SG.ERG/ABS). Ergative and biabsolutive constructions involving 1st or 2nd person external arguments can be differentiated on the basis of person agreement on the auxiliary with an absolutive argument, such as as in (iia) vs. (iib).

(ii) *Case syncretism and person agreement* (Kazenin 1998: 99):
a. na q̄at̄a buw-nu bu-r
I.I.ERG house.III.ABS III.build-CON.PAST III.AUX-3SG
'I have built the house.'
b. na q̄at̄a buw-nu u-ra
I.I.ABS house.III.ABS III.build-CON.PAST I.AUX-1SG
'I have built the house.'

argument controls gender agreement on the lexical verb, whereas the absolutive external argument controls both gender and person agreement on the auxiliary in the examples in (8).[5]

(8) *Person agreement in analytic Biabsolutives* (Kazenin 1998: 98-99):

a. rasul ču buwh-u-nu u-r
Rasul.I.ABS horse.III.ABS III.catch-PAST-CON I.AUX-3SG
'Rasul has caught the horse.'

b. na ču buwh-u-nu u-ra
I.I.ABS horse.III.ABS III.catch-PAST-CON I.AUX-1SG
'Rasul has caught the horse.'

c. ninu na uh-l-ej du-r
mother.II.ABS I.I.ABS I.catch-DUR-CON.PRES II.AUX-3SG
'Mother is catching me.'

If we turn now to synthetic ergative and biabsolutive constructions, we see that they exhibit the same basic agreement pattern as their respective analytic counterparts. The internal argument controls gender agreement in both instances and person agreement in the ergative constructions in (9a). However in biabsolutive constructions (9b), person agreement is with the absolutive external argument (paralleling agreement on auxiliaries in periphrastic biabsolutives). Provided that without an overt auxiliary there is no need to postulate Aux in synthetic biabsolutives and that synthetic and analytic biabsolutives show the same agreement pattern, I assume that they involve the same structure. I propose that AuxP can be dispensed with (contra Gagliardi et al. 2014, Ganenkov 2016) in the constructions at hand and that T hosts φ-probes in both synthetic and analytic biabsolutives (the only difference being that T spells out an overt auxiliary in the latter).

(9) *Synthetic Ergative and Biabsolutive constructions* (Kazenin 2013: 59):

a. Ga-n-al na uhlahi-s:a-ra
3SG-OS-ERG 1SG.I.ABS catch.I.PROG-ASSRT-1SG
'He is catching me.'

---

[5] Note that glossing of the examples is adopted from the literature and does not necessarily reflect on my view on the structure of auxiliaries.

b. Ga na uhlahi-s:a-r
3SG.ABS 1SG.I.ABS catch.I.PROG-ASSRT-3SG
'He is catching me.'

Turning to case assignment in Lak, I follow Gagliardi et al. (2014), Ganenkov (2016) and Radkevich (2017) in assuming that ergative is assigned to the external argument inside of vP, while a second absolutive case assignment (i.e., to the external argument in biabsolutives) must come from outside the vP. This is supported by the availability of the respective cases in vP *-awu*-masdars (only ergative but no absolutive on the external argument) in (6) above and TP *-šiwu*-masdars (absolutive on the external argument) in (7) above.

Moreover, Radkevich (2017) suggests that, unlike dative, the ergative in Lak is not an inherent case. While subjects of ergative constructions can surface with absolutive in biabsolutives, dative subjects cannot show this alternation (10).

(10) *Dative constructions* (Radkevich 2017):

a. Aˤli-n matematika q:a-d-urč'laj
Ali.I.SG-DAT math.IV.SG.ABS NEG-IV.SG-understand.PROG
d-ur-∅.
IV.SG-AUX-3
'Ali does not understand math.'

b. *Aˤli matematika q:a-d-urč'laj
Ali.I.SG.ABS math.IV.SG.ABS NEG-IV.SG-understand.PROG
∅-ur-∅.
I.SG-AUX-3

For the purpose of this paper, I follow Radkevich (2017) in assuming that case assignment in Lak is configurational (Marantz 1991, McFadden 2004, Bobaljik 2008, Levin 2017, i.a.). As illustrated in (11), dependent ergative is assigned to a DP which c-commands another DP within vP, provided the former did not receive inherent case before.

(11) *Disjunctive case hierarchy* (Radkevich 2017):

a. inherent/lexical case is assigned
b. dependent case (ergative) is assigned to a DP which c-commands another DP in the minimal vP.

c. default case (absolutive) is assigned

Tying in with the approaches by both Ganenkov (2016) and Radkevich (2017), who show that the external argument has to be dislocated from the vP in order to surface as absolutive subjects in biabsolutive constructions, I propose that the external argument moves to a position outside of TP, namely Spec,TopP, from which it can control agreement on the auxilary.

Biabsolutives in Lak (and other Nakh-Daghestanian languages) are typically associated with agent topicalization (Kazenin 1998, Schulze 2007, Forker 2012, 2019) or at least agent emphasis (Gagliardi et al. 2014). Thus, the agent is interpreted as the "semantic centre of the construction [i.e., biabsolutives]" (Forker 2012: 80) and is affected by the progressive action (Gagliardi et al. 2014), as shown in the biabsolutive construction in (12).

(12) *Biabsolutive agent emphasis* (Gagliardi et al. 2014: 144):
Aˤli q:ata b-ullaj ∅-ur.
Ali.I.ABS house.III.ABS III-do.PROG I-AUX
'Ali is in the state of building a house.' (=house-building currently affects his life)

Moreover, Lak biabsolutives are not possible with inanimate external arguments. According to Forker (2012: 84), (13a) is rejected on the basis of an implied voluntary action by the wind, while the ergative construction in (13b) can host an inanimate external argument. Thus, Forker (2012) assumes a ban on inanimate subjects to be agent topicalized in Biabsolutive constructions.

(13) *Inanimate subjects* (Forker 2012: 83):
a. *marč nuz t'it'l-ej b-u-r
wind.III.ABS door.IV.ABS open-DUR-CVB III-AUX-3SG
b. murčal nuz t'it'l-ej d-u-r
wind.III.ERG door.IV.ABS open-DUR-CVB IV-AUX-3SG
'The wind is opening the door.'

Similar differences in interpretation are reported for other Nakh-Daghestanian languages, such as Inguish (14) where agents are topicalized in biabsolutives and Mehweb (15), where non-agentive subjects are also banned from biabsolutives.

(14) *Ingush Biabsolutives and agent topicalization* (Forker 2012: 80-81):

a. txy naana maasha b-ezh
1PL.EX.GEN mother(J) homespun(B) B-make.CVB
j-ar.
J-PROG.PST
'Our mother made homespun.' (='Our mother was one of the people who could make homespun')

b. txy naanaz maasha b-ezh
1PL.EX.GEN mother.ERG homespun(B) B-make.CVB
b-ar (so dwachyvealcha).
B-PROG.PST 1SG PREV.in.V.go.TEMP.CVB
'Our mother was making homespun (when I came in).'

(15) *Mehweb non-agentive biabsolutive* (Ganenkov 2019: 228):

a. ??ʁʷaˤr ʁut'-be šiš d-uk'-aq-uwe
wind.ABS tree-PL.ABS move NPL-LV.IPF-CAUS-CVB.IPFV
le-b.
AUX-N
'The wind is shaking the trees.'

b. *c'a qul-le ig-uwe le-b.
fire.ABS house-PL.ABS burn.IPF-CVB.IPFV AUX-N
'A fire is burning the house.'

Based on these data, it seems that external arguments in biabsolutives are the topic of the clause. Assuming that topic is reflected in syntactic representation (Polinsky and Potsdam 2001: 593 and references therein), I take the relevant position for biabsolutive subjects expressing topic to be the specifier of TopP immediately dominating TP (Culicover 1991, Müller and Sternefeld 1993, Hoekstra and Zwart 1994, Rizzi 1997, Polinsky and Potsdam 2001).

With the basic assumptions about the syntax of Lak biabsolutives in place, I will rule out two conceivable alternatives below before I turn to my analysis on the basis of anti-local Agree in Section 3.3.

### 3.2. Against Case-Based and Biclausal Approaches

Evidence against treating the biabsolutive constructions in Lak on the basis of (morphological) case (i.e., absolutive arguments control agreement on the nearest verbal heads, while ergative arguments cannot control agreement) and in favour of a syntactic analysis comes from different dialects of closely-related Dargwa (Nakh-Dagestanian, Russia). Tanti Dargwa (Sumbatova 2014, 2019, Sumbatova and Lander 2014*a*,*b*, Belyaev 2016), Aquasha Dargwa (van den Berg 1999, Ganenkov 2018) and Sanzhi Dargwa (Forker 2019) exhibit so-called alternating agreement whereby in a periphrastic construction the ergative external argument optionally controls agreement on the auxiliary, while the lexical verb agrees with the absolutive internal argument. In (16a), (17a) and (18), the agreement morphology on the auxiliary corresponds to the ergative argument.

(16) *Tanti Dargwa* (Belyaev 2016: 88):

a. murad-li t’ant’i-b qali b-irq’.u.le sa-j
Murad.M-ERG in.Tanti-N house.N N-building AUX-M

b. murad-li t’ant’i-b qali b-irq’.u.le sa-b
Murad.M-ERG in.Tanti-N house.N N-building AUX-N
‘Murad is building the house.’

(17) *Aqusha Dargwa* (Ganenkov 2018: 531):

a. unra-ni kaʁar b-uč’-uli saj
neighbour-ERG letter.ABS N-read:IPF-CONV AUX.M

b. unra-ni kaʁar b-uč’-uli sabi
neighbour-ERG letter.ABS N-read:IPF-CONV AUX.N
‘The neighbour is reading a letter.’

(18) *Sanzhi Dargwa* (Forker 2019: 385):
it-i-l di-c:e d-urs-ul ca-r
that-OBL-ERG 1SG-IN NPL-tell-ICVB AUX-F
‘She tells (stories) to me.

Structurally, the alternating agreement constructions in Dargwa and the biabsolutives in Lak bear a striking similarity: in periphrastic constructions auxiliary agreement alternates between different controllers, while agreement of the lexical verb is always with the internal argument. A further similarity is the reliance on topicality. As shown above in Section 3.1, biabsolutive constructions in Lak are only possible if the external argument receives an agentive topic reading (Kazenin 1998, Schulze 2007, Forker 2012); they are not possible with inanimate external arguments (Forker 2012). According to Sumbatova (2014), Sumbatova and Lander (2014*a*), Belyaev (2016), the alternation between auxiliary agreement controllers in Tanti Dargwa is also conditioned by topicality and animacy. In (19), agreement between the ergative argument (‘dog’) and the auxiliary is ruled out because the absolutive internal argument (‘brother’) outranks it in animacy and topicality (Sumbatova and Lander 2014*a*, Belyaev 2016).

(19) *Tanti Dargwa* (Sumbatova and Lander 2014*a*: 453):
ʕe$^{ʕ}$la uc:i.li-ž se b-it.arg.ur.se? – hi.t ca χ:$^{w}$e-li
thy brother.M-DAT what.N N-happened that.M one dog.N-ERG
uc.ib =s:a-j / *=s:a-b
bite AUX-M AUX-N
‘What happened to your brother? A dog bit him.’

Bearing in mind those similarities, it seems that Biabsolutive constructions and alternating agreement constructions are virtually the same phenomena with the exception of (morphological) case exponence. While in Biabsolutive constructions the external argument is assigned absolutive if it controls agreement on the auxiliary (in contrast to ergative in the baseline construction), the agreement-controlling external argument in alternating agreement constructions still receives ergative case-marking. Clearly, a case-based approach, whereby only absolutive arguments are possible goals for agreement, is not available for the latter construction. I conclude that this also excludes a case-based analysis for Biabsolutives in Lak. Given the common structural and semantic properties, the two phenomena in closely-related languages should receive the same theoretical explanation.

Another potential line of analysis assumes a biclausal structure for biabsolutive constructions. In these kinds of approaches, the external argument is the subject of a separate clause headed by the auxiliary, while the internal argument is the only argument of a clause headed by the lexical verb.[6] In such a structure, the external argument is assigned absolutive as it is the sole argument of the clause and thus can control agreement on the auxiliary.

Evidence against these kinds of approaches come from Ā-movement. According to Gagliardi et al. (2014) and Radkevich (2017), Lak Ā-movement

---

[6] I am aware of the fact that *external argument* and *internal argument* are somewhat unsuitable terms in this context. External argument refers to the subject of the (surface) sentence alternating between ergative and absolutive case, while internal argument refers to the absolutive object.

is always clause-bound. In (20a), the *wh*-pronoun *ci* is contained within the embedded clause. Moving *ci* out of the embedded clause into the matrix clause, as in (20b), results in ungrammaticality.

(20) *Lak wh-movement is clause-bound* (Radkevich 2017):

a. Nit:i-n k'ul-s:a-r-iw, Rasul ci
mother-DAT know-ASSRT-PRS-Q Rasul.I.SG.ABS what.IV.ABS
d-ullaj-s:a-r-iw?
IV.SG-do.PRG-ASSRT-PRS-Q
'Does mother know what Rasul is building?'

b. *Ci$_i$ nit:i-n k'ul-s:a-r-iw, Rasul
what.IV.ABS mother-DAT know-ASSRT-PRS-Q Rasul.I.SG.ABS
$t_i$ d-ullaj-s:a-r-iw?
IV.SG-do.PRG-ASSRT-PRS-Q

In Lak Biabsolutive constructions, both arguments can undergo *wh*-movement: Thus in (21a), the external argument *cu wh*-moves to the edge of the clause. Crucially, the internal argument *ci* can also be Ā-moved to the edge of the clause, as shown in (21b). Given the fact that Ā-movement in Lak is clause-bound, we can conclude that this movement step does not cross a clause-boundary as it results in a grammatical expression.

(21) *Ā-movement in Biabsolutives* (Gagliardi et al. 2014: 148):

a. Cu$_i$ $t_i$ q:at:a b-ullaj Ø-ur?
who.I.ABS house.III.ABS III-do.PROG I-AUX
'Who is building the house?'

b. Ci$_i$ Aˤli $t_i$ b/d-ullaj Ø-ur?
what.ABS Ali.I.ABS III/IV-do.PROG I-AUX
'What is Ali building?'

This poses a problem for biclausal approaches (Kazenin 1998) or Basque-style analyses involving an embedded PP (Laka 2006). In both instances, the grammaticality of Ā-movement of the internal argument in Biabsolutive constructions (such as in (21b)) would be unexpected, as this movement step would cross a clause boundary (see also Forker 2012: 93–96 and Gagliardi et al. 2014: 161–162 for discussion). Biclausal approaches, thus, would not be able to generate these examples without postulating construction-specific exceptions for clause-bound Ā-movement. It seems reasonable to assume Biabsolutive constructions not to involve two separate clauses. Furthermore, ergative marking of the external argument in alternating agreement constructions (see above) could not be explained straightforwardly if it was the sole argument of the clause.

### 3.3. Deriving Biabsolutives

I propose that the difference in agreement on the auxiliary between ergative and biabsolutive constructions in Lak stems from an anti-locality constraint on Agree, repeated here in (22). In ergative constructions, the external argument in Spec,vP is too close to the φ-probe on T to engage in an agreement relation. In biabsolutives, on the other hand, the external argument dislocates to Spec,TopP and is, thus, anti-local enough for agreement with T.

(22) *Generalized Anti-Locality*: =(2)
$*[\ \ldots \alpha \ldots \beta \ \ldots]$ (where $\alpha$ and $\beta$ are participants in an Agree relation) unless there is a $\Gamma$ such that

a. $\Gamma$ is in the non-edge domain of a phrase XP.
b. $\alpha$ c-commands XP.
c. $\beta$ is reflexively included in $\Gamma$.

To recap, in ergative constructions the internal argument controls agreement on both the lexical verb and the auxiliary; c.f. (23). In (24), we see that Agree between v and the internal Argument is not blocked by anti-locality as $\text{DP}_{\text{Int}}$ is in the non-edge domain (the complement) of VP and v c-commands VP.[7] This successful derivational step yields agreement on the lexical verb. However, when the φ-probe on T (responsible for agreement on the auxiliary) triggers Agree, anti-locality excludes the in-situ external argument from being a possible goal. As $\text{DP}_{\text{Ext}}$ is in the edge domain of vP, there is no $\Gamma$ in the non-edge domain of a phrase c-commanded by T between the participants of the desired Agree relation. Instead, T also finds the internal argument in its c-command domain and agrees with it. According to the case hierarchy in

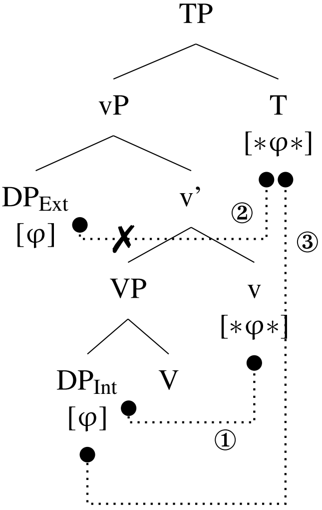
---

[7] The notation for the features triggering syntactic operations is adopted from Heck and Müller (2007): Agree is triggered by probe features on a head [∗F∗], while (Internal) Merge is triggered by structure-building features [•F•].

(11), the external argument is assigned ergative and the internal argument receives default absolutive.

(23) *Ergative construction* (Gagliardi et al. 2014: 144): =(1a)
Aˤli-l q:ata b-ullaj b-ur.
Ali.I-ERG house.III.ABS III-do.PROG III-AUX
'Ali is building a house.'

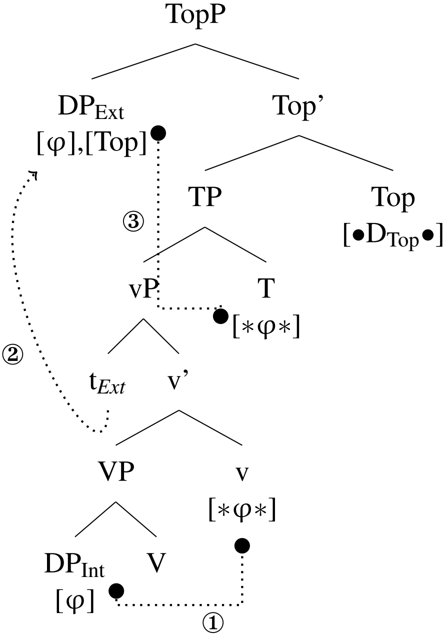
(24)



In biabsolutive constructions like (23), v again successfully probes for the φ-feature on the internal argument. Again, the probe in T would not be able to undergo Agree with the external argument in Spec,vP because of the anti-locality constraint. Suppose however, that probes in Lak have a directional bias for upward Agree (see Section 2) and that probing is suspended at least until the next functional head is merged to allow upward Agree to apply. When Top with a structure-building feature $[\bullet \text{D}_{\text{Top}} \bullet]$ is merged, it attracts DP$_{\text{Ext}}$ to its specifier (we assumed in Section 3.1 that the external argument is topicalized in biabsolutive constructions). The external argument is now in a position to undergo Agree with T, as T is in the non-edge domain of TP c-commanded by DP$_{\text{Ext}}$ (with DP$_{\text{Ext}}$ and T being participants of the Agree relation in the sense of the definition in (22)). At this point, the external argument cannot receive ergative as it is no longer in vP. Thus, both arguments are assigned absolutive; c.f. (26).[8]

(25) *Biabsolutive construction* (Gagliardi et al. 2014: 144): =(1b)
Aˤli q:ata b-ullaj 0-ur.
Ali.I.ABS house.III.ABS III-do.PROG I-AUX
'Ali is in the state of building a house.'

(26)



We have seen that anti-local Agree correctly derives the observed differences in agreement on the auxiliary between ergative and biabsolutive constructions. This difference does not stem from varying probes on T (responsible for auxiliary agreement) or the inability of ergative arguments to control agreement (see the discussion in Section 3.2); rather, it can be traced back to the position that the external argument occupies in the respective constructions. In an ergative construction, agreement between T and the in-situ subject in Spec,vP is ruled out by anti-locality: the external argument is not a valid goal for Agree

---

[8] Note that for Dargwa dialects with alternate agreement where agreement-controlling subjects still bear ergative (Section 3.2), we would have to assume a case mechanism where the smallest case domain still includes includes TopP, not just vP, as in (11b).

as it is too close to the probe. In biabsolutives, on the other hand, the subject moves to Spec,TopP and controls upward agreement from this position while still satisfying the anti-locality constraint.

## 4. Problems with Cyclicity

There are, however, two problems concerning cyclicity that come with the approach outlined in Section 3.3. Firstly, the internal argument should already have undergone cyclic spell-out (in the sense of phases; Chomsky 2000, 2001, 2004) before T probes and secondly, suspending bidirectional Agree on T until the external argument is dislocated to Spec,TopP violates the Cyclic Principle (Perlmutter and Soames 1979), and, accordingly, the Earliness Principle (Pesetsky 1989, Pesetsky and Torrego 2001) and Featural Cyclicity (Richards 2001). If taken seriously, auxiliary agreement should not be possible in Lak counter to the fact.

The first problem arises under a phase-based approach to syntax where phase complements are spelled out cyclically and, as per the Phase Impenetrability Condition (PIC), only the phase head and edge are available to operations outside of the phase (Chomsky 2000, 2001, 2004; see also Richards 2012, Gallego 2020). If v is a phase head, as widely assumed, then its complement should be transferred upon merging the next functional head T. Consequently, the internal argument inside VP should be inaccessible for the φ-probe on T (step 3 in (24)). Thus, assuming downward probing as in this case, anti-locality defines the upper limit of T search space and the PIC the lower limit. In this scenario, both constraints conspire to make both arguments inaccessible for T, incorrectly predicting auxiliary agreement to be ungrammatical in Lak ergative constructions.

Fortunately, this problem can be solved trivially if T simply agrees with the valued φ-features on v instead of on any of the arguments. After successful probing (step 1 in (24)), v acquires the internal argument's φ-values and acts as a goal for Agree on T (similar to Cyclic Agree; Legate 2005). In fact, v is the only possible goal within the narrowed search space: it is a phase head (complying with the PIC) and in the non-edge domain of a phrase c-commanded by T (complying with Generalized Anti-Locality). Thus, auxiliary agreement in Lak ergative constructions is only seemingly with the internal argument, more precisely it should be analyzed as an Agree relation between T and v.[9]

The second problem emerges from the bidirectionality of Agree that we assumed for the purpose of the analysis. T has to be able to probe downwards (to agree with the internal argument/v) and upwards (to agree with the external argument in Spec,TopP). Given the Cyclic Principle (Earliness Principle, Featural Cyclicity), Agree on T should apply as soon as possible. This would, in turn, mean that Agree on T is always downwards with v, considering v is available earlier (i.e., as soon as T is merged) than the desired goal in Spec,TopP. Delaying probing until the external argument is re-merged in Spec,TopP to derive auxiliary agreement via upward Agree thus violates cyclicity.

In Section 3.3, this problem was circumvented with the help of a directional bias: Earliness is not violated as the structural description of Agree (at least for Lak) includes that upward Agree is always favored over downward Agree. Probing on T, thus, applies only if enough structure is built-up to realize anti-locality-satisfying upward Agree that (i.e., after Top is merged).

Interestingly, the directional bias can be dispensed with if Top is merged cyclically before T can discharge its probing feature.[10] Then, agreement between T and v would be ruled out by the Strict Cycle Condition (SCC; Chomsky 1973) as this operation would apply to a proper subdomain of the current cycle (i.e., TopP). This state of affairs can be achieved if we tentatively assume that subcategorization can be triggered by a feature on the lower head: In biabsolutive constructions, T has among its features a structure-building feature [•Top•] that triggers subcategorization of Top at the root TP (reciprocal subcategorization; see Popp and Tebay 2019 and references therein). Suppose now that this structure-building feature is ordered before the probing feature on T $[\bullet\text{Top}\bullet] \succ [*\varphi*]$ (see Heck and Müller 2007, Müller 2009, Georgi 2014, 2017 for sequential ordering of features). Then, Top is merged before T has the chance to probe and agree with v. As downward Agree between T and v is now blocked by the SCC, successful Agree is effectively delayed until the external argument is re-merged in Spec,TopP and becomes a valid goal.

The supposed directional bias of Agree in Lak thus emerges as a consequence

---

[9] This type of analysis also lends itself to agreement between T and nominative objects in Icelandic quirky case constructions (Taraldsen 1995, Chomsky 2001) and German unaccusative constructions (Grewendorf 1989).

[10] Thanks to Sören Tebay (p.c.) for this idea; see also Privizentseva (2023).

of "early" subcategorization and strict cyclicity. Moreover, as the SCC blocks operations between T and v,"delayed" agreement between Spec,TopP and T no longer violates the Cyclic Principle, the Earliness Principle or Featural Cyclicity, given that Agree cannot apply earlier than after the movement step to Spec,TopP (there is no other possible goal in the derivation).

## 5. Conclusion

In this paper, I have argued that Generalized Anti-Locality constrains Agree and derives agreement phenomena where arguments more distant from the probe can be the controller of agreement, while arguments closer to the probe cannot trigger agreement.

The biabsolutive pattern in Lak is derived by movement of an agent external argument to Spec,TopP, from where it can control upward agreement on T restricted by Generalized Anti-Locality. In ergative constructions, on the other hand, the in-situ external argument in Spec,vP is too local for agreement with T in both possible positions. Thus, anti-locality rules out too-close syntactic dependencies also in regard to agreement. The emerging picture sees Agree being subject both to anti-locality (excluding extremely local controllers) and locality (selecting the closest anti-locality-obeying controller).

Problems with cyclicity arising from restricting the search space via anti-locality and delaying upward probing of T until an anti-local enough goal can be found (in Spec,TopP) are dealt with by "early" merging of Top via reciprocal subcategorization (Popp and Tebay 2019) ordered before φ-Agree on T. Consequently, strict cyclicity will prevent T from probing downward, paving the way for Agree between T and the external argument in Spec,TopP as the only remaining option.

## References

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

Baker, Mark C. (2008): *The syntax of agreement and concord*. Cambridge University Press, Cambridge.

Belyaev, Oleg (2016): Ergative gender agreement in Dargwa "backward control" or feature sharing?. *In:* D. Arnold, M. Butt, B. Crysmann, T. H. King and S. Müller, eds, *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure*

*Grammar and Lexical Functional Grammar, Polish Academy of Sciences, Warsaw, Poland*. CSLI Publications, Stanford, CA, pp. 83–103.

Bhatt, Rajesh (2005): ‘Long Distance Agreement in Hindi-Urdu’, *Natural Language & Linguistic Theory* **23**(4), 757–807.

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

Bobaljik, Jonathan D. (2008): Where’s Phi? Agreement as a Post-Syntactic Operation.. *In:* D. Harbour, D. Adger and S. Béjar, eds, *Phi-Theory: Phi features across interfaces and modules*. Oxford University Press, Oxford, pp. 295–328.

Bobaljik, Jonathan David and Susi Wurmbrand (2005): ‘The domain of agreement’, *Natural Language & Linguistic Theory* **23**, 809–865.

Bošković, Željko (1994): ‘D-Structure, Theta-Criterion, and Movement into Theta Positions’, *Linguistic Analysis* **24**, 247–286.

Bošković, Željko (2016): ‘On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling’, *The Linguistic Review* **33**(1), 17–66.

Branan, Kenyon (2022): ‘Locality and Anti-Locality: The Logic of Conflicting Requirements’, *Linguistic Inquiry* **54**(1), 1–38.

Brillman, Ruth and Aaron Hisch (2016): An Anti-Locality Account of English Subject/Non-Subject Asymmetries. *In:* R. Burkholder, C. Cisneros, E. R. Coppess, J. Grove, E. A. Hanink, H. McMahan, C. Meyer, N. Pavlou, Özge Sarıgül, A. R. Singerman and A. Zhang, eds, *Proceedings of the 50th annual meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, IL, pp. 73–87. Proceedings of CLS 50.

Carstens, Vicki (2016): ‘Delayed Valuation: A Reanalysis of Goal Features, “Upward” Complementizer Agreement, and the Mechanics of Case’, *Syntax* **19**(1), 1–42.

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

Chomsky, Noam (2000): Minimalist inquiries: The framework. *In:* R. Martin, D. Michaels, J. Uriagereka and S. J. Keyser, eds, *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*. MIT Press, Cambridge, MA, pp. 89–155.

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

Chomsky, Noam (2004): Beyond explanatory adequacy. *In:* A. Belletti, ed., *The Cartography of Syntactic Structures, vol. 3: Structures and Beyond*. Oxford University Press, New York, NY, pp. 104–131.

Chumakina, Marina and Oliver Bond (2016): Competing controllers and agreement potential. *In:* O. Bond, G. G. Corbett, M. Chumakina and D. Brown, eds, *Archi:*

*Complexities of Agreement in Cross-Theoretical Perspective*. Oxford University Press, Oxford, pp. 77–117.

Culicover, Peter W. (1991): Topicalization, inversion, and complementizers in English. *In:* D. Delfitto, M. Everaert, A. Evers and F. Stuurman, eds, *Going Romance and Beyond*. OTS Working Papers, Utrecht, pp. 48–62.

Deal, Amy Rose (2019): 'Raising to Ergative: Remarks on Applicatives of Unaccusatives', *Linguistic Inquiry* **50**(2), 388–415.

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

Forker, Diana (2012): 'The bi-absolutive construction in Nakh-Daghestanian', *Folia Linguistica* **46**(1).

Forker, Diana (2019): *A grammar of Sanzhi Dargwa*. Languages of the Caucasus 2, Language Science Press, Berlin.

Gagliardi, Annie, Michael Goncalves, Maria Polinsky and Nina Radkevich (2014): 'The biabsolutive construction in Lak and Tsez', *Lingua* **150**, 137–170.

Gallego, Ángel J. (2020): Strong and weak "strict cyclicity" in phase theory. *In:* A. Bárány, T. Biberauer, J. Douglas and S. Vikner, eds, *Syntactic architecture and its consequences II: Between syntax and morphology*. Language Science Press, Berlin, pp. 207–226.

Ganenkov, Dmitry (2016): 'Lak Biabsolutives are Raising Structures', *Higher School of Economics Research Paper.* WP BRP 46/LNG/2016.

Ganenkov, Dmitry (2018): 'Gender agreement alternation in Aqusha Dargwa', *Studies in Language* **42**(3), 529–561.

Ganenkov, Dmitry (2019): Case and agreement in Mehweb. *In:* M. Daniel, N. Dobrushina and D. Ganenkov, eds, *The Mehweb language: Essays on phonology, morphology and syntax*. Language Science Press, Berlin, pp. 189–234.

Georgi, Doreen (2014): Opaque Interactions of Merge and Agree: On the Nature and Order of Elementary Operations. PhD dissertation, University of Leipzig, Leipzig.

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

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

Grohmann, Kleanthes K. (2003): *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Benjamins, Amsterdam.

Grohmann, Kleanthes K. (2011): Anti-Locality: Too-Close Relations in Grammar. *In:* C. Boeckx, ed., *The Oxford Handbook of Linguistic Minimalism*. Oxford University Press, Oxford, pp. 260–290.

Heck, Fabian and Gereon Müller (2007): Extremely local optimization. *In:* E. Bainbridge and B. Agbayani, eds, *Proceedings of WECOL 17*. California State University, Departmentof Linguistics, Fresno, CA, pp. 170–183.

Himmelreich, Anke (2016): Case matching as bidirectional Agree. *In:* K. Barnickel, M. G. Naranjo, J. Hein, S. Korsah, A. Murphy, L. Paschen, Z. Puškar and J. Zaleska, eds, *Replicative Processes in Grammar*. Linguistische Arbeitsberichte 93, Universität Leipzig, Leipzig, pp. 259–300.

Hoekstra, Eric and Jan-Wouter Zwart (1994): 'De Structuur van de CP: Functionele Projecties voor Topics en Vraagwoorden in het Nederlands', *Spektator* **23**, 191–212.

Kazenin, Konstantin (1998): On Patient Demotion in Lak. *In:* L. Kulikov and H. Vater, eds, *Typology of Verbal Categories*. Niemeyer, Tübingen, pp. 95–115.

Kazenin, Konstantin (2013): 'Sintaksis lakskogo jazyka'. Ms., Moscow.

Laka, Itziar (2006): Deriving Split Ergativity in the Progressive: The case of Basque. *In:* A. Johns, D. Massam and N. J, eds, *Ergativity: Emerging Issues*. Studies in Natural Language & Linguistic Theory, Springer, Dordrecht, pp. 173–195.

Lee, Hyunjung (2020): 'Reconstruction Effects for Binding and Antilocality'. Ms., Universität Leipzig.

Legate, Julie Anne (2005): Phases and Cyclic Agreement. *In:* M. McGinnis and N. Richards, eds, *Perspective on Phases*. MIT Working Papers in Linguistics, MIT Press, Cambridge, MA, pp. 147–156.

Levin, Theodore (2017): 'Successive-cyclic case assignment: Korean nominative-nominative case-stacking', *Natural Language & Linguistic Theory* **35**(2), 447–498.

Marantz, Alec (1991): Case and licensing. *In:* G. Westphal, B. Ao and H.-R. Chae, eds, *Proceedings of the 8th Eastern States Conference on Linguistics (ESCOL 8)*. CLC Publications, Ithaca, NY, pp. 234–253.

McFadden, Thomas (2004): The position of morphological case in the derivation: A study on the syntax-morphology interface. PhD thesis, Univeristy of Pennsylvania.

Müller, Gereon (2009): Ergativity, accusativity, and the order of Merge and Agree. In: *Explorations of phase theory: Features and arguments*. de Gruyter, Berlin, pp. 269–308.

Müller, Gereon (2020): 'Current concepts in Syntactic Theory: Anti-locality'. Lecture notes, Universität Leipzig.

Müller, Gereon and Wolfgang Sternefeld (1993): 'Improper Movement and Unambiguous Binding', *Linguistic Inquiry* **24**(3), 461–507.

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

Pesetsky, David (1989): 'Language-particular processes and the Earliness Principle'. Ms., MIT.

Pesetsky, David and Esther Torrego (2001): T-to-C Movement: Causes and Consequences. *In:* M. Kenstowicz, ed., *Ken Hale: a Life in Language*. MIT Press, Cambridge, MA, pp. 355–426.

Polinsky, Maria (2016): Agreement in Archi from a Minimalist perspective. *In:* O. Bond, G. G. Corbett, M. Chumakina and D. Brown, eds, *Archi: Complexities of*

*Agreement in Cross-Theoretical Perspective*. Oxford University Press, Oxford, pp. 184–232.

Polinsky, Maria and Eric Potsdam (2001): 'Long-distance agreement and topic in Tsez', *Natural Language & Linguistic Theory* **19**, 583–646.

Popp, Marie-Luise and Sören Eggert Tebay (2019): 'No possessor inversion in German PPs', *Glossa* **4**(1).

Preminger, Omer and Maria Polinsky (2015): 'Agreement and semantic concord: a spurious unification'. lingbuzz/002363.

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.

Radkevich, Nina (2017): 'Two types of split ergativity= two sides of the same coin'. Talk held at LAGB Annual Meeting, University of Kent.

Richards, Marc D. (2012): On feature inheritance, defective phases, and the movement-morphology connection. *In:* Ángel J. Gallego, ed., *Phases*. de Gruyter, Berlin, pp. 195–232.

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

Rizzi, Luigi (1997): The Fine Structure of the Left Periphery. *In:* L. Haegeman, ed., *Elements of Grammar*. Kluwer Academic Publishers, Dordrecht, pp. 281–337.

Saito, Mamoru and Keiko Murasugi (1999): Subject predication within IP and DP. *In:* K. Johnson and I. Roberts, eds, *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*. Kluwer, Dordrecht, pp. 167–188.

Schulze, Wolfgang (2007): 'The Lak Language'. Ms., LMU Munich.

Sumbatova, Nina (2014): V poiskax podležaščego: kontrol' klassnogo soglasovani-jai priznaki grammatičeskogo prioriteta v darginskom jazyke [Lookingfor the subject: Gender agreement and grammatical priority in Dargwa].. *In:* M. Danièl, E. Ljutikova, V. Plungian, S. Tatevosov and O. Fedorova, eds, *Jazyk. Konstanty. Peremennye. Pamjati Aleksandra Evgen'eviča Kibrika, [Language.Constants. Variables. In memoriam Alexander E. Kibrik].*. Aleteja, St. Petersburg.

Sumbatova, Nina (2019): 'Agreement of essive adverbials in Tanti Dargwa'. Talk presented at 52nd Annual Meeting of the Societas Linguistica Europaea, Leipzig.

Sumbatova, Nina and Yury Lander (2014*a*): *Darginskij govor selenija Tanty: Grammatičeskij očerk, voprosy sintaksisa* [Tanti Dargwa: Grammar survey and essays on syntax]. Jazyki slavjanskoj kul'tury, Moscow.

Sumbatova, Nina and Yury Lander (2014*b*): The Dargwa Variety of the Tanti Village. A Grammatical Sketch. Issues in Syntax. Research paper no. wp brp, 12.

Taraldsen, Knut Tarald (1995): On agreement and nominative objects in Icelandic. *In:* H. Haider, S. Olsen and S. Vikner, eds, *Studies in comparative Germanic syntax*. Kluwer, Dordrecht, pp. 307–327.

van den Berg, Helma (1999): ‘Gender and person agreement in Akusha Dargi’, *Folia Linguistica* **33**(1-2), 153–168.

Zyman, Erik (2021): ‘Antilocality at the Phase Edge’, *Syntax* **24**(4), 510–556.

# Towards a Unified Explanation of Apparent Cases of Counter-cyclicity

Fabian Heck*

**Abstract**

The Strict Cycle Condition has proven to be an essential constraint on syntactic derivations. Despite this, various analyses have been proposed over the years that (explicitly or implicitly) assume syntactic operations applying in a counter-cyclic fashion. Presupposing that both the SCC and the gist of these proposals are correct, the following questions arise: a) Is it possible to come up with strictly cyclic reformulations of these proposals that preserve their general gist? b) Is there a uniform strictly cyclic account that covers all types of analyses? The present paper answers both questions in a constructive way by offering such a uniform and strictly cyclic account of the different types of apparently counter-cyclic analyses in terms of non-monotonic derivations.

## 1. Introduction

The Strict Cycle Condition (SCC) was introduced by Chomsky (1973) as a means to constrain syntactic derivations. In essence, the SCC states that a cyclic domain D that has been subject to syntactic operations at earlier stages of the derivation must not be revisited and thus be modified at later stages if the modification exclusively affects D. (1) displays the original formulation of the SCC given in Chomsky (1973: 243-245).

(1) *Strict Cycle Condition*:
No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

The exact nature of cyclic domains is left open in (1). In what follows, I adopt the most restrictive view that every syntactic node generated by Merge

---

*Acknowledgments go to the audience of the Mini-Workshop “(Strict) Cyclicity” of the RU “Cyclic Optimization” at Leipzig University. I especially thank Eric Fuß, Gereon Müller, and Mariia Privizentseva for helpful comments.

*Cyclicity*, 297–328
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

(Chomsky 1993, 1995) constitutes a cyclic domain. (1) is also unspecific as to the syntactic operations that are subject to the SCC. It seems that in later work, strict cyclicity was mostly conceived of being a property of Merge (cf. the Extension Condition in Chomsky 1993, 1995 or the No-Tampering Condition in Chomsky 2008). Embracing the original view of Chomsky (1973), I assume that the SCC applies to all syntactic operations, thus comprising both (internal and external) Merge as well as Agree (Chomsky 2000, 2001).

A detailed motivation of the SCC cannot be provided in the present paper. It may seem that with the changes that syntactic theory underwent in the last 50 years, many of the arguments in favor of the SCC have lost their force (see Freidin 1978, 1999, Browning 1991, Boeckx 2003). While this may be correct, there remain good reasons to assume the SCC. In fact, it seems to me that the original idea presented in Chomsky (1973), which motivated the SCC as a means to enforce Minimality (back then: Superiority), can be maintained today, albeit in a way that adapts to more modern theorizing (cf. Riemsdijk and Williams 1986, Freidin 1992, Kitahara 1997, Bošković and Lasnik 1999, Heck 2018 for relevant discussion). In what follows, I therefore assume that the SCC is well motivated and in good health today.

Against this background, it might be surprising that various types of analyses have been proposed over the years that (explicitly or implicitly) assume that syntactic operations may apply in a counter-cyclic fashion. Taking these proposals seriously, the question arises how they can be reconciled with the SCC. In particular, one may ask whether a uniform approach is possible that reformulates each of these counter-cyclic proposals in a strictly cyclic manner while at the same time preserving the gist of the respective analysis. The present paper contains such a proposal. §2 lists the counter-cyclic analyses that have been proposed in the literature that I am aware of. §3 briefly introduces the background that the present proposal is based on, the theory of non-monotonic derivations. §4 contains a strictly cyclic reformulation for each of the counter-cyclic proposals. Finally, §5 concludes.

## 2. Counter-cyclic Proposals

### 2.1. Head-Movement/Undermerge

The first counter-cyclic operation to be discussed is head-movement. Head-movement is a widespread and well-established analytical tool (cf. already

McCawley 1968, 1970). It is typically motivated by contexts where one head shows up in various positions (within or across languages). A textbook example is the difference between English and French with respect to the placement of finite main verbs relative to VP-adverbs. While in French a finite main verb precedes a manner adverb such as *often*, a finite main verb in English must follow the same type of adverb (Kayne 1975, Emonds 1976, Pollock 1989). According to the head-movement analysis, finite main verbs in French move out of vP to combine with the higher, preceding T-head, thereby crossing the adverb (2a). In contrast, no such head-movement takes place in English (2b).

(2) a. [$_{\mathrm{TP}}$ Nous embrass-ons [$_{\mathrm{vP}}$ souvent __ Marie ]].
we kiss-1PL.PRES often Marie
'We often kiss Marie.'
b. [$_{\mathrm{TP}}$ We [$_{\mathrm{vP}}$ often kiss Mary ]].

In many cases, the targeted c-commanding head has an overt exponent. For instance, the analysis in (2a) assumes that the T-head is realized by the inflectional affix *-ons*. This indicates that the higher head is not replaced ('substituted') by movement of the lower head. Rather, the lower head adjoins to the higher head, forming a complex head (see Baker 1988; but cf. Roberts 2010 for an alternative analysis of head-movement). Adjunction to the higher head appears to violate the SCC. It applies to a cyclic domain, the T-head in (2a), that is (immediately) dominated by another cyclic domain, the TP.

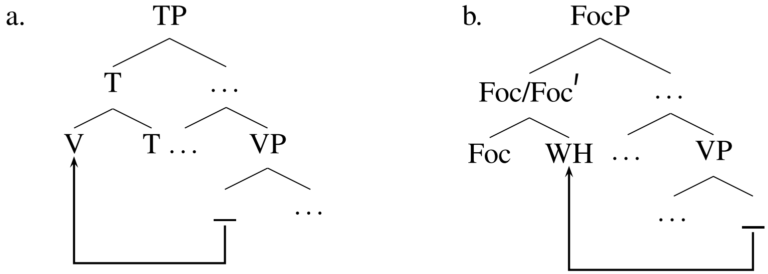
A counter-cyclic operation that is closely related to head-movement has been proposed by Pesetsky (2013) and is called Undermerge. Just like head-movement, Undermerge combines a category with a higher head. Unlike head-movement, however, the moved category targeted by Undermerge is a phrase (see already Sportiche 2005). Yuan (2017) offers an analysis of *wh*-movement in Kikuyu in terms of Undermerge. One of the motivations for this Undermerge analysis comes from the fact that the moved *wh*-phrase in Kikuyu follows the focus head *nĩ*, which is assumed to be the movement trigger:

(3) [$_{\mathrm{FocP}}$ nĩ kĩĩ [$_{\mathrm{TP}}$ mwana a-ta-na-rug-a __ ]]?
FOC what child 1SM-NEG-PST-cook-FV
'What didn't the child cook?'

According to Pesetsky (2013), a phrase that undergoes Undermerge literally becomes the complement of a higher head. Alternatively, to make the parallelism to head-movement even clearer, one might think of the moved phrase as adjoining to the higher head. In any event, Undermerge is counter-cyclic: The moved phrase targets a cyclic domain (the focus head in (3)) that is dominated by another cyclic domain (the FocP).

The parallelism between head-movement (in French) and Undermerge (in Kikuyu) is illustrated in (4a,b).

(4)



One might try to avoid a violation of the SCC by head-movement/Undermerge by stipulating that adjunction (in contrast to Merge) is not subject to the SCC (and by assuming that Undermerge involves adjunction). This is not sufficient to capture other cases of apparent counter-cyclicity, however.

## 2.2. Minimality

In many languages, an experiencer blocks raising to SpecT out of an embedded infinitive (see, e.g., Italian (5a), French (5b), Icelandic (5c); see McGinnis 1998 and references therein).

(5) a. *Gianni sembra a Piero [$_{TP}$ _ fare il suo dovere ].
Gianni seems to Piero to.do the his duty
'Gianni seems to Piero to do his duty.'

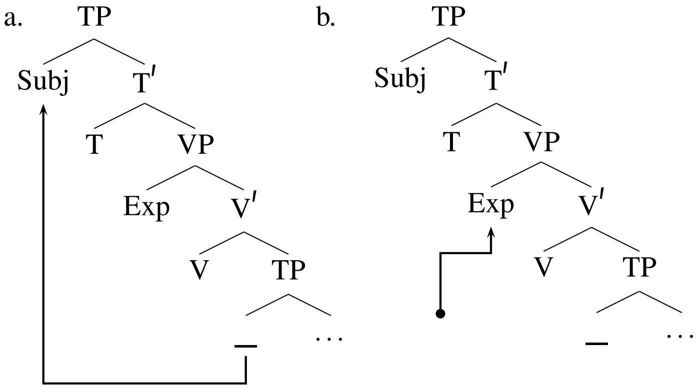
b. *Jean semble à Marie [$_{TP}$ _ avoir du talent ].
Jean seems to Marie have of.the talent
'Jean seems to Marie to be gifted.'

c. *Ólafur virðist mér [$_{TP}$ _ vera gáfaður ].
Olaf.NOM seems me.DAT to.be intelligent
'Olaf seems to me to be intelligent.'

Curiously, in English such raising is fine, see (6). There, the experiencer does not seem to induce the intervention effect that is usually interpreted as a violation of Minimality, see the analysis in (7a). The alternative analysis, raising to SpecT followed by counter-cyclic merger of the experiencer (Stepanov 2001*a*,*b*), solves the Minimality problem for English, but only at the costs of violating the SCC (7b).

(6) John seems to Mary [$_{TP}$ _ to be smart ].

(7)

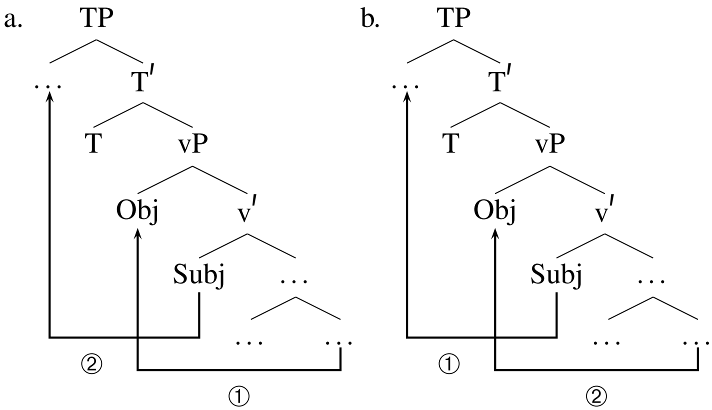


Another analysis that employs counter-cyclicity in order to come to grips with a Minimality problem is presented in Stepanov (2004). In a nutshell, the proposal is as follows: In a theory of ergativity where ergative case is assigned to the Subj by v and absolutive case is assigned to the DObj by T (Campana 1992, Murasugi 1992), one may expect the Subj in Specv to block absolutive assignment due to Minimality, contrary to fact. Stepanov's (2004) solution is to merge the Subj after the DObj has been assigned case, which is obviously counter-cyclic.

As a final example, one may approach an old problem arising with Scandinavian object shift in terms of late merger. There are reasons to assume that object shift (cf. (15)) targets an outer Specv, above the Subj (Chomsky 1993, Holmberg and Platzack 1995, Bobaljik and Jonas 1996, Anagnostopoulou

2003). This, however, creates the puzzle of how the Subj can undergo raising to SpecT (Branigan 1992, Chomsky 1993, 2000, 2001, Koizumi 1993, Kitahara 1997: chapter 3, Hiraiwa 2001, Dikken 2007). The shifted object, which is closer to T, should prevent such raising via Minimality (as first noted by Vikner 1989), see (8a).

(8)



Although it has not been proposed in the literature (but cf. Heck 2016 and section 4 below), there is an alternative counter-cyclic derivation, which first raises the Subj to SpecT and then performs object shift (late internal merger), see (8b). This derivation avoids the Minimality issue at the cost of weakening (or abandoning) strict cyclicity. (The proposal in Holmberg 1999 comes close to this type of analysis, however, it ultimately eschews counter-cyclicity by placing object shift in the PF-branch.)

## 2.3. Reconstruction

It is usually assumed that *wh*-movement (in general: Ā-movement) shows obligatory reconstruction behavior with respect to Principle C (Riemsdijk and Williams 1981, Lebeaux 1988, 1990). The ungrammaticality of (9a,b) can thus be traced back to the same source: a Principle C violation.

(9) a. *He$_i$ denied [$_{\text{DP}}$ the claim that John$_i$ was asleep ].

b. *[$_{\text{DP}}$ Which claim that John$_i$ was asleep ] did he$_i$ deny _?

c. [$_{\text{DP}}$ Which claim that John$_i$ was asleep ] did he$_i$ deny
[$_{\text{DP}}$ ~~which … John$_i$ …~~ ]?

A popular account of these facts involves the idea that movement leaves behind a copy. This copy is not spelled out at PF (indicated by strike through: ~~copy~~). Since the moved category in (9b) contains the R-expression *John*, so does the copy (see (9c)). It is this latter instance of *John* which remains in the c-command domain of the co-indexed pronoun *he* in (9b,c), thus triggering the Principle C violation (Chomsky 1995, Sauerland 1998, Fox 1999).

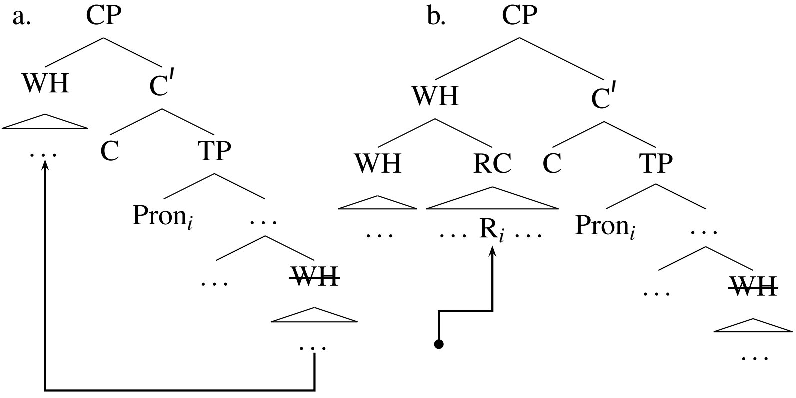
While (9) involves a complement clause to a noun, it has been observed (Lebeaux 1988, 1990) that in the case of a relative clause (often assumed to be adjoined to the nominal projection), the Principle C effect observable for (9b) vanishes (10b). This is surprising if Ā-movement always leaves a copy.

(10) a. *He$_i$ later denied [$_{\text{DP}}$ the claim that John$_i$ had made ].

b. [$_{\text{DP}}$ Which claim that John$_i$ had made _ ] did he$_i$ later deny _?

A common interpretation of this effect (due to Lebeaux 1988, 1990) is that the relative clause may enter the derivation *after* the noun has moved (e.g., Chomsky 1995, Fox 1999, Fox and Nissenbaum 1999, Takahashi 2006, Lebeaux 2009, Takahashi and Hulsey 2009). This is called late merger (but cf. Sportiche 2019 for criticism). Thus, the copy left behind by movement actually does not contain the relative clause (and therefore not the offending R-expression). The relevant steps of the derivation (*wh*-movement and late merger of the relative clause) are displayed in (11a,b).

(11)



Late merger in (11b) is counter-cyclic. Assuming that the relative clause is adjoined, one may resort to the idea that adjunction is exempt from the SCC.

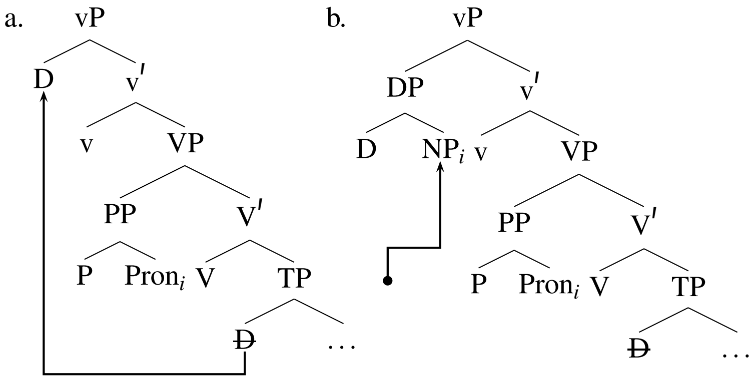
However, Takahashi and Hulsey (2009) and Lebeaux (2009) argue that late merger may also apply to NP complements of determiners. This assumption is motivated by the fact that in contrast to *wh*-movement (Ā-movement), raising (A-movement) appears to show optional reconstruction for Principle C (i.e., a Principle C violation can be avoided by A-movement; Chomsky 1995, Fox 1999, Lebeaux 2009), see (12a,b).

(12) a. [$_{\text{DP}}$ The boys ]$_i$ seemed to each other$_i$ [ _ to be smart ].

b. [$_{\text{DP}}$ John$_i$'s mother ] seems to him$_i$ [ _ to be beautiful ].

The idea is to account for the lack of obligatory reconstruction in (12a,b) by a derivation that involves the steps displayed in (13a,b): A-movement of a bare determiner D plus subsequent late merger of the complement of D.[1,2]

(13)



Again, late merger in (13b) violates the SCC. Moreover, since this is arguably a case of merging a complement, one cannot resort to the stipulation that adjunction is exempt from strict cyclicity.

Finally note that in order to avoid such late merger of complements of D

---

[1] There is reason to believe that the preposition *to* in (12) does not hinder c-command by the pronoun over the R-expression (see Chomsky 1995, Pesetsky 1995, McGinnis 1998).

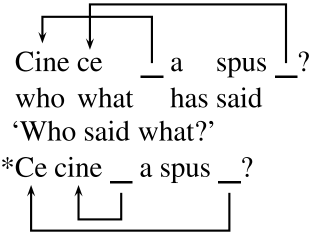
[2] (12a) suggests that only the complete DP *the boys* forms an R-expression, not the definite determiner on its own. (12b) seems to require that the complement *mother (of) John('s)* undergoes late merger, followed by subsequent counter-cyclic DP-internal raising of *John('s)*. This means that not only external Merge but also internal Merge must be able to apply counter-cyclically (cf. Lechner 2019).

with Ā-movement, where reconstruction for Principle C is obligatory (recall (9b)), Takahashi and Hulsey (2009) propose a constraint to the effect that NP must merge with D before case is assigned to the DP (cf. Lebeaux 2009, Stanton 2016, Lechner 2019). In raising contexts such as (12a,b), nominative case is assigned to the raised DP by the matrix T-head. Takahashi and Hulsey (2009) assume that raising makes an intermediate stop in Specv (as indicated in (13); see also Legate 2003, Sauerland 2003, Richards 2004, Deal 2009). In this position, the raised D is outside the c-command domain of the co-indexed pronoun, and the head assigning its case has not been merged yet. Thus late merger can still apply. No such point of the derivation is available for cases of Ā-movement like (9b). In (9b), v assigns accusative case to the Ā-moving DObj. Thus, the D-head of the DObj in (9b) must be merged with its complement while it is still the sister of V. But then the R-expression in the complement of D will be c-commanded by the co-indexed pronoun in Specv, leading to a Principle C violation.

## 2.4. Tucking-In

If a head H triggers multiple instances of movement, and if these movements are of the same type, i.e., are triggered by the same feature, then the moved categories usually target multiple specifiers of H in an order preserving way. Put differently, the movement paths show a crossing pattern.[3]

This generalization is made explicit in McGinnis (1998), and it is well established (see, e.g., Müller 1997, Richards 1997). (14) illustrates the crossing paths that show up with multiple *wh*-movement in Romanian (Rudin 1988). (15) shows crossing paths arising with multiple object shift in Danish (Vikner 1989).



(14) a. Cine ce _ a spus _?
who what has said
'Who said what?'
b. *Ce cine _ a spus _?

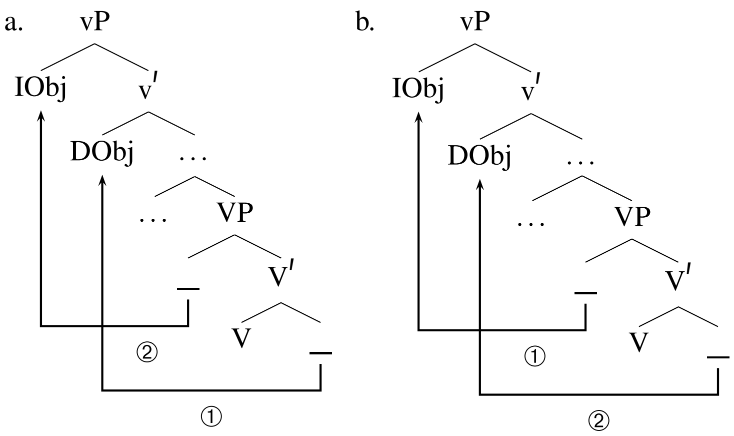
---

[3] In contrast, if multiple movements to the same specifier domain are triggered by different features, then the movement paths are nested, flipping the order of the moved categories.

(15) a. Peter viste hende den jo _ _.
Peter showed her it indeed
'Peter indeed showed it to her.'
b. *Peter viste den hende jo _ _.

Order preserving movement that generates crossing paths (exemplified in (16) by multiple object shift) seems to either violate some kind of Minimality (Rizzi 1990, Fanselow 1991, Chomsky 1995), as in (16a), or the SCC, as in (16b), depending on what moves first. In (16a), the DObj moves to an inner specifier first, across the c-commanding IObj, violating Minimality; in (16b), the DObj moves second but targets an inner specifier, violating the SCC.

(16)



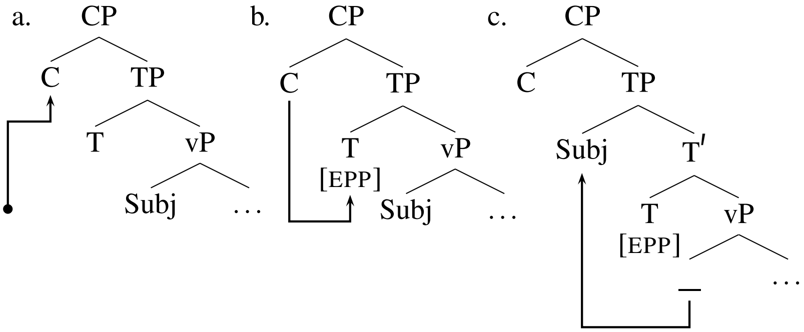
The most popular assumption is that the derivation in (16b) is the correct one. The hypothesis is that a specifier S created by internal Merge targets the lowest position within the specifier domain. If there already is a specifier S′ present in the same domain, then S undergoes 'tucking-in' below S′ (Richards 1997, 1999, Mulders 1997; see also Řezáč 2002, Doggett 2004, Branigan 2014, Bošković 2016, Safir 2019 for discussion). Thus, the tucking-in hypothesis involves counter-cyclic (internal) Merge.

## 2.5. Feature Inheritance

Chomsky (2007, 2008) proposes that certain $\phi$-features (probes in the sense of Chomsky 2000, 2001) and EPP-features are a property of phase heads

(C and v) only. If such features show up on another head H, such as for instance T, then this is because H inherits the relevant feature from a phase head under c-command. (See Richards 2007 for conceptual motivation for the idea of feature inheritance; cf. Broekhuis 2016 for empirical criticism.) The proposal seems to imply that the satisfaction of the inherited feature applies counter-cyclically (Richards 2007). For instance, C must merge with TP in order to be able to hand down its EPP-feature to T (17a,b).

(17)



Satisfaction of the inherited EPP-feature by subject raising to SpecT (17c) then applies within a cyclic domain, the TP, that is properly included within another cyclic domain, the CP, in violation of the SCC.

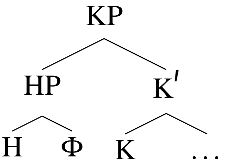
## 3. Non-monotonic Derivations

This section introduces the background that the strictly cyclic reanalysis is based on: the theory of non-monotonic derivations (Heck 2016, 2022).

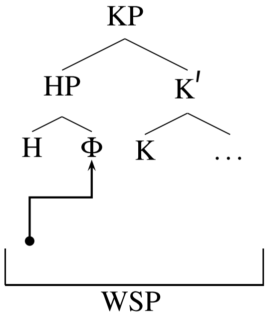
### 3.1. Necessity of Workspaces

Given the SCC and the existence of complex specifiers, it is clear that syntactic derivations must be able to construct different syntactic objects in parallel. The common assumption is that the derivation may employ different 'workspaces' (WSP), which serve to built up and hold ready various syntactic objects (Uriagereka 1999). To illustrate, suppose that the structure in (18) is to be generated. A (partial) derivation of (18), such as the one in (19), which makes use of only one WSP, violates the SCC: In order to generate the complex category HP Φ must be merged counter-cyclically.

(18)

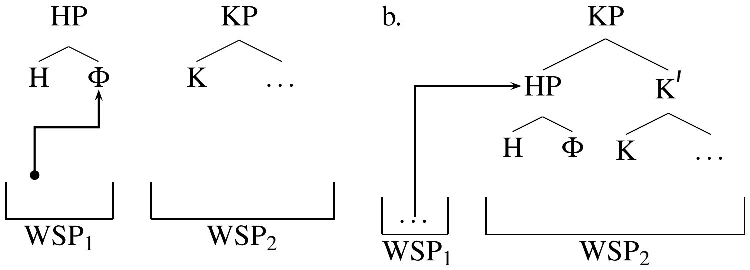


(19)



In contrast, the derivation in (20a,b), which makes use of multiple WSPs, is able to generate (18) without violating the SCC: The complex category HP is first generated in $WSP_1$; later HP is merged from $WSP_1$ to $WSP_2$, thereby becoming a specifier of KP.

(20) a.



### 3.2. Making Further Use of WSPs

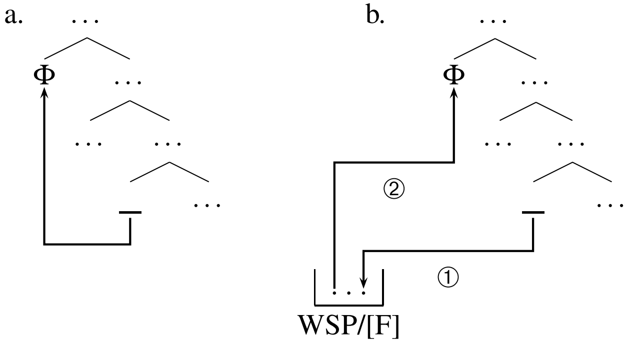
Following Heck (2016, 2022), I assume that syntactic derivations may make further use of multiple WSPs (see Nunes 2001, 2004; but cf. also Bianchi and Chesi 2014, Jayaseelan 2017, Adger 2017, Thoms 2019, and Thoms and Heycock 2022 for related proposals). In particular, movement of Φ in (21a) may be decomposed into two operations. First, removal of Φ applies (cf. Müller 2017, 2018, Pesetsky 2016), shifting Φ to another WSP, see step ① in (21b).[4] (For ease of exposition, the second WSP, which would host the tree

---

[4] This reminds of sideward movement (Nunes 2001, 2004, Hornstein 2001). These proposals mainly deal with phenomena where the category shifted to another WSP undergoes external Merge to pick up a second theta role (Control, parasitic gaps, across-the-board movement, but cf. Nunes 2004 on reconstruction and head-movement). The present discussion focuses on cases where internal Merge (implying a c-commanding movement trigger) is involved.

being moved from, is not displayed in (21).) Second, Φ is remerged from the WSP to the current tree (② in (21b)).

(21)



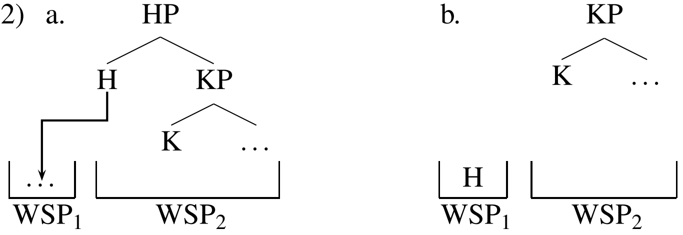
The probe [F] that attracts Φ acts as a pointer to the WSP that Φ is temporarily moved to (in (21b), this is indicated by displaying the probe below the WSP that hosts the attracted category).

If no other operation is interspersed between ① and ② in (21b), the derivation is equivalent to the one in (21a), where movement applies in one fell swoop. It becomes interesting when such interspersion takes place.

### 3.3. Shrinking Trees

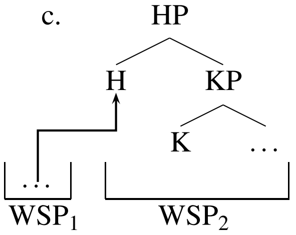
Head-movement, like phrasal movement, may proceed via some WSP (recall (21)). Assume some KP, immediately dominated by an HP. Suppose next that the head H is removed and placed in some $WSP_1$ (22a). Then, by assumption, the projection of H, HP, ceases to exist temporarily (cf. Heycock and Kroch 1993, Takano 2000). What remains in $WSP_2$ is thus KP (22b).

(22)



In other words, going from (22a) to (22b), the representation has shrunk. (Put yet another way: The representations of the derivation are not monotonously growing. This is where the notion of a non-monotonic derivation comes from.) Later, H may be remerged. This re-establishes the HP (22c).[5]

(22) c.



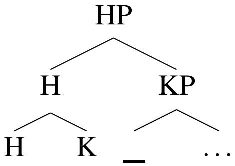
It should be mentioned that I assume that removal of a head H does not lead to the disappearance of HP under two conditions. First, if there is a specifier within HP, removing HP would leave this specifier unconnected to the rest of the representation. Second, if HP is the complement to another head, removing HP would make the higher head lose its connection to its complement. In these configurations, HP is maintained (see Heck 2016).

## 4. Strictly Cyclic Reformulation

### 4.1. Head-Movement/Undermerge

Bobaljik (1995) (see also Bobaljik and Brown 1996, Nunes 2004) proposes to render head-movement strictly cyclic by invoking sideward movement (Nunes 2001, 2004). To illustrate, assume that the representation to be generated is the one in (23), where head-movement of K to H has applied.

(23)



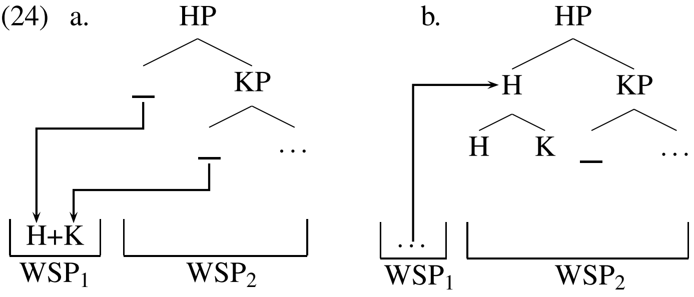
The derivation proceeds as follows. Before H is merged with KP, K is removed from KP and placed into a separate WSP that already hosts H (sideward movement). Being part of the same WSP as H, K may adjoin to H in a strictly cyclic fashion. Afterwards, the thus generated complex head is

---

[5] Head-movement as in (22a-c) seems pointless as H ends up in the same position it started from. The motivation for this maneuver will be given in section 4.

remerged. In Heck (2016), the proposal of Bobaljik (1995) is slightly modified by adding the assumption that the attracting head H is first merged to a position c-commanding the head K. In this configuration, H probes K, which triggers the subsequent removal to the two heads to the same WSP (24a). The remains of the derivation match the derivation proposed in Bobaljik (1995): First K adjoins to H (noted as H+K in (24a)), then the complex head formed in $WSP_1$ is remerged to the tree in $WSP_2$ (24b).

(24)



As a consequence, head-movement is strictly cyclic.

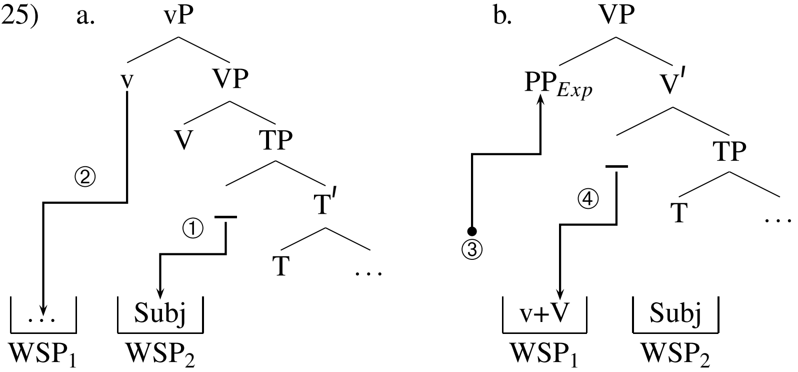
Given the structural parallelism between representations that are generated by head-movement and representations generated by Pesetsky's (2013) operation Undermerge (recall (4)), it appears that Undermerge may be treated in the same way as head-movement.

### 4.2. Minimality

The following analysis of apparent Minimality violations in the context of Subj-raising across an experiencer in English was proposed by Heck (2016). It pursues an idea already put forward in Stepanov (2001*a*,*b*) according to which such raising is possible because the experiencer is not yet part of the structure when raising applies. Rather, the experiencer is merged late, after raising. The crucial difference between Heck (2016) and Stepanov's (2001a,b) analysis is that late merger of the experiencer is strictly cyclic in the former but counter-cyclic in the latter. As will become clear, the analysis makes crucial use of the theory of head-movement presented in section 4.1.

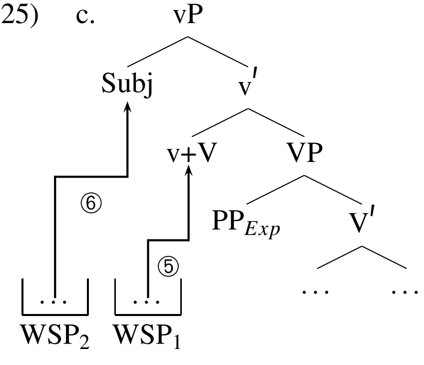
The derivation is given in (25a-c). Suppose that instead of merging the experiencer in SpecV right away, v is merged with VP. The Subj is attracted by v and is placed in the separate $WSP_2$, see step ① in (25a).[6,7] As there is no experiencer present, such raising respects Minimality. Next, v is displaced to $WSP_1$ in order to initiate V-to-v head-movement (step ②; see section 4.1). With v being removed, the vP-shell vanishes, too, and the tree shrinks, becoming a VP again. Accordingly, the experiencer can now be merged to SpecV, respecting the SCC (③ in (25b)). Also, V joins v in $WSP_1$, to form a complex head (step ④).

(25) a. b.



In the remaining steps, the complex v+V-head is remerged with VP from $WSP_1$, thereby re-establishing the vP of the matrix clause, and the (to-be-raised) Subj is merged from $WSP_2$ to Specv (steps ⑤ and ⑥).

(25) c.



---

[6] I am assuming here that A-movement must make an intermediate stop in Specv in order to comply with the Phase Impenetrability Condition (PIC Chomsky 2000, 2001); see also the references in section 2.3.

[7] The clausal spine in (25) is hosted by yet another WSP, not displayed here.

From there, the Subj may move to SpecT at some later point. Related analyses are available to other cases of apparent Minimality violations (Heck 2016; cf. also Thoms 2023 for a similar idea), as for instance those mentioned in section 2.2 above.

Note, however, that the above maneuver that allows the Subj to get past the experiencer crucially relies on two ingredients. First, the derivation must allow to procrastinate Merge of the experiencer. In languages that do not exhibit raising across an experiencer (see section 2.2), it may be the case that such procrastination is not allowed (see Heck 2016, 2022, and Privizentseva 2022 for some ideas what might regulate the availability of late merger). Second, head-movement must take place. Without V-to-v movement applying in (25a-c), the vP-projection would not vanish temporarily. And without the tree shrinking to the size of a VP, late merger of the experiencer to SpecV would violate the SCC. In other words, non-monotonic derivations of the kind illustrated in (25a-c) are contingent on particular conditions and are not simply available across the board. This means that there may be (apparent) Minimality violations that are not amenable to this technique.

### 4.3. Reconstruction

Before turning to the analysis, it is useful to clarify some background assumptions. In contrast to much contemporary research on reconstruction that is based on the copy theory of movement, I assume here that reconstruction (at least reconstruction with respect to binding) is the result of a derivational interpretation of binding principles (see Burzio 1986, Belletti and Rizzi 1988, Lebeaux 1988, 2009, Heycock 1995, Sabel 1995, 1998). This means that, for instance, Principle C is violated if an R-expression is c-commanded by a coreferential expression at any point of the derivation. Accordingly, I am also not adopting the copy theory of movement (see below). Moreover, I am assuming that semantic interpretation proceeds cyclically (Epstein et al. 1998), as determined, for instance, by phases (Chomsky 2000, 2001, 2008).

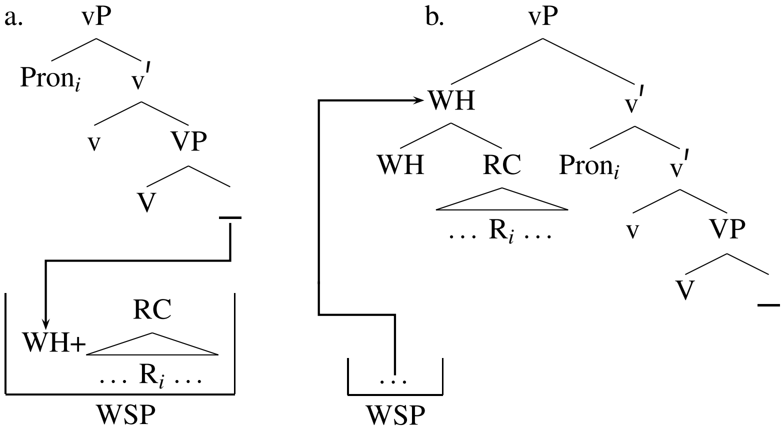
Returning to the main plot, consider first Lebeaux's (1988, 1990) observation that reconstruction of Ā-movement with respect to Principle C is not obligatory if the R-expression is embedded within a relative clause (26).

(26) Which argument [ that John$_i$ made __ ] did he$_i$ believe __?

The strictly cyclic analysis pursued here maintains the idea of late merger (Lebeaux 1988) but adds the assumption that late merger applies when the Ā-moved category is shifted to the WSP that also hosts the relative clause. Merge (or: adjunction) of the relative clause may then apply to the root and therefore respect the SCC. The proposal can already be found in Nunes (2004: 146-151). Here, it will be extended (and slightly adapted to present assumptions) to integrate the discussion in Takahashi and Hulsey (2009).

Relevant steps of the derivation are displayed in (27). In (27a), the *wh*-phrase moves to the WSP containing the relative clause. There, *wh*-phrase and relative clause combine. In the next step (27b), the constituent consisting of *wh*-phrase and relative clause is remerged to the main clause.[8]

(27)



As the relative clause (containing the R-expression) is not c-commanded by the co-indexed pronoun at any point, Principle C is not violated.

Turning to complement clauses, recall that in this case reconstruction with respect to Principle C is obligatory (28). Thus, any derivation of (28) employing late merger of the kind illustrated in (27) must be blocked.

(28) *Which argument [$_{\text{CP}}$ that John$_i$ is a genius ] did he$_i$ believe _?

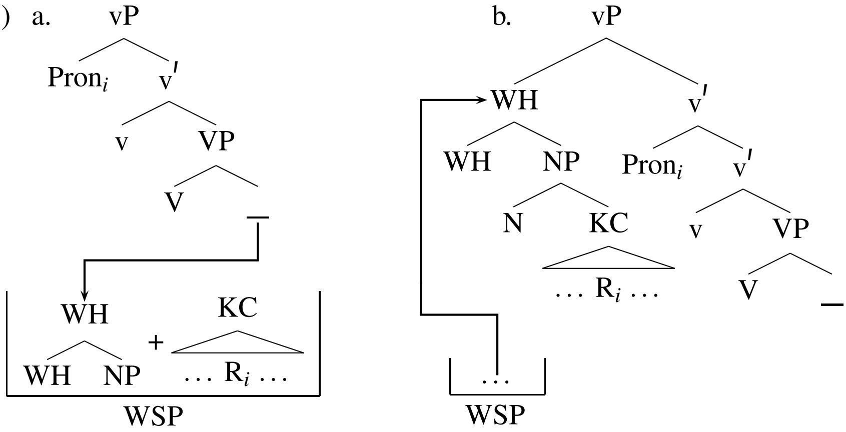
Assuming that complement clauses are merged as the sister of the noun while relative clauses are adjoined to the DP they modify (irrespective of whether

---

[8] Just like A-movement, *wh*-movement makes an intermediate stop in Specv, enforced by the PIC, cf. section 4.2.

one is dealing with a restrictive or an appositive relative clause),[9] it follows right away why a derivation fails that merges the *wh*-phrase (including the D-head and its NP-restrictor) in argument position and only later combines it with the complement clause (KC) in a separate WSP: Merge of the complement clause as the sister of N (see (29a)) violates the SCC as it does not target the root (cf. the representation in (29b)).[10]

(29)


a. b.

For the same reason, the proposal accounts for a reconstruction asymmetry in multiple modifier constructions noted by Tada (1993) (see also Sauerland 1998, Stanton 2016). Reconstruction of an internal modifier does not enforce reconstruction of an outer modifier: In (30a), the reduced relative clause *compatible with his* (inner modifier) reconstructs (for variable binding), and the full relative clause (outer modifier) does not reconstruct, thereby avoiding a Principle C violation. In contrast, reconstruction of the outer modifier *does* enforce reconstruction of an inner modifier, see (30b).

(30) a. [ Which computer compatible with his$_j$ that Mary$_i$ knew how to use ] did she$_i$ tell every boy$_j$ to buy _?

---

[9] For reasons of interpretation, it is often assumed that restrictive relative clauses are adjoined lower than appositive ones (Partee 1975). But cf. Frosch (1995) and Sternefeld (2006), who cast doubt on the necessity of this structural distinction; cf. Heim and Kratzer 1998: §4.5.

[10] In Takahashi and Hulsey (2009), the derivation in (29) is assumed to result in a non-interpretable LF, based on the copy theory of movement.

b. *[ Which computer compatible with Mary$_i$'s that he$_j$ knew how to use ] did she$_i$ tell every boy$_j$ to buy __?

The theoretical interpretation of this asymmetry under the present assumptions is clear. In multiple modifier constructions late merger may not target an internal modifier to the exclusion of an outer modifier because such late merger would be counter-cyclic (the internal modifier has to be merged in between the DP and the outer modifier). In contrast, nothing prevents late merger of the outer modifier in the presence of an internal modifier.

Returning to the case of complement clauses (recall (28)), a second derivation that has to be blocked involves external merge of a bare D-head to an argument position plus subsequent Ā-movement to a separate WSP, where D then undergoes late merger with an NP-restrictor containing a complement clause. I assume here that such a derivation violates the $\theta$-criterion of Chomsky (1981): A bare D cannot pick up the $\theta$-role that is assigned to the argument position; only a fully fledged argument DP is able to do so. When D finally merges with NP, it no longer occupies an argument position, and thus the resulting DP remains without a $\theta$-role.[11]

Finally, consider the case of A-movement, where reconstruction with respect to Principle C can be avoided, witness (31).

(31) Every argument [$_{\text{CP}}$ that John$_i$ is a genius ] seems to him$_i$ to be __ flawless.

As noted in section 2.3, Takahashi and Hulsey (2009) assume that (31) involves merge of a bare D, followed by movement of D to a position c-commanding the pronoun, and subsequent late merger of the restrictor NP containing the complement clause with D (see (13)). Under present assumptions, this analysis is not available for $\theta$-theoretic reasons (see above).

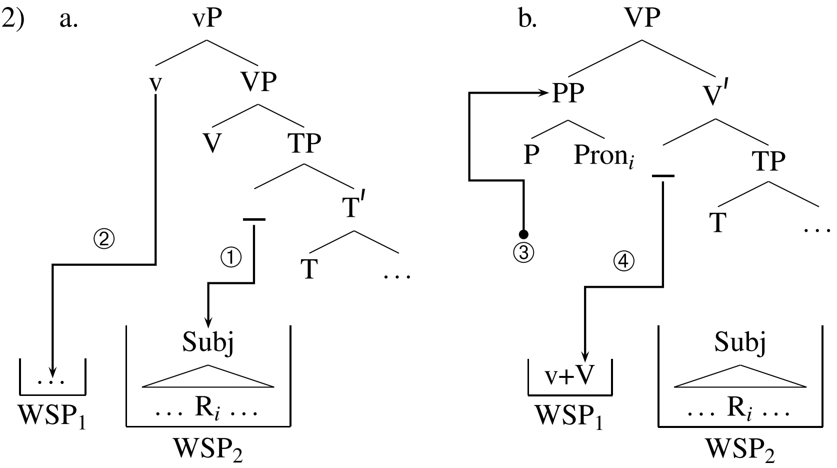
However, the lack of obligatory reconstruction for Principle C with A-movement already falls out from the analysis presented in section 4.2 (as

---

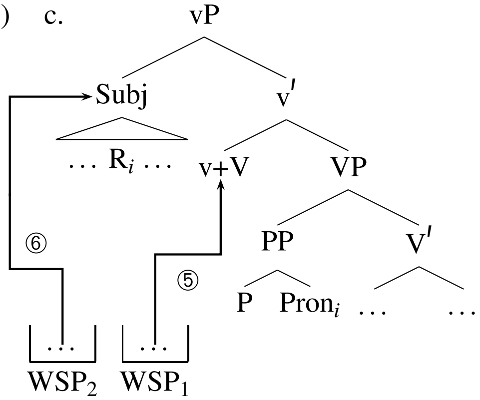
[11] Under the assumptions of Takahashi and Hulsey (2009) no such problem arises (cf. (13), section 2.3): There, the $\theta$-role is assigned to a copy of the moved element, which is semantically enriched by the process of trace conversion (due to Fox 1999) and therefore can receive the $\theta$-role. Accordingly, Takahashi and Hulsey (2009) invoke another constraint to block this derivation, see section 2.3. Nunes (2004), discussing a slightly different but related derivation, makes problems with copy deletion responsible for its failure.

already noted in Heck 2016). The offending co-indexed pronoun in (31) is contained within an experiencer PP. A-movement of the Subj to SpecT of the matrix clause crosses this experiencer, in apparent violation of Minimality. According to the analysis in section 4.2, this Minimality violation is avoided because the experiencer is merged late, after A-movement (to a separate WSP) has taken place. As a consequence, there is no point in the derivation where the pronominal experiencer c-commands the R-expression contained in the A-moved Subj. Therefore, no Principle C violation arises. This is illustrated in (32a-c).

(32)



(32)



In this way, the lack of Principle C effects with A-movement in English is related to the independent (and somewhat exceptional) property of A-movement across an experiencer in the same language.[12]

This makes the prediction that in languages that, generally, do not allow for raising across an experiencer Principle C effects should return in scenarios where such raising becomes exceptionally possible (but does not require a non-monotonic derivation of the kind in (32)). As pointed out in Heck (2016), the prediction is testable for such languages if raising applies across a cliticized experiencer (where cliticization helps to void Minimality). In such a scenario, there is a point of the derivation where the experiencer c-commands the to-be-raised Subj. This should trigger a Principle C effect if the experiencer is a pronoun co-indexed with a referential Subj (i.e., a reflexive). The prediction appears to be borne out. (33a) illustrates for French (McGinnis 1998). Similar facts hold for Italian (Rizzi 1986).

(33) a. *Jean$_i$ se$_i$ semble __ __ avoir du talent.
Jean SELF.DAT seems to.have of.the talent
'Jean seems to himself to be gifted.'

b. Jean$_i$ lui$_j$ semble __ __ avoir du talent.
Jean him.DAT seems to.have of.the talent
'Jean seems to him to be gifted.'

If the clitic is a non-coreferential pronoun, then the result is well-formed (33b), as expected. See Heck (2022) for further discussion of reconstruction effects in terms of non-monotonic derivations.

### 4.4. Tucking-In

One way to rephrase tucking-in in a way that obeys the SCC makes use of a buffer that is organized as a stack or a queue. Different versions of this proposal have been put forward (without being fully aware of previous works), see Doggett (2004) (who mentions the idea in a footnote, attributing it to David Pesetsky), Stroik (2009), Unger (2010), and Heck and Himmelreich (2017). The main point relevant here is that such a buffer can be straightfor-

---

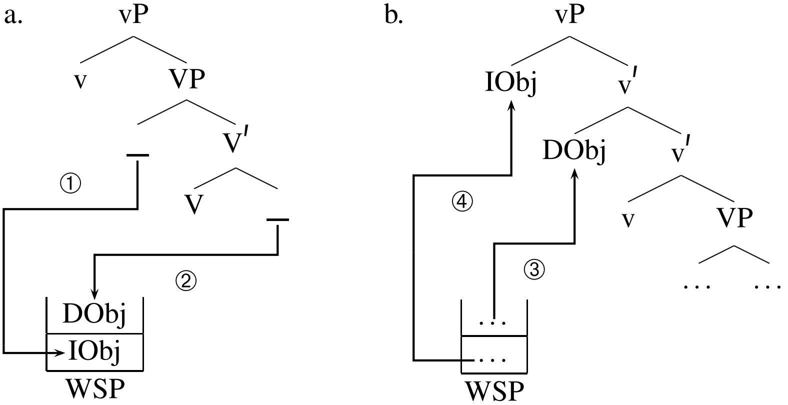
[12] A remaining problem for the approach is the fact that A-movement in English may reconstruct for Principle A and variable binding; see Heck (2016) for some discussion.

wardly described as a WSP pointed to by a single probe attracting multiple categories (Heck 2016).

The idea is as follows. The probe may only attract the highest category (respecting Minimality). Each attracted category is stored on top of the same stack in a separate WSP. If some category has been attracted to the stack, the next higher one becomes accessible to the probe. Once all categories have been attracted, the topmost category is removed from the stack and is remerged as the innermost specifier. Such removal makes the second topmost category of the stack accessible, which is then remerged as the next higher specifier (following strict cyclicity). This procedure continues until the stack is empty. In this way, the attracted categories show up as specifiers in an order (bottom up) that is the inverse of the order of attraction, leading to crossing paths.

A sample derivation is illustrated in (34a,b).[13]

(34)



(34a) shows that attraction obeys Minimality. First the closer IObj is attracted. Once removed and placed on the stack, the probe gets access to the DObj, attracts it and places it on top of the IObj on the stack (steps ① and ②). All objects have been attracted, and thus the remerge procedure starts, beginning with the DObj, which occupies the top of the stack. After the DObj is remerged (as the innermost specifier), the IObj is accessible and becomes the outermost specifier, see steps ③ and ④ in (34b), which obey the SCC.

---

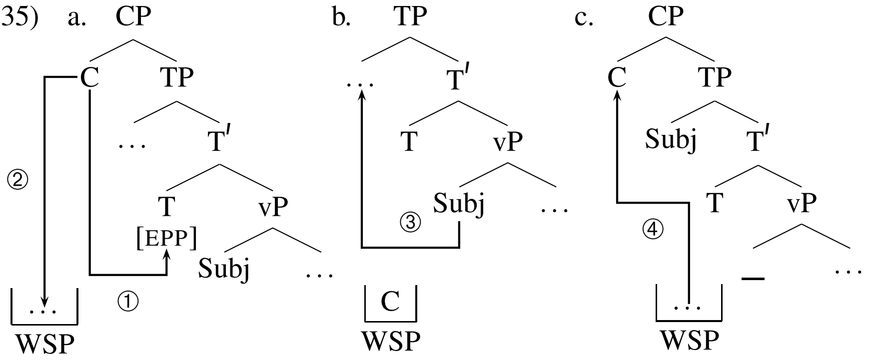
[13] (34) could, for instance, instantiate multiple object shift or multiple successive cyclic *wh*-movement of two objects.

### 4.5. Feature Inheritance

Finally, feature inheritance may receive a strictly cyclic interpretation if one makes the additional assumption that it shares the property of head-movement to temporarily remove the higher head (placing it in a separate WSP) after having handed down its features (Heck 2016).[14]

(35a-c) illustrates the effect of removing the higher head by means of a non-monotonic derivation:

(35)



Step ① in (35a) represents feature inheritance. Once C has assigned its EPP to T, it is removed to the WSP (step ②). With C removed, its projection vanishes, too (cf. section 4.1). The current tree shrinks, temporarily becoming a TP again. Accordingly, T can now satisfy the EPP it inherited from C by attracting the subject without violating the SCC (see step ③ in (35b)). Finally, C is remerged from the WSP, restoring the CP-layer (step ④ in (35c)).

## 5. Conclusion

To briefly conclude, in the present study I argued that non-monotonic derivations that make use of additional WSPs may be fruitfully put to use when approaching the problem of (apparent) counter-cyclic operations in syntax (including head-movement/Undermerge, Minimality, reconstruction, tucking-

---

[14] Given that it is not an operation involving the higher head itself that violates the SCC but rather an operation that involves the inheriting head (the SCC-violation being caused indirectly by the presence of the higher head), such an assumption is perhaps less motivated than it was for the case of head-movement.

in, and feature inheritance) in a uniform manner. Whether there are instances of apparent counter-cyclicity that are beyond the scope of this approach remains to be seen.

## References

Adger, David (2017): A Memory Architecture for Merge. Ms., Queen Mary University of London.

Anagnostopoulou, Elena (2003): *The Syntax of Ditransitives: Evidence from Clitics*. Studies in Generative Grammar, Mouton De Gruyter, Berlin.

Baker, Mark (1988): *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.

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

Bianchi, Valentina and Christiano Chesi (2014): 'Subject islands, reconstruction, and the flow of the computation', *Linguistic Inquiry* **45**, 525–569.

Bobaljik, Jonathan (1995): Morphosyntax: The Syntax of Verbal Inflection. PhD thesis, MIT, Cambridge, Massachusetts.

Bobaljik, Jonathan and Dianne Jonas (1996): 'Subject Positions and the Roles of TP', *Linguistic Inquiry* **27**, 195–236.

Bobaljik, Jonathan and Samuel Brown (1996): 'Interarboreal Operations: Head Movement and the Extension Requirement', *Linguistic Inquiry* **28**, 345–365.

Boeckx, Cedric (2003): *Islands and Chains*. John Benjamins, Amsterdam.

Bošković, Željko (2016): 'On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking-in from labeling', *The Linguistic Review* **33**, 17–66.

Bošković, Željko and Howard Lasnik (1999): 'How Strict Is the Cycle?', *Linguistic Inquiry* **30**, 691–703.

Branigan, Phil (1992): Subjects and Complementizers. PhD thesis, MIT, Cambridge, Massachusetts.

Branigan, Phil (2014): Approaching the Edge in Syntactic Derivation. Ms., Memorial University.

Broekhuis, Hans (2016): Feature inheritance versus extended projections. *In:* G. Legendre, M. Putnam, H. de Swart and E. Zaroukian, eds, *Optimality-Theoretic Syntax, Semantics, and Pragmatics*. Oxford University Press, Oxford, pp. 136–157.

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

Burzio, Luigi (1986): *Italian Syntax: A Government-Binding Approach.* Reidel, Dordrecht.

Campana, M (1992): A Movement Theory of Ergativity. PhD thesis, McGill University.

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

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

Chomsky, Noam (1993): A Minimalist Program for Linguistic Theory. *In:* K. Hale and S. J. Keyser, eds, *The View from Building 20*. MIT Press, Cambridge, Massachusetts, pp. 1–52.

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

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* R. Martin, D. Michaels and J. Uriagereka, eds, *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, Massachusetts, pp. 89–155.

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

Chomsky, Noam (2007): Approaching UG from Below. *In:* U. Sauerland and H.-M. Gärtner, eds, *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. Mouton de Gruyter, Berlin, pp. 1–30.

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, C. P. Otero and M. L. Zubizarreta, eds, *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*. MIT Press, Cambridge, Massachusetts, pp. 133–166.

Deal, Amy Rose (2009): 'The Origin and Content of Expletives: evidence from "Selection"', *Syntax* **12**, 285–323.

Dikken, Marcel den (2007): 'Phase Extension. Contours of a Theory of the Role of Head Movement in Phrasal Extraction', *Theoretical Linguistics* **33**, 1–41.

Doggett, Teal Bissell (2004): All Things Being Unequal: Locality in Movement. PhD thesis, MIT.

Emonds, Joseph (1976): *A Transformational Approach to English Syntax*. Academic Press, New York.

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

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

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

Fox, Danny and Jon Nissenbaum (1999): Extraposition and scope: A case for overt

QR. *In:* S. Bird, A. Carnie, J. D. Haugen and P. Norquest, eds, *Proceedings of WCCFL 18*. Cascadilla Press, Somerville, Massachusetts, pp. 132–144.

Freidin, Robert (1978): ‘Cyclicity and the Theory of Grammar’, *Linguistic Inquiry* **9**, 519–549.

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

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

Frosch, Helmut (1995): ‘Zwischen Semantik und Pragmatik: Appositive und restriktive Relativsätze’, *Sprachtheorie und germanistische Linguistik* **2**, 7–19.

Heck, Fabian (2016): Non-monotonic derivations. Habilitation thesis, Leipzig University.

Heck, Fabian (2018): Successive-Cyclic Crossing Paths and Tucking-In in English. Presentation at CGSW 33, Universität Göttingen.

Heck, Fabian (2022): Ways to sidestep Minimality (and how to diagnose them). *In:* G. Bîlbîie, B. Crysmann and G. Schaden, eds, *Empirical Issues in Syntax and Semantics 14*. CSSP, Paris, pp. 85–112.

Heck, Fabian and Anke Himmelreich (2017): ‘Opaque Intervention’, *Linguistic Inquiry* **48**, 47–97.

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

Heycock, Caroline (1995): ‘Asymmetries in Reconstruction’, *Linguistic Inquiry* **26**, 547–570.

Heycock, Caroline and Anthony Kroch (1993): ‘Verb Movement and the Status of Subjects: Implications for the Theory of Licensing’, *Groninger Arbeiten zur Germanistischen Linguistik* **36**, 75–102.

Hiraiwa, Ken (2001): EPP: Object Shift and Stylistic Fronting in Scandinavian. *In:* K. Megerdoomian and L. A. Bar-el, eds, *Proceedings of WCCFL 20*. Cascadilla Press, Sommerville, Massachusetts, pp. 290–303.

Holmberg, Anders (1999): ‘Remarks on Holmberg’s Generalization’, *Studia Linguistica* **53**, 1–39.

Holmberg, Anders and Christer Platzack (1995): *The Role of Inflection in Scandinavian Syntax*. Oxford University Press, Oxford.

Hornstein, Norbert (2001): *Move! A Minimalist Theory of Construal*. Blackwell, Oxford.

Jayaseelan, Karattuparambil (2017): Parallel Work Spaces in Syntax and the Inexistence of Internal Merge. *In:* G. Sengupta, S. Sircar, M. G. Raman and R. Balusu, eds, *Perspectives on the Architecture and Acquisition of Syntax*. Springer, Singapore, pp. 115–134.

Kayne, Richard (1975): *French Syntax: The Transformational Cycle*. MIT Press, Cambridge, Massachusetts.

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

Koizumi, Masatoshi (1993): Object Agreement Phrases and the Split VP Hypothesis. *In:* J. Bobaljik and C. Phillips, eds, *Papers on Case and Agreement I*. Vol. 18 of *MIT Working Papers in Linguistics*, MIT Press, Cambridge, Massachusetts, pp. 99–148.

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

Lebeaux, David (1990): Relative Clauses, Licensing, and the Nature of the Derivation. *In:* S. Rothstein, ed., *Perspectives on Phrase Structure: Heads and Licensing*. Academic Press, San Diego California.

Lebeaux, David (2009): *Where does Binding Theory Apply?*. MIT Press, Cambridge, Massachusetts.

Lechner, Winfried (2019): A calculus for reconstruction and anti-reconstruction. *In:* M. Schenner and M. Krifka, eds, *Reconstruction Effects in Relative Clauses*. Akademie Verlag, Berlin, pp. 113–145.

Legate, Julie Anne (2003): 'Some Interface Properties of the Phase', *Linguistic Inquiry* **34**, 506–516.

McCawley, James D. (1968): Lexical Insertion in a transformational grammar without Deep Structure. *In:* B. J. Darden, C.-J. N. Bailey and A. Davison, eds, *Proceedings from the fourth Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, pp. 71–80.

McCawley, James D. (1970): 'English as a VSO language', *Language* **46**, 286–299.

McGinnis, Martha (1998): Locality in A-Movement. PhD thesis, MIT, Cambridge, Massachusetts.

Mulders, Iris (1997): Mirrored Specifiers. *In:* J. A. Coerts and H. de Hoop, eds, *Linguistics in the Netherlands 1997*. John Benjamins, Amsterdam, pp. 135–146.

Müller, Gereon (1997): Order Preservation, Parallel Movement, and the Emergence of the Unmarked. *In:* F. J. d'Avis and U. Lutz, eds, *Zur Satzstruktur im Deutschen*. Universität Tübingen, pp. 171–214. Arbeitspapiere des Sonderforschungsbereichs 340.

Müller, Gereon (2017): 'Structure Removal. An Argument or Feature-Driven Merge', *Glossa* **2**, 1–35.

Müller, Gereon (2018): 'Structure Removal in Complex Prefields', *Natural Language and Linguistic Theory* **36**, 219–264.

Murasugi, Kumiko (1992): Crossing and Nested Paths: NP Movement in Accusative and Ergative Languages. PhD thesis, MIT, Cambridge, Massachusetts.

Nunes, Jairo (2001): 'Sideward Movement', *Linguistic Inquiry* **32**, 303–344.

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

Partee, Barbara (1975): 'Montague Grammar and Transformational Grammar', *Linguistic Inquiry* **6**, 203–300.

Pesetsky, David (1995): *Zero Syntax. Experiencers and Cascades*. MIT Press, Cambridge, Massachusetts.

Pesetsky, David (2013): *Russian Case Morphology and the Syntactic Categories*. MIT-Press, Cambridge, Massachusetts.

Pesetsky, David (2016): Exfoliation: Towards a derivational theory of clause size. Talk given at Workshop on Shrinking Trees, Leipzig University.

Pollock, Jean-Yves (1989): 'Verb Movement, Universal Grammar, and the Structure of IP', *Linguistic Inquiry* **20**, 365–424.

Privizentseva, Mariia (2022): Inverse case attraction in Moksha Mordvin relative clauses and the nature of Merge. PhD thesis, Leipzig University.

Řezáč, Milan (2002): Cyclic domains and (multiple) specifier constructions. Ms., IKER, Bayonne.

Richards, Marc (2004): Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax. PhD thesis, University of Cambridge.

Richards, Marc (2007): 'On Feature-Inheritance: An Argument from the Phase Impenetrability Condition', *Linguistic Inquiry* **38**, 563–572.

Richards, Norvin (1997): What moves where when in which language. PhD thesis, MIT, Cambridge, Massachusetts.

Richards, Norvin (1999): Featural Cyclicity and the Ordering of Multiple Specifiers. *In:* S. D. Epstein and N. Hornstein, eds, *Working Minimalism*. MIT Press, Cambridge, Massachusetts, pp. 127–158.

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

Riemsdijk, Henk van and Edwin Williams (1986): *Introduction to the Theory of Grammar*. MIT Press, Cambridge, Massachusetts.

Rizzi, Luigi (1986): On Chain Formation. *In:* H. Borer, ed., *The Syntax of Pronominal Clitics*. Vol. 19 of *Syntax and Semantics*, Academic Press, New York, pp. 65–96.

Rizzi, Luigi (1990): *Relativized Minimality*. MIT Press, Cambridge, Massachusetts.

Roberts, Ian (2010): *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. MIT Press, Cambridge, Massachusetts.

Rudin, Catherine (1988): 'On Multiple Questions and Multiple Fronting', *Natural Language and Linguistic Theory* **6**, 445–501.

Sabel, Joachim (1995): Restrukturierung und Lokalität. PhD thesis, Frankfurt University.

Sabel, Joachim (1998): Principles and Parameters of Wh-Movement. Habilitation, Frankfurt University.

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

Sauerland, Uli (1998): The Meaning of Chains. PhD thesis, MIT, Cambridge, Massachusetts.

Sauerland, Uli (2003): 'Intermediate Adjunction with A-Movement', *Linguistic Inquiry* **34**, 308–314.

Sportiche, Dominique (2005): Division of labor between Merge and Move: Strict locality of selection and apparent reconstruction paradoxes. Ms., UCLA.

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

Stanton, Juliet (2016): 'Wholesale Late Merger in Ā-Movement: Evidence from Preposition Stranding', *Linguistic Inquiry* **47**, 89–126.

Stepanov, Arthur (2001*a*): Cyclic Domains in Syntactic Theory. PhD thesis, University of Connecticut, Storrs.

Stepanov, Arthur (2001*b*): 'Late adjunction and minimalist phrase structure', *Syntax* **4**, 94–125.

Stepanov, Arthur (2004): Ergativity, Case and the Minimal Link Condition. *In:* A. Stepanov, G. Fanselow and R. Vogel, eds, *Minimality Effects in Syntax*. Mouton de Gruyter, Berlin, pp. 367–399.

Sternefeld, Wolfgang (2006): *Syntax – Eine morphologisch motivierte generative Beschreibung des Deutschen*. Stauffenburg, Tübingen.

Stroik, Thomas (2009): *Locality in Minimalist Syntax*. MIT Press, Cambridge, Massachusetts.

Tada, Hiroaki (1993): A/A$'$ Partition in Derivation. PhD thesis, MIT.

Takahashi, Shoichi (2006): Decomposition and identity. PhD thesis, MIT.

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

Takano, Yuji (2000): 'Illicit Remnant Movement: An Argument for Feature-Driven Movement', *Linguistic Inquiry* **31**, 141–156.

Thoms, Gary (2019): Antireconstruction as layering. *In:* M. Baird and J. Pesetsky, eds, *Proceedings of NELS 49*. GLSA, Amherst.

Thoms, Gary (2023): Getting out of the way in Icelandic. Talk given at GLOW 46, Vienna.

Thoms, Gary and Caroline Heycock (2022): Deriving 'late merge' from External Remerge. *In:* O. Bakay, B. Pratley, E. Neu and P. Deal, eds, *Proceedings of NELS 52*. GLSA, Amherst.

Unger, Christina (2010): A computational approach to the syntax of displacement and the semantics of scope. PhD thesis, Universiteit Utrecht.

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

Vikner, Sten (1989): ‘Object Shift and Double Objects in Danish’, *Comparative Germanic Syntax* **44**, 119–141.

Yuan, Michelle (2017): More on Undermerge: Phrasal and head movement interactions in Kikuyu. *In:* C. Halpert, H. Kotek and C. van Urk, eds, *A Pesky Set: Papers for David Pesetsky*. MIT Press, Cambridge, Massachusetts, pp. 543–552.

# Late Merge and Cyclicity

Mariia Privizentseva

## Abstract

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

## 1. Background

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

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

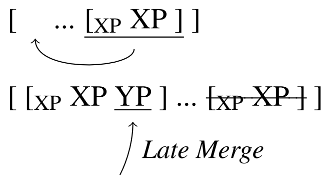
*Cyclicity*, 329–350
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

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

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

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

(3) a.
b.



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

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

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

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

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

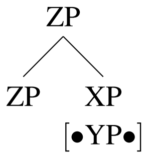
When considering their relation to late Merge, I will assume an approach to syntax under which all instances of Merge, including late Merge, are driven by features. Following Heck and Müller (2007), I will indicate features that trigger Merge as [•F•] and features that trigger Agree as [$*$F$*$]. I will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

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

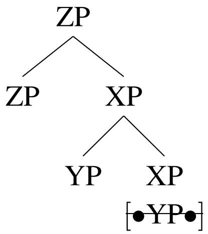
## 2. Late Merge and the Earliness Principle / Featural Cyclicity

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

(8) Step 1: Merge XP and ZP



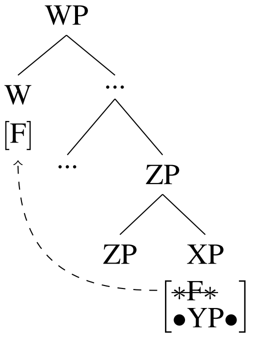
(9) Step 2: Late Merge of YP



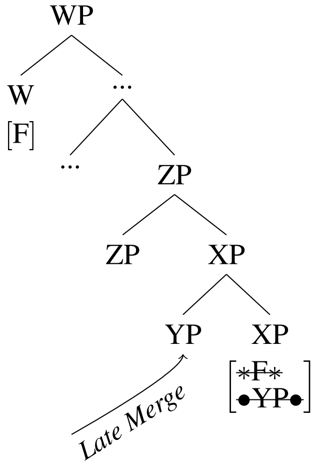
The state of affairs is different if unsatisfied syntactic features are assumed to be ordered (see Stabler (1997) and Müller (2011)), and if Merge and Agree features can be interleaved. In that case, only one feature appears on the top of the stack and can be active. Features ordered after it must wait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

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

(10) Step 1: Upward Agree



(11) Step 2: Late Merge of YP



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

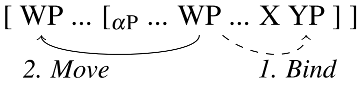
Late Merge as in (11), however, is peculiar and differs from most proposed cases of late Merge in that a phrase within which late Merge applies remains deeply embedded. Being developed for deriving anti-connectivity effects of moved phrases, late Merge typically occurs in configurations where a constituent,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

(12) a. [ [$_{\text{DP}}$ DP $\underline{\text{PP}}$ ]$_j$ ... ~~[$_{\text{DP}}$ DP ]$_j$~~ ]
↑ *Late Merge*

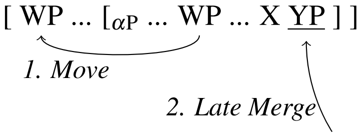
b. [$_{\text{DP}}$ Which pictures [$_{\text{PP}}$ near John$_i$ ] ]$_j$ did he$_i$ look at [$_{\text{DP}}$ ~~which pictures~~ ]$_j$?

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

(13) WP binds YP before movement:



(14) Late Merge counterfeeds binding:



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

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

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

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

(16) $[_{ZP}$ Z $[_{XP}$ X $[_{WP}$ WP YP ] ] ]
④ ② ① ③

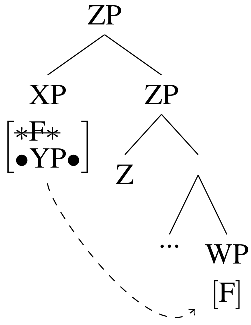
Epstein et al. (1998) use derivational c-command to account for data as in (17), where personal pronoun *she* can be co-referent with *Mary* and thus should not c-command it at any stage of the derivation, while *every student* must c-command *he* to allow for bound variable interpretation.

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

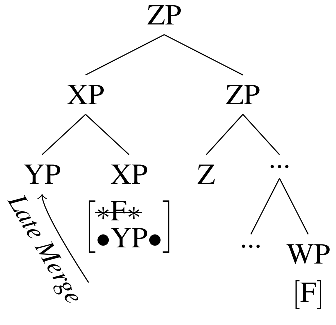
Given derivational c-command all c-command relations are as required if the relative clause is late merged after the pronoun *she* enters the derivation but before the noun phrase *every student* is introduced. This application of late Merge, thus, does not involve movement of the host and is, in this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

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

(18) Step 1: Downward Agree



(19) Step 2: Late Merge of YP

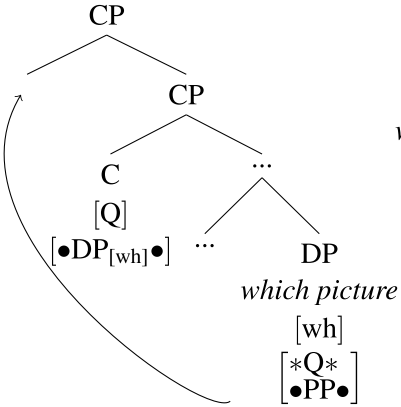


The XP targeted by late Merge in this derivation may be either first merged as a specifier of ZP or moved to this position. In the first case, the presence or absence of the late Merge of YP has no further effect. In the second case, the derivation is the one that is typically proposed to account for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

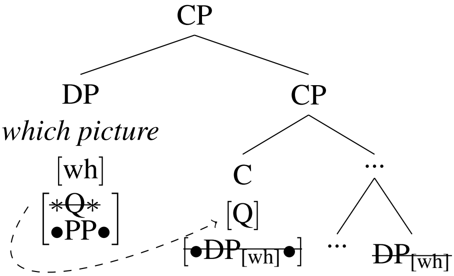
(20) Which pictures near John$_i$ did he$_i$ look at __ ?

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

(21) Step 1: Wh-movement

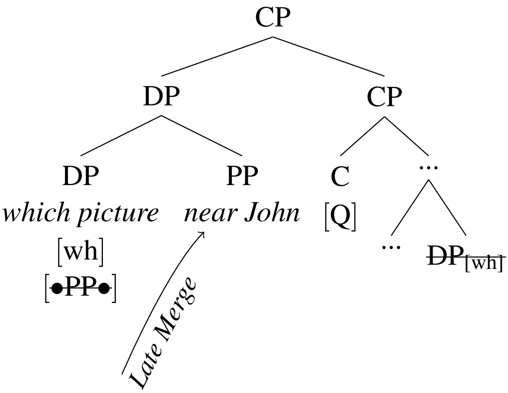


(22) Step 2: Agreement



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

(23) Step 3: Late Merge of the adjunct

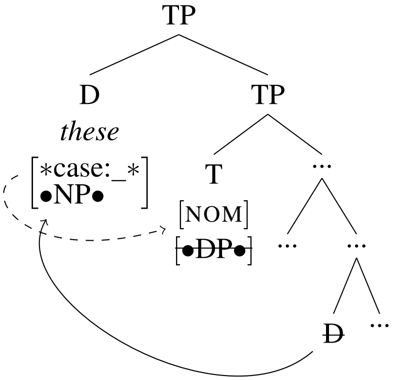


I will next turn to anti-connectivity attested for arguments of A-moved phrases as in (24).

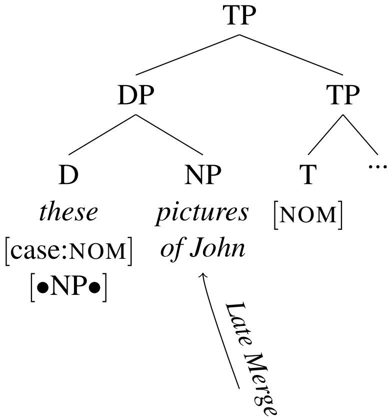
(24) [ These pictures of John$_i$ ]$_j$ seemed to him$_i$ [ _$_j$ to be very good].

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

(25) Steps 1-2: Move and Agree



(26) Step 3: Late Merge

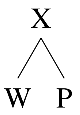


Interestingly, the distribution of late Merge under the EP/FC plus downward agree is similar to the distribution of late Merge in a model where the application of Merge is restricted by yet another cyclicity principle: the Peak Novelty Condition (PNC). The PNC was introduced by Safir (2019), and unlike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

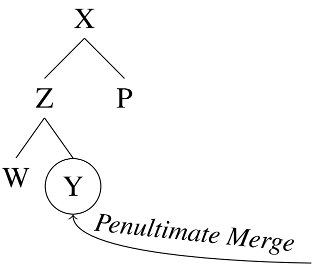
(27) Peak Novelty Condition
After every instance of Merge, $M_i$, the undominated node U of the resulting structure immediately dominates a node that U did not immediately dominate before $M_i$. (see Safir (2019: 292))

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

(28) Before Merge



(29) Penultimate Merge



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

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

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

---

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

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

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

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

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

---

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

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

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

## 3. Late Merge and Strict Cycle Condition / Extension Condition

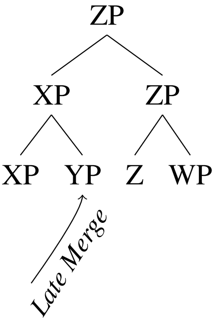
In this section, I will turn to the two stronger notions of cyclicity: the SCC and the EC. The definitions I will rely on here are repeated below.

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

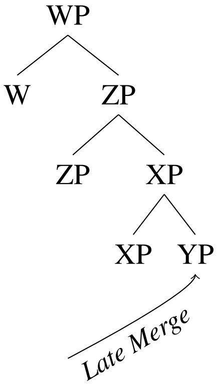
(32) Extension Condition (EC):
A syntactic derivation can only be continued by applying operations to the root of the tree. (see Chomsky (1993, 1995) and Adger (2003: 75) for this formulation)

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

(33) Late Merge of YP

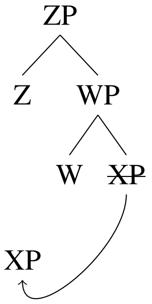


(34) Late Merge of YP

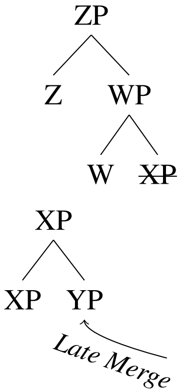


Nevertheless, syntax restricted by the SCC/EC can, in fact, incorporate late Merge if additional assumptions are made. As noted in the previous section, since late Merge is used to derive anti-connectivity effects, it applies inside constituents that undergo movement. This opens up the possibility to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

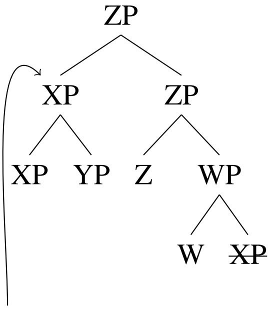
(35) Step 1



(36) Step 2



(37) Step 3



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

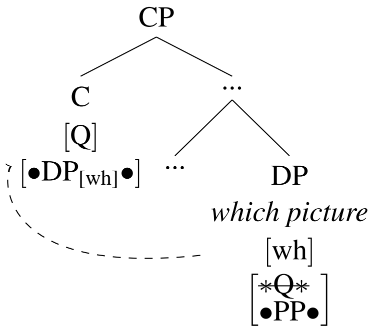
Let us see how this applies to the core cases of anti-connectivity discussed above. Again, I will start with the adjuncts of Ā-moved wh-phrases in (38) repeated from (2b).

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

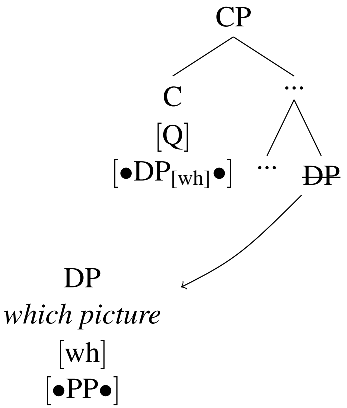
Similarly to the previous section, I assume that wh-dependency involves [•DP$_{[wh]}$•] on the C head and [∗Q∗] on the wh-phrase, but in this derivation the agreement probe must be discharged via upward agreement before movement as shown in (39). After deletion of the agreement probe, the Merge feature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

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

(39) Step 1: Agree

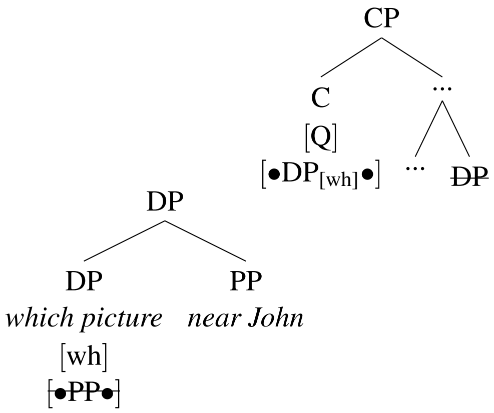


(40) Step 2: Merge



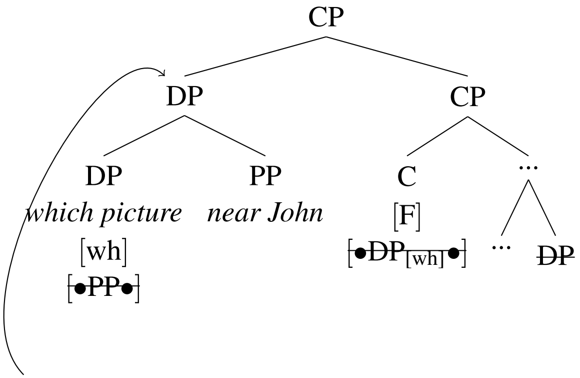
After this, the active Merge feature can be discharged without violating the SCC/EC (see (41)).

(41) Step 3: Late Merge



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

(42) Step 4: Merge into Spec,CP

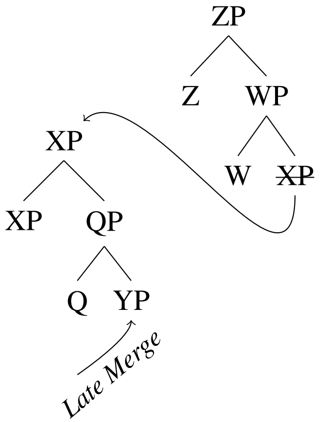


Anti-connectivity for arguments of A-moved phrases, as in (43), is derived in the same vein with the only differences being that the whole NP is late merged, and it is a case probe that can ensure a delayed discharge of the merge feature.

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

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

(44) Impossible late Merge



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

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

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

## 4. Summary

Late Merge is often used for deriving anti-connectivity effects, but is at the same time extensively criticized as being inherently counter-cyclic. This paper shows that it can, in fact, be implemented in a way fully compatible with all major concepts of cyclicity. In each case, however, additional assumptions are necessary. To incorporate late Merge into models regulated by the EP/FC, it is required to assume that features on syntactic nodes are organized in ordered stacks. Models restricted by the SCC/EC require to assume that movement involves copying of a syntactic object to the workspace. At the same time, despite the general possibility to implement late Merge, none of the considered models by themselves predict late Merge in exactly those contexts where it is needed to derive anti-connectivity. Coupling with widely assumed restrictions on late Merge related to case and adjunct/complement status also does not automatically yield a correct distribution.

## References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

# Incorporation as a Repair Mechanism and Cyclicity

Helene Streffer*

## Abstract

Turkana (Eastern Nilotic; Kenya) shows a pattern where non-nominal modifiers incorporate into the head noun prenominally while they appear unbound postnominally. In this paper, I develop a two-step analysis with (i.) regular phrasal movement to SpecDP followed by (ii.) incorporation of the non-nominal modifier from the specifier into the D head. The analysis exploits the limits of cyclicity and constitutes, as such, a good testing ground for fine-grained notions of cyclicity varying in their degree of strictness. After presenting arguments for the analysis and sketching an implementation in Harmonic Serialism (McCarthy 2008, Heck and Müller 2013), I compare the analysis to three different formulations of the Strict Cycle Condition (SCC, Chomsky 1973): the Extension Condition (Chomsky 1995), the Peak Novelty Condition (Safir 2019) and a formulation of the SCC in Müller (2018). Incorporation as an immediate repair mechanism of previously built structure constitutes, thereby, an argument for a less strict version of the SCC which maintains at the same time a strong notion of cyclicity.

## 1. Introduction

One way to distinguish between different concepts of cyclicity is to investigate the strictness of cyclicity concepts like the Strict Cycle Condition (SCC, Chomsky 1973). Various formulations varying in their degree of strictness

---

*I am grateful to Jannet Akwom, Geoffrey Edapal Edato, and Amos Nakwa Emoru for their valuable work as linguistic consultants. If not otherwise indicated, data in this paper come from fieldwork with them. Turkana exhibits grammatical tone which distinguishes tense on verbs and case on nouns (Dimmendaal 1983: 37). However, since in-depth research on tone in Turkana is still missing, I will leave tone aside for the moment. Accordingly, I have chosen not to transcribe tone in order to avoid misrepresenting the data.

I would also like to thank Maria Kouneli, Paula Fenger, Sören E. Tebay, Imke Driemel, Johannes Hein, Coppe van Urk, Bronwyn M. Bjorkman, the extended group of the project 'Layers of Morphosyntactic Number in Eastern Sudanic', and the audiences of the Leipzig Morphlogy/Syntax Colloquium, the Strict Cycle Workshop, the retreat of the Research Unit on Cyclic Optimization and the Syntax Colloquium at Goethe-Universität Frankfurt for feedback on various stages of this work.

*Cyclicity*, 351–370
Mariia Privizentseva, Felicitas Andermann & Gereon Müller (eds.)
LINGUISTISCHE ARBEITS BERICHTE 95, Universität Leipzig 2023

have been proposed for the SCC in the literature (e.g. Chomsky 1995, Safir 2019, Müller 2018). In order to distinguish between these different concepts, we need to examine the degree to which grammar exploits the limits of cyclicity. In this paper, I will discuss a word order pattern in the Turkana DP (Eastern Nilotic; Kenya) where non-nominal modifiers incorporate into the head noun prenominally while they appear unbound postnominally. I lay out a two-step analysis that derives the pattern through (i.) regular phrasal movement of the modifier to SpecDP, followed by (ii.) an optimization step where non-nominal elements can incorporate into the noun in order to adhere to noun-initiality. The second step of incorporation exploits the limits of cyclicity. Hence, if this analysis is on the right track, it represents a good testing ground in order to investigate the limits of cyclicity, i.e. the strictness of cyclicity concepts like the SCC. While the present analysis violates the strictest versions of this condition, it fits into Müller's (2018) formulation of the SCC that allows operations to target everything within the current phrase.

The rest of the paper is structured as follows: In section 2.1, I describe the puzzle in the Turkana data, then provide a more detailed picture of modifiers in the prenominal domain in 2.2. Subsequently, in section 3, I show that modifiers which appear prefixed to the noun are actually incorporated before phonology. The analysis is laid out in section 4 and is followed by a discussion of cyclicity, where I compare the analysis to three formulations of the SCC which vary in their degree of strictness.

## 2. Data

### 2.1. The Puzzle

The head noun in the Turkana DP (Eastern Nilotic; Kenya) precedes all modifiers in the unmarked case. Thus, the DP in Turkana is generally noun-initial. An example that illustrates this strong preference in Turkana is shown in (1).

| | | | | |
|---|---|---|---|---|
| (1) | ŋa-kinekine | ŋa-tʃɛ | ŋa-uni | ŋa-kɛŋ |
| | F.PL-goat | F.PL-other | F.PL-three | F.PL-3SG.POSS |
| | 'his three other goats' | | | |

The language exhibits three different genders, which are marked on the noun with a prefixed gender marker (see (2)). This becomes relevant if one considers the quantifier *-tʃɛ* ‘other’. It is possible to move the quantifier in front of the noun due to information structure reasons. However, while it appears unbound in the postnominal position (3a), it appears prenominally between the nominal gender marker and the noun itself (3b). Thus, it appears as a bound prefix of the noun. Note that this happens in a context where a non-nominal element is moved to a prenominal position which goes against the general preference for noun-initiality in the Turkana DP.

(2) a. e-kile
M.SG-man

b. a-bɛrʊ
F.SG-woman

c. ɪ-ŋɔq
N.SG-dog

(3) a. a-bɛrʊ a-tʃɛ
F.SG-woman F.SG-other
‘another woman’

b. a-tʃɛ-bɛrʊ
F.SG-other-woman

As with all other modifiers, nominal possessors appear postnominally in the unmarked case. However, as can be seen in (4b), it is also possible to move the nominal possessor in front of the head noun for information structure reasons. In contrast to the quantifier *-tʃɛ*, which prenominally appears between the gender marker and the noun, the nominal possessor appears unbound in front of the head noun.[1] Unlike the moved quantifier, the prenominal appearance of the nominal possessor is still in line with the general preference for noun-initial DPs.

(4) a. ŋɪ-dɛ a e-tuko
PL-child of M.SG-zebra
‘children of a zebra’

---

[1] While I take the example in (4b) to show that the nominal possessor can be moved in front of the head noun, the example is not conclusive. One could also interpret the example as a case of possessor raising where *e-tuko* ‘zebra’ does not form a constituent with *ŋɪ-dɛ* ‘children’. I thank Mariia Privizentseva and anonymous reviewers of GLOW 46 and ACAL 54 for pointing this out. One way to test the constituency would be to try to move *e-tuko ŋɪ-dɛ* to the preverbal domain. Barabas-Weil (2022) notes that the preverbal domain in Turkana can only host a single constituent. Thus, if the nominal possessor does not form a constituent with the head noun, one would expect this movement test to be ungrammatical. Unfortunately, this has to await future research. Note, however, that the general puzzle - why non-nominal modifiers appear incorporated into the head noun prenominally while they appear unbound postnominally - is generally independent of the data point concerning nominal possessors. Additionally, the analysis in 4 would still work even if this data point turns out to have a different interpretation.

b. tɔ-dɛm-ara-ɪ e-tuko ŋɪ-dɛ
3.SUBS-take-ITIVE-ASP M.SG-zebra PL-child
'The children were taken away from the zebra'

Hence, on the one hand, there are modifiers which can appear unbound in front of the noun (like a nominal possessor), and on the other hand, there are modifiers which prenominally appear as a bound prefix (like the quantifier *-tʃɛ*).

## 2.2. Modifiers in the Prenominal Domain

This section provides an overview of various modifiers when they are moved for information structure reasons to the prenominal domain. Section 2.1 already showed that one can distinguish between two different prenominal positions: an unbound prenominal position vs. an incorporated prenominal position between the nominal gender marker and the noun. We saw that nominal possessors appear in the unbound prenominal position while the quantifier *-tʃɛ* appears in the incorporated prenominal position. The corresponding example with *-tʃɛ* is repeated in (5), with the addition of (5c), which shows that the quantifier *-tʃɛ* cannot surface in the unbound prenominal position.

(5) a. a-bɛrʊ a-tʃɛ
F.SG-woman F.SG-other
'another woman'
b. a-tʃɛ-bɛrʊ
F.SG-other-woman
c. *a-tʃɛ a-bɛrʊ
F.SG-other F.SG-woman

In contrast to the simple quantifier *-tʃɛ*, a more complex quantifier like *-kidikidio* in (6) shows the opposite pattern. This quantifier cannot appear in the incorporated position between the nominal gender marker and the noun, but it can surface in the unbound prenominal position due to information structure reasons even though this violates the general noun-initiality preference.

(6) a. ŋa-kipi ŋa-kidikidio
F.PL-water F.PL-few
'small amount of water'

b. *ŋa-kidikidio-kipi
F.PL-few-water
c. ŋa-kidikidio ŋa-kipi
F.PL-few F.PL-water

The same behavior can be found with numerals. As shown in (7), a numeral can only surface prenominally in the unbound position.

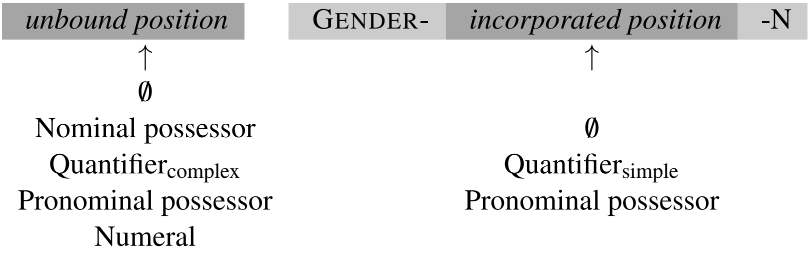
(7) a. ŋa-bɛr ŋa-kan-k-omwɔn
F.PL-woman F.PL-five-LINK-four
'nine women'
b. *ŋa-kan-k-omwɔn-bɛr
F.PL-five-LINK-four-woman
c. ŋa-kan-k-omwɔn ŋa-bɛr
F.PL-five-LINK-four F.PL-woman

Finally, one can observe that pronominal possessors surface in both prenominal positions. The example in (8b) shows that a pronominal possessor can appear between the nominal gender marker and the noun, and the example in (8c) demonstrates that it can also surface in the unbound prenominal position.

(8) a. ŋa-ki ŋa-kon
F.PL-ear F.PL-2SG.POSS
'your ears'
b. ŋa-kon-ki
F.PL-2SG.POSS-ear
c. ŋa-kon ŋa-ki
F.PL-2SG.POSS F.PL-ear

An overview of the pattern in the prenominal domain can be found in (9). Since Turkana exhibits a strong preference for noun-initiality in the DP, no modifier appears in the prenominal domain in the unmarked case (indicated with ∅ in (9)). If a modifier moves to the prenominal domain for information structure reasons, one can observe two different positions: (i) an incorporated position between the nominal gender marker and the noun where simple quantifiers and pronominal possessors can appear and (ii) an unbound prenominal position where nominal possessors, complex quantifiers, pronominal possessors and numerals surface prenominally. While most modifiers are restricted to one prenominal position, pronominal possessors surface in both prenominal positions. Finally, it can be noted that the appearance of complex quantifiers, pronominal possessors, and numerals in the unbound prenominal position is, at least at first sight, a violation to the general noun-initiality in the Turkana DP.

(9) *The prenominal domain*



## 3. Incorporation before Phonology

In this section, I will present three empirical arguments that the appearance of a prenominal modifier between the nominal gender marker and the noun (like *-tʃɛ* in (5)) is the result of incorporation which applies before phonology.

The first argument makes use of a specific type of nominal concord called ‘restrictive agreement’ in Dimmendaal (1983: 217), which distinguishes the form of the nominal gender marker from the agreement marker prefixed to modifiers. In previous examples involving non-restrictive agreement, the nominal gender marker and the agreement marker prefixed to modifiers were identical in form. In order to argue for incorporation, one has to show that the gender marker prefixed to the simple quantifier (like *-tʃɛ in* (5), repeated in (10c)) is, in fact, the nominal gender marker and not a modifier with regular agreement marking in front of a noun without a nominal gender marker. The example in (10a) shows an instance of restrictive agreement with a postnominal modifier. While the nominal gender marker is *a-*, the restricitve agreement marker prefixed to the quantifier is *na-*. Crucially, it is impossible to retain the restrictive agreement marking of the quantifier if the quantifier appears prenominally (see (10b)). Thus, the gender marking in (10c) is, in fact, the nominal gender marker.

(10) a. a-bɛrʊ **na**-tʃɛ
F.SG-woman F.RESTR-other
'another woman'

b. ***na**-tʃɛ-bɛrʊ
F.RESTR-other-woman
'another woman'

c. **a**-tʃɛ-bɛrʊ
F.SG-other-woman
'another woman'

Furthermore, taking into account why the quantifier *-tʃɛ* does not show its own gender agreement when it appears in the incorporated prenominal position provides an answer to the timing of incorporation. In principle, there could be different explanations for this pattern. However, they all predict that the moved quantifier forms a complex head with the noun before phonology. One potential explanation could be that at the point where an agreement node would be inserted, the quantifier has already incorporated into the noun and cannot get its agreement node anymore (cf. the argument for morphological wordhood of Bulgarian denominal adjectives in Harizanov 2018). Another explanation could be that an agreement node of the moved quantifier next to the node hosting the nominal gender marker with nearly identical features induces a haplological dissimilation rule sensitive to morphosyntactic features, which deletes the agreement node. The relevant domain for such a process has been argued to be a complex head (see Nevins 2012), which shows that the quantifier incorporates into the noun before phonology.

In addition, one can observe that the size of the modifiers plays a role for the prenominal position where the element appears. While the complex quantifier cannot appear in the incorporated position, the simple quantifier can. Thus, complex elements cannot appear in the incorporated prenominal position. The same size-based requirement for the incorporated position can be found with a modified quantifier. The simple quantifier *-di* can occur in the incorporated prenominal position (see (11b)). As can be seen in the unmarked postnominal word order in (11c), the quantifier *-di* can be modified by *tʃɪtʃɪk* 'somewhat'. However, the whole modified phrase cannot appear in the incorporated prenominal position (see (11d)). This pattern is straightforwardly explained if the incorporation step is a result of head movement which can only target single heads and not more complex material.

Since this operation needs to take place before phonology, it is again an argument that incorporation takes place before phonology.

(11) a. ŋa-kile ŋa-di
F.PL-milk F.PL-some
'some milk'
b. ŋa-di-kile
F.PL-some-milk
c. ŋa-kile ŋa-di tʃɪtʃɪk
F.PL-milk F.PL-some somewhat
'some small amount of milk'
d. *ŋa-di-tʃɪtʃɪk-kile
F.PL-some-somewhat-milk

The previous paragraphs presented arguments for incorporation taking place before phonology. This predicts that phonological processes treat the incorporated element as already part of the noun because incorporation happened earlier. This prediction can be confirmed by looking at vowel harmony in Turkana. The language exhibits [ATR]-vowel harmony which is generally root-controlled if there is no strong suffix (Dimmendaal 1983: 19-27). The example in (12b) shows that there is no vowel harmony between the incorporated quantifier and the noun.[2] However, Dimmendaal (1983: 192) notes that compounds do not exhibit vowel harmony. Thus, it is expected to see no vowel harmony with incorporation, either. Instead, the incorporated element interrupts the vowel harmony between the nominal gender marker and the noun. The nominal gender marker in (12a) is in the harmony domain of the [+ATR] noun if there is no incorporated element and surfaces accordingly with the [+ATR] vowel *e*-. However, as soon as the quantifier is incorporated and opens a new vowel harmony domain (in (12b)), the nominal gender marker is not in the harmony domain of the noun anymore. Accordingly, the nominal gender marker harmonizes with the incorporated [-ATR] quantifier and surfaces with the [-ATR] vowel *ɛ*- in this case. This demonstrates that the quantifier has already incorporated when it comes to the phonological process of vowel harmony.

---

[2] My data differ here from Dimmendaal (1983: 303-304, 343-344), who noted vowel harmony between the incorporated element and the noun.

(12) a. **e**-kile ɛ-tʃɛ ye
M.SG-man M.SG-other that
'that other man'
b. **ɛ**-tʃɛ-kile ye
M.SG-other-man that

## 4. Analysis and Discussion

An analysis for the Turkana DP data has to capture three main points: (i) the two different positions in the prenominal domain, (ii) the *unbound* postnominal appearance vs. *bound* prenominal appearance of the modifiers surfacing in the incorporated prenominal position, and (iii) the correct split between the modifiers which appear in the two different prenominal positions, including the twofold behavior of pronominal possessors who can appear in both positions.

The data discussion of the complex quantifier *-kidikidio* and the modified quantifier *-di tʃɪtʃɪk* suggests that the size of the prenominal modifiers constitutes a crucial factor for the division between the different prenominal positions. As indicated earlier, it is straightforward to derive this distinction through phrasal movement vs. head movement. Thus, at first sight, one could think about an analysis where some modifiers move via phrasal movement to SpecDP, the unbound prenominal position, while other modifiers move via head movement to D, the bound prenominal position. However, that approach has two major drawbacks. First, long head movement from the base position of the modifier, which is a specifier of a functional projection in the nominal spine, violates locality constraints proposed for head movement (Travis 1984, Koopman 1984).[3] Second, it would require an ad hoc stipulation to explain

---

[3] Van Urk (2015) describes a phenomenon in the clausal domain in Dinka that shows strong similarities to the pattern in the Turkana DP. Dinka exhibits a V2-effect where the finite verb moves to C, and the clause-initial position in front of the finite verb can be occupied by the argument which serves as a topic or focus of the clause. He reports that this clause-initial position is restricted to nominals. If a PP adjunct moves to this position, it is only the embedded nominal which surfaces in the clause-initial position, and the preposition of the adjunct incorporates into the finite verb in C. Van Urk (2015) proposes an analysis for this pattern where the preposition undergoes a long head movement step followed by phrasal movement of the embedded noun into the clause-initial position. While this movement step violates the HMC (Travis 1984, Koopman 1984), it additionally imposes a look-ahead problem. The preposition incorporates before any conflicting structure, i.e. a non-nominal in the clause-initial position, exists. I propose that changing the order of operations in van Urk's (2015) analysis solves the

why some modifiers move to the prenominal domain via phrasal movement while others move probably for the same information structure feature via head movement.

In contrast, I pursue an analysis where all modifiers undergo regular phrasal movement in a first step induced by the same information structure feature. This is then followed by an optimization step where the derivation tries to adhere to the general noun-initiality preference. I propose that incorporation is a possible repair mechanism available for small elements. Thus, if a simple quantifier moves to SpecDP, the structure will be repaired via incorporation of the quantifier into D.

### 4.1. A Harmonic Serialist Approach

I assume that the nominal gender marker in Turkana is located on D (see e.g. Dimmendaal (1983: 307) for similarity between nominal gender markers and demonstratives in Turkana) and that noun-initiality is derived through N-to-D movement (see, e.g., Carstens 2017 for N-to-D movement in Shona and Kouneli 2020 for Kipsigis). I take from Minimalism the assumption that phrases which consist only of one head are both maximal and minimal at the same time. Locality restrictions on head movement (Koopman 1984, Travis 1984) predict that these elements can only be addressed as minimal, i.e. as a head, from a local viewpoint. Any attempt to address a phrase which consists of only one head as minimal, i.e. as a head, from a distant point in the tree will be hindered by locality constraints on head movement. Thus, from a distant viewpoint, they will always be perceived as maximal, i.e. as a phrase.

I implement the analysis in Harmonic Serialism (McCarthy 2008, Heck and Müller 2013), a strictly derivational OT-model. A crucial property of this model is that every evaluation step includes maximally one operation. Thus, output candidates can only vary from the input by applying at most one operation to the input structure. The output candidate with the best constraint profile is chosen as the input for the next evaluation step. The derivation stops when optimization is no longer possible, i.e. when the best output candidate is

---

locality and the look-ahead problem. The gist of the analysis would then be the same as the analysis proposed in section 4.1 for the Turkana DP. This is a welcome result since it seems to be the case that we are dealing with the same underlying phenomenon in Turkana and Dinka: a structure where a dispreferred element appearing in a restricted initial position can be repaired by incorporation of that element into the next lower head.

identical to the input (Heck and Müller 2013). The present Harmonic Serialist analysis makes use of two constraints: NOMINALFIRST$_{\text{DP}}$ in (13) which reflects the strong preference in the language for noun-initial DPs and the MERGE CONDITION in (14) which drives feature-based merge and movement.

(13) NOMINALFIRST$_{\text{DP}}$ (NF):
(van Urk 2015, Driemel and Kouneli 2022)[4]
Assign a violation for every non-nominal element in SpecDP.

(14) MERGE CONDITION (MC):
(Chomsky 1995, 2001, Heck and Müller 2013)
Assign a violation for every unchecked [•F•].

The following tableaux in (15) and (16) illustrate the derivation of a simple quantifier like *-tʃɛ*. Subsequently, I will lay out how the analysis captures the rest of the modifiers. The tableau in (15) starts at the point of the derivation where the Ā-feature on D induces phrasal movement of the modifier to SpecDP.[5,6] Note that movement of the quantifier to SpecDP will violate the NOMINALFIRST$_{\text{DP}}$ constraint. However, movement is still carried out because a violation of the higher ranked MERGE CONDITION due to not moving and leaving the Ā feature unchecked would be worse (see $O_1$ in (15)).

---

[4] A restriction on nominals in the initial position has also been observed for the clause-initial position in van Urk (2015) for Dinka (Western Nilotic; South Sudan) and in Driemel and Kouneli (2022) for Kipsigis (Southern Nilotic; Kenya). In addition, Barabas-Weil (2022) seems to observe the same restriction for the preverbal position in Turkana (Eastern Nilotic; Kenya). The account presented in this paper shows that the same restriction can be found in the nominal domain in Turkana. Thus, the strong preference for nominals in the initial position of the clausal or nominal domain could be a general property of Nilotic languages.

[5] The derivation in (15) starts at a point in the derivation where N-to-D movement has happened already. I follow Carstens's (2017) proposal for N-to-D movement here. Following the HMC, this includes the noun raising through every intermediate head to D. Thus, on its way to D it will necessarily pick up the heads of the phrases that build the nominal spine. Accordingly, the noun is part of a complex head structure when it arrives at D. However, for illustratory reasons, I will still represent it as N in the following structures.

[6] I am using a generic Ā-feature here since there is not enough semantic work on Turkana to determine which information structure feature triggers the movement in the Turkana nominal domain.

(15) *Regular phrasal movement*

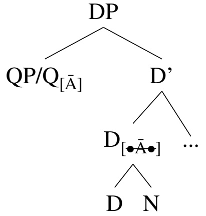
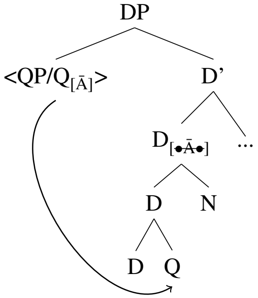
| I: $[_{DP}$ D$_{[\bullet\bar{A}\bullet]}$ N [ ... $[_{XP}$ QP/Q$_{[\bar{A}]}$ $[_{X'}$ X [ ...NP ]]]]] | MC | NF |
|---|---|---|
| O$_1$: $[_{DP}$ D$_{[\bullet\bar{A}\bullet]}$ N [ ... $[_{XP}$ QP/Q$_{[\bar{A}]}$ $[_{X'}$ X [ ...NP ]]]]] | *! | |
| ☞ O$_2$: $[_{DP}$ QP/Q$_{[\bar{A}]}$ $[_{D'}$ D$_{[\bullet\bar{A}\bullet]}$ N [ ... $[_{XP}$ <QP/Q> $[_{X'}$ X [ ...NP ]]]]]] | | * |

An integral part of Harmonic Serialism is that the derivation will only stop when no further optimization is possible. Therefore, it is naturally the case in this model that the derivation tries to optimize the constraint profile and repair the structure with a non-nominal quantifier in the initial position of the DP. I propose that Turkana exhibits incorporation into D as a repair mechanism. Thus, if the non-nominal element in SpecDP is small enough to undergo incorporation into D, i.e. a single head, the structure can be optimized further. Incorporation results in an optimized constraint profile because the non-nominal element is not in SpecDP anymore, which resolves the previous violation of NOMINALFIRST$_{DP}$. Since a simple quantifier like *-tʃɛ* is at the same time maximal and minimal, i.e. consists of only a single head, it will be able to undergo incorporation into D and optimize the DP structure. The optimization step with a simple quantifier is shown in (16). Crucially, incorporation only becomes possible at that point and not earlier since the quantifier and D are only now in a local relationship with one another.[7]

---

[7] The general incorporation step from a specifier into the head is well-known from Matushansky's (2006) M-Merger approach.

(16) *An optimizing step: Incorporation as a repair*

| I: [$_{\text{DP}}$ QP/Q$_{[\bar{\text{A}}]}$ [$_{\text{D}'}$ D$_{[\cancel{\bullet\bar{\text{A}}\bullet}]}$ N [ ... [$_{\text{XP}}$ <QP/Q> [$_{\text{X}'}$ X [ ...NP ]]]]]] | MC | NF |
|---|---|---|
| O$_{21}$:  | | *! |
| ☞ O$_{22}$:  | | |

The first step of phrasal movement is the same for all modifiers that are marked with an Ā-feature. It is the second step of optimization where differences between the modifiers arise. Since the incorporation step is restricted to heads, the repair mechanism is not available to phrases consisting of more than a single head. Thus, the constraint profile cannot become better for complex non-nominal elements after Ā-movement to SpecDP, and they will surface in the unbound prenominal position. This is the case for the complex quantifier *-kidikidio* and numerals. In Turkana, numerals starting from six are built through an associative construction, which indicates that there is more structure involved than a single head. Similarly, the form of the complex quantifier *-kidikidio* seems to be much more complex than its simple counterpart *-di*.[8] Furthermore, Dimmendaal (1983: 168) notes that *-kidikidio* can appear in verbal constructions. I take this to mean that this quantifier is more complex than others. In contrast to complex non-nominal modifiers, movement of the nominal possessor will never violate NOMINALFIRST$_{\text{DP}}$ since it fulfills the nominal requirement. Therefore, incorporation will never be needed as a repair mechanism for nominal possessors, and they surface unbound in SpecDP.

---

[8] Heine (1981) reports that both quantifiers have the same meaning.

Finally, the question arises: what enables pronominal possessors to appear in both the unbound and the incorporated prenominal position? For an answer to this question, it is interesting to take a look at the clausal domain in Turkana. Barabas-Weil (2022) shows that Turkana, which has VSO word order in the unmarked case, exhibits a preverbal focus position. For this clause-initial position, she observes a distinction between weak and strong pronouns. Both pronouns appear in the postverbal position, but only the strong pronoun *ájɔ́ŋ* can appear in the preverbal position (see (17)).

(17) a. é-múdʒí (àjɔ́ŋ/àŋ) ákíríŋ
1SG-eat I.NOM meat.ABS
'I am eating meat' (Barabas-Weil 2022)

b. ájɔ́ŋ/*áŋ é-múdʒí ákíríŋ
I.ABS 1SG-eat meat.ABS
'I am eating meat' (Barabas-Weil 2022)

I assume that strong and weak pronouns correspond to structures with different levels of complexity.[9] More precisely, I assume that strong pronouns exhibit a complex structure while weak pronouns consist of a single head. If the strong vs. weak pronoun distinction is also maintained with pronominal possessors (even though they cannot be distinguished morphologically here) it is straightforward that pronominal possessors can appear in both the unbound and the incorporated prenominal position. Since strong pronouns consist of more than a single head, they cannot optimize their constraint profile by incorporation and surface unbound in SpecDP. In contrast, weak pronouns can undergo the repair mechanism and optimize their constraint profile because they are at the same time maximal and minimal. Hence, we can observe pronominal possessors in both prenominal positions.[10]

To sum up, the two-step analysis straightforwardly captures all three main points of the data summarized at the beginning of section 4. First, it derives both prenominal positions while maintaining the same Ā-movement trigger for all elements. Second, this analysis provides an explanation for why

---

[9] Depending on the concrete implementation, structures of different complexity with pronominal possessors could also correlate with being nominal vs. non-nominal (see e.g. Déchaine and Wiltschko 2002).

[10] This would mean that the nominal domain and the clausal domain in Turkana show a similar phenomenon with respect to pronominal possessors, except that there exists a repair mechanism in the nominal domain which allows weak pronouns to surface in the incorporated position.

modifiers like simple quantifiers appear bound in the prenominal position and unbound in the postnominal position. There is simply no reason to undergo an incorporation repair in their postnominal position because a postnominal modifier does not violate NOMINALFIRST$_{\text{DP}}$. And third, the correct divide between the elements appearing in the unbound or the incorporated position (summarized in (9)) follows from the characterization of the repair mechanism only being available for single heads.

### 4.2. Cyclicity

The proposed analysis behaves in a derivational fashion. Structure is built by checking features, and if this produces a dispreferred structure, the derivation tries to repair it. Crucially, the repair mechanism in this analysis applies in the immediate next step after the dispreferred structure has been built. Thus, the analysis has a cyclic characteristic principally. However, it is worth taking a closer look at the structure in order to detect the more fine-grained differences. An influential implementation of cyclicity in derivations is the Strict Cycle Condition from Chomsky (1973) in (18). Interestingly, the degree of strictness is not fixed in this formulation but depends on the notion of a 'cyclic node'. In the following, I will consider three different formulations of the SCC that vary in their degree of strictness.

(18) *Strict Cycle Condition (SCC)*
No rule can apply to a domain dominated by a cyclic node *A* in such a way as to affect solely a proper subdomain of *A* dominated by a node *B* which is also a cyclic node. (Chomsky 1973: 243)

The strictest notion of the SCC is formulated in Chomsky's (1995) Extension Condition in (19), which states that Merge and movement have to apply at the root. The first step in the analysis presented above, regular phrasal movement to SpecDP, obeys this condition. However, the second step of incorporation in cases where the repair mechanism applies does not target the root node. Accordingly, the incorporation step violates the strictest version of the SCC.[11]

---

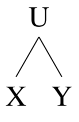
[11] Note that without additional assumptions, head movement taking place in the syntax is generally excluded under the definition of the EC. Thus, the N-to-D movement earlier in the derivation would already be problematic under the strictest version of the SCC.

(19) *Extension Condition (EC)* Chomsky (1995)
Merge and movement have to extend the structure at the root.

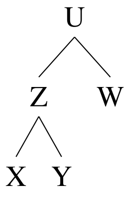
A slightly weaker version of the SCC is formulated in Safir's (2019) Peak Novelty Condition in (20). This version of the SCC permits Merge to apply at more places than only the root. Safir (2019: 292-293) explains the difference using the trees in (21) and (22). Both trees show an instance of Merge ($M_i$) where X has just been merged. The structure in (21) would be in line with both the EC and the PNC. The structure in (22) (with X being the element which has been merged last) violates the EC. However, (22) fulfills the PNC because Z is assumed to be a new node that the undominated node U immediately dominates after the application of Merge. Safir (2019) notes that this makes operations like head movement possible. Thus, a structure like the input in (15), where N has moved to D (abbreviated in (23)), is permitted by the PNC. The subsequent step of regular phrasal movement of the modifier to SpecDP is again in line with the PNC. However, the second step of incorporation is once again problematic under the definition in (20). Moving a modifier to SpecDP extends the structure so that Merge cannot target the D head again since the resulting new node would not be immediately dominated by the undominated node after the specifier position has been filled. Thus, the incorporation step does not obey the PNC either.[12]

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

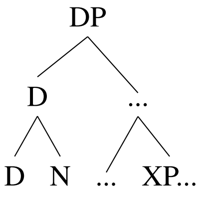
(21)



(22)



(23)



---

[12] In addition, the modifier targets an even lower projection of D in the complex D head during its repair step than N-to-D movement targeted. Depending on whether the definition in (20) counts nodes or labels, this could also rule out the repair step independently of the first step in the analysis with regular phrasal movement.

However, the incorporation step fits into Müller's (2018) formulation of the SCC in (24). This formulation allows the derivation to target every node within the current XP. Hence, targeting D for incorporation after movement of the modifier to SpecDP is in accordance with Müller's (2018) formulation of the SCC. Note that this formulation constitutes the least strict version of the SCC out of the three discussed versions in this section. Nevertheless, it maintains a notion of strict cyclicity. For the present analysis, this means that spec-head incorporation can only repair a structure created in the step immediately before it.

(24) *Strict Cycle Condition in Müller (2018)*
Within the current XP $\alpha$, a syntactic operation may not exclusively target some item $\delta$ in the domain of another XP $\beta$ if $\beta$ is in the domain of $\alpha$. (Müller 2018: 241)

## 5. Conclusion

To sum up, I have presented data from the Turkana DP that show two different prenominal positions: an unbound position in front of the noun and a position between the nominal gender marker and the noun itself. There are various arguments making reference to vowel harmony, nominal concord, and the size of the modifiers in both positions that point towards an analysis of the bound position as incorporation into the noun before phonology. The presented two-step analysis derives the pattern through (i.) regular phrasal Ā-movement to SpecDP followed by (ii.) an optimization step where a structure that violates the general noun-initiality preference in the DP can be repaired by incorporation of the non-nominal element into D if the non-nominal element is maximal and minimal at the same time.

If this analysis is on the right track, it is an interesting test case for the evaluation of different versions of the SCC. A comparison of three different degrees of strictness (EC, PNC, and SCC in Müller 2018) shows, on the one hand, that the incorporation repair mechanism does not fit into the strictest versions of the SCC. On the other hand, however, it also shows that the formulation in Müller (2018) can capture the described repair mechanism by restricting a cycle to the current domain. This demonstrates that analyses exhausting the limits of cyclicity, given that they can be argued to be accurate, can provide the space to evaluate different fine-grained notions of cyclicity.

## References

Barabas-Weil, Leonie (2022): Long Distance wh-Movement in Turkana. Handout of a talk presented at the Syntax Colloquium in Frankfurt.

Carstens, Vicki (2017): Noun-to-Determiner-Movement. In: *The Wiley Blackwell Companion to Syntax*. 2nd edn, pp. 1–26.

Chomsky, Noam (1973): Conditions on Transformations. In: *A Festschrift for Morris Halle*. Academic Press, New York, p. 232–286.

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

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

Déchaine, Rose-Marie and Martina Wiltschko (2002): 'Decomposing Pronouns', *Linguistic Inquiry* **33**, 409–442.

Dimmendaal, Gerrit Jan (1983): *The Turkana Language*. Foris Publications, Dordrecht.

Driemel, Imke and Maria Kouneli (2022): A dedicated topic position in Kipsigis. available at lingbuzz/006571.

Harizanov, Boris (2018): 'Word Formation at the Syntax-Morphology Interface: Denominal Adjectives in Bulgarian', *Linguistic Inquiry* **49**, 283–333.

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.

Heine, Bernd (1981): *The Non-Bantu languages of Kenya*. Vol. 2 of *Language and Dialect Atlas of Kenya*, Reimer, Berlin.

Koopman, Hilda (1984): *The Syntax of Verbs: From Verb Movement in the Kru Languages to Universal Grammar*. Foris Publications, Dordrecht.

Kouneli, Maria (2020): N-to-D movement in the Kipsigis DP. Handout of a talk presented at Linearising Constituents Across Domains.

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

McCarthy, John (2008): 'The Serial Interaction of Stress and Syncope', *Natural Language and Linguistic Theory* **26**, 499–546.

Müller, Gereon (2018): 'Structure Removal in Complex Prefields', *Natural Language & Linguistic Theory* **36**, 219–264.

Nevins, Andrew (2012): Haplological dissimilation at distinct stages of exponence. *In:* J. Trommer, ed., *The morphology and phonology of exponence*. Oxford University Press, Oxford, pp. 84–116.

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

Travis, Lisa (1984): Parameters and effects of word order variation. PhD thesis, Massachusetts Institute of Technology.

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