# Morpheme structure constraints solve three puzzles for theories of blocking in nonderived environments*

Ezer Rasin

Pre-print, August 2022

To appear in *Linguistic Inquiry*

**Abstract**

In Nonderived Environment Blocking (NDEB), a phonological process applies across morpheme boundaries or morpheme-internally when fed by another phonological process but is otherwise blocked. I present a theory of NDEB that attributes blocking to an interaction between morpheme structure constraints (which constrain possible URs in the lexicon) and the usual phonological mapping from URs to surface forms. The theory has some unusual aspects that make it conceptually suspicious, but I will argue that it receives empirical support. Using several case studies, I discuss three puzzles for theories of NDEB previously proposed in the literature, including the Strict Cycle Condition (Mascaró, 1976), Kiparsky's (1993) theory of underspecification, Sequential Faithfulness (Burzio, 2000), Coloured Containment (van Oostendorp, 2007), and Optimal Interleaving with Candidate Chains (Wolf, 2008). I show that none of those theories can deal with all three puzzles and that the proposed theory with morpheme structure constraints succeeds. This result supports a dual-component architecture of phonology (as in SPE) over architectures that eliminate language-specific morpheme structure constraints (i.e., the principle of Richness of the Base in Optimality Theory).

## 1 Introduction

In Nonderived Environment Blocking (NDEB), a phonological process applies across morpheme boundaries or morpheme-internally when fed by another phonological process but is otherwise blocked. To illustrate the phenomenon of NDEB, I will use a simple artificial example modeled after the description of Finnish assibilation in Kiparsky

*For helpful feedback and discussion, I thank Adam Albright, Arto Anttila, Eric Baković, Edward Flemming, Morris Halle, Roni Katzir, Michael Kenstowicz, Daniel Margulis, Joan Mascaró, Miriam Nussbaum, David Pesetsky, Roger Schwarzschild, Sam Zukoff, the audiences at NELS 46 in Concordia University and AMP 2015 in the University of British Columbia, participants of the MIT Workshop of Fall 2014, and especially Donca Steriade. This research was supported by the Israel Science Foundation (grant No. 2872/21) and the Azrieli Foundation.

(1973, 1993) which I will refer to as *Finnish' assibilation*.[1] In this artificial example, assibilation turns the stop /t/ into the strident [s] before the high vowel /i/. The process applies before suffixes that begin with /i/, as in (1a). Morpheme-internally, it applies only when the high vowel is the result of final-vowel raising (which raises /e/ to [i] word-finally), as in (1b). Otherwise, assibilation does not apply within morphemes (1c). The sequence /ti/ is often referred to as a *derived environment* in (1a) and (1b) and as a *nonderived environment* in (1c).

(1) Assibilation in the artificial language Finnish'
   a. Assibilation applies across a morpheme boundary:
      lut-a ∼ lus-i (/lut-i/)
   b. Assibilation applies morpheme-internally when fed by final-vowel raising ($e \rightarrow i$ / __ #):
      vete-pa ∼ vesi (/vete/ → ve**ti** → [vesi])
   c. Otherwise, assibilation is blocked morpheme-internally:
      - tila
      - niti

NDEB is an instance of under-application opacity that poses a challenge to both rule-based phonology (Chomsky and Halle 1968) and Optimality Theory (OT; Prince and Smolensky 1993/2004). In rule-based phonology, a rule of assibilation that turns the stop /t/ into the strident [s] before the high vowel /i/ would incorrectly apply to nonderived /ti/ sequences if no conditions on its application are posited. Similarly, in OT, the markedness constraint *ti would equally penalize derived and nonderived surface sequences of [ti]. And if *ti is allowed to be repaired by assibilation in derived environments (by appropriately ranking it over faithfulness constraints like IDENT[cont]), assibilation would incorrectly apply in nonderived environments as well. More generally, if $P$ is a process that is blocked in nonderived environments, the challenge in both theories is to partition the set of environments of application of $P$ into two subsets – corresponding to derived and nonderived environments – and block the application of $P$ precisely in nonderived environments (in what follows, I will use $P$ as shorthand for a process that is blocked in nonderived environments). Previous works that have tried to address the challenge include Mascaró (1976), Kiparsky (1993), Burzio (2000) Inkelas (2000), Łubowicz (2002), McCarthy (2003) van Oostendorp (2007), Kula (2008), Wolf (2008), and Anttila (2009), among others.

This paper will develop and defend a theory of NDEB which is an extension of Kiparsky's (1993) underspecification theory of NDEB. According to Kiparsky, a pro-

[1] The literature on Finnish assibilation following Kiparsky (1973, 1993) has challenged the original description and offered alternative analyses of Finnish assibilation according to which the process is not blocked in nonderived environments (see Hammond (1992), Wolf (2008), and especially Anttila (2006)). The process as originally described is, to my knowledge, unique in its simplicity among reported cases of NDEB as it demonstrates two types of NDEB effects simultaneously, each of which is particularly simple by itself. For purposes of clarity and simplicity of exposition, it will be convenient to use an artificial variant of the Finnish case before turning to more complex cases from natural languages later in the paper. To be clear, no argument in this paper will be based on this artificial variant, and this paper will not contribute to the debate about the right analysis of actual Finnish.

cess which shows NDEB is structure-building, which means that it can apply to underspecified but crucially not to fully specified structure. On this view, Finnish' assibilation, stated in (2a) below, would be a feature-filling rule that applies to underspecified /T/ but cannot apply to fully specified /t/ (/T/ here stands for a variant of the voiceless alveolar stop /t/ in which the feature [continuant] is not specified). A default rule which applies after assibilation would convert any underspecified /T/ that did not undergo assibilation into [t], as in (2b).

(2) Grammar of Finnish' according to Kiparsky's (1993) theory

a. T → s / __ i

b. T → t

Example (3) shows the derivation of hypothetical [timas-i], which alternates with [timat-a] and includes two potential environments for assibilation, one that is fully contained in the stem and one that spans a morpheme boundary. According to Kiparsky, the second environment contains an underspeficied /T/ which ends up undergoing assibilation and the first environment contains a fully-specified /t/ which is protected from assibilation.

(3) Derivation of [timasi]

| UR | /timaT-i/ |
|---|---|
| T → s / __ i | timas-i |
| T → t | - |
| SR | [timasi] |

As noted by Burzio (2000) and discussed in detail by Inkelas (2000), Kiparsky's analysis of NDEB misses a crucial component, as it leaves the underlying distribution of underspecified /T/ and fully-specified /t/ as an accident of the lexicon. Nothing in the analysis prevents fully-specified /t/ from occurring stem-finally and blocking assibilation before a suffix-initial /i/. The grammar thus freely generates ungrammatical forms such as *[rat-i] in which assibilation has not applied across a morpheme boundary, as illustrated in (4).

(4) Derivation of ungrammatical *[rati]

| UR | /rat-i/ |
|---|---|
| T → s / __ i | - |
| T → t | - |
| SR | *[rati] |

In other words, Kiparsky's theory cannot represent NDEB processes that apply *obligatorily*. For every hypothetical form that surfaces with assibilation (e.g., [· · ·s-i] from /· · ·T-i/), the grammar also generates a minimally different form in which assibilation does not apply ([· · ·t-i] from /· · ·t-i/). However, obligatory NDEB processes are attested in natural language, and we will see some concrete examples later.

In this paper, I propose to respond to Burzio's and Inkelas' challenge for Kiparsky's theory directly, using Morpheme Structure Constraints (MSCs) from early generative phonology which capture phonological generalizations over single morphemes in the lexicon (see Halle 1959, 1962; Stanley 1967; Chomsky and Halle 1968; and Booij 2011 for an overview). With MSCs, the grammar can regulate the distribution of underspecified /T/ and fully-specified /t/ in the lexicon of Finnish'. In particular, I will propose that the grammar of Finnish' includes the MSC in (5). This MSC bans a fully-specified /t/ from underlying morpheme-final position and thus blocks monomorphemic URs like /rat/ which derive ungrammatical forms under Kiparsky's theory.

(5) MSC IN FINNISH' (informal; to be implemented below)
/t/ occurs before /i/; /T/ occurs elsewhere

The MSC in (5) seems like a suspicious addition to the grammar, since it makes a distributional statement in the lexicon that applies to exactly the same environments that trigger assibilation in the phonology. It is thus reminiscent of an old observation by Chomsky and Halle (1968, p. 382) and Kisseberth (1970) that MSCs sometimes duplicate the environment of phonological rules (see also Kenstowicz and Kisseberth 1979, chapter 3). This observation played a role in the abandonment of MSCs in later literature. Within Lexical Phonology and Morphology, Kiparsky (1982) proposed to replace MSCs with phonological processes that apply at the stem level before suffixation. Within OT, MSCs were replaced with surface-oriented constraints, and the ban against MSCs was made explicit using a representational principle, *Richness of the Base*, which prevents language-specific MSCs from being stated:

(6) Richness of the Base (ROTB; Prince and Smolensky 1993, p. 191, Smolensky 1996, p. 3):
  a. All systematic language variation is in the ranking of the constraints.
  b. In particular, there are no language-specific constraints on URs.

With ROTB in place, the hope was that surface-oriented constraints could replace MSCs while also accounting for the phonological mapping (but see Gouskova (to appear) for a recent argument that this is not always the case). To date, however, it has been hard to find empirical evidence for one architecture over the other, and the choice between MSCs and ROTB has been largely left as a matter of taste (though see Rasin and Katzir 2020 for a recent empirical argument for MSCs from learnability). In this paper, I argue that the phenomenon of NDEB can help distinguish between the two architectures empirically. I will present several case studies that pose three puzzles for previous theories of NDEB and argue that the proposed MSC theory is the only theory that can solve all three puzzles. If this argument is correct, then NDEB would provide new empirical support for MSCs.

To preview the structure of the argument, the success of the MSC theory of NDEB will be due to the following properties of MSCs:

(7) a. MSCs can capture generalizations over URs of single morphemes.
  b. In particular, MSCs can filter URs with non-contrastive phonological features.

Independent motivation for capturing generalizations over single morphemes as in (7a) comes from phonological generalizations that seem to apply in the lexicon but not on the surface (Booij 2011): In English, for example, no mono-morphemic form can end in [ɡd], but [ɡd] arises through suffixation, as in [hʌɡ-d] *hugged*. An MSC that bans [ɡd] from the lexicon correctly rules out [ɡd] within morphemes without affecting morphologically-complex words. In Dutch, nasal consonants are homorganic with a following obstruent within morphemes, as in [dɑmp] 'damp', [tɑnt] 'tooth', and [dɑŋk] 'thanks', and no morpheme ends in the sequences [-mt], [-mk], [-ŋp], [-ŋt]; however, such sequences can arise through suffixation: the 3rd singular present verbal suffix /-t/ can be preceded by all three types of nasals, as in [klɪm-t] 'climbs', [zon-t] 'sunbathes', and [zɪŋ-t] 'sings'. Again, an MSC that enforces nasal-place assimilation in the lexicon correctly bans [-mt], [-mk], [-ŋp], [-ŋt] within morphemes without banning them from complex words.

Property (7a) will allow the MSC theory to formalize the characterization of NDEB in (8):

(8) *A characterization of NDEB*
$P$ is blocked in some environment when this environment is present in the UR of some morpheme.

When an environment of $P$ would be present in the UR of some morpheme (e.g., as in the UR of [tila] in Finnish'), MSCs like (5) would ensure that $P$ is unable to apply in that environment (e.g., by requiring the input /t/ to be fully-specified). Environments created outside of URs will avoid the blocking MSC.

Some theories of NDEB do not rely on the characterization in (8) and deny a direct role for URs in blocking. Examples of such theories are the Strict Cycle Condition (Mascaró, 1976), Coloured Containment (van Oostendorp, 2007), and a cyclic variant of the proposed MSC theory that will be discussed below. The first two puzzles discussed in section 3 – *Persistent blocking* and *Blocking within suffixes* – will directly support (8), providing an argument for the MSC theory over such theories.

Other theories of NDEB have followed (8) but have formalized it within architectures that adopt ROTB and reject MSCs. Such theories include Sequential Faithfulness (Burzio, 2000), Kula's (2008) theory, and Optimal Interleaving with Candidate Chains (Wolf, 2008).[2] Since such theories adopt ROTB, they lack the property in (7b) and thus differ from the MSC theory in whether they count non-contrastive features as 'derived'. To illustrate the difference, consider flapping in some varieties of English, which is fully predictable: simplifying, /t/ is flapped between a preceding stressed vowel and a following unstressed vowel (e.g., [ǽɾəm] 'atom') but flaps never appear elsewhere (e.g., *[ɾæɡ]). To account for the distribution of flaps, an MSC-based account can use the MSC in (9) along with a rule that creates flaps in the right environment. The MSC in (9) will ensure that flaps never surface except when generated by the rule. On this account, URs like /ɾæɡ/ are never generated and thus flaps are necessarily derived.

(9) MSC in English
No flaps in the lexicon

[2] In fact, Wolf (2008) follows a variant of (8), but the differences will be irrelevant for the purpose of the current discussion.

Differently from the MSC-based account of flapping, an account using ROTB will generate URs like /ɾæɡ/, with underlying non-contrastive flaps, and enforce the correct distribution of flaps through the mapping from URs to surface forms. Significantly, given ROTB, flaps are not necessarily derived. As we will see, theories of NDEB that formalize (8) without MSCs run into trouble due to their inability to filter URs with non-contrastive features, a problem of *Too-Many-URs*. The problem of Too-Many-URs arises in languages like Romanian, where a process that is blocked in nonderived environments is triggered by non-contrastive features (the puzzle of *Non-contrastive triggers* below). While non-contrastive features must count as derived for the purposes of NDEB, ROTB generates them liberally in URs, causing (8) to misidentify some derived environments as nonderived and vice versa. The MSC theory of NDEB, given the property in (7b), can ban derived environments with non-contrastive features from URs using MSCs like (9) and correctly distinguish derived from nonderived environments at the UR level.

The paper is structured as follows. First, in section 2, I implement an architecture that uses MSCs to regulate the distribution of underspecified and fully-specified structure in the lexicon and develop an analysis of Finnish' assibilation within this architecture. Then, in section 3, I present the three puzzles using several case studies from natural languages and show how they are accounted for using the MSC theory of NDEB. In section 4, I discuss the predictions of previous approaches to NDEB, including the Strict Cycle Condition (Mascaró, 1976), Coloured Containment (van Oostendorp, 2007), Sequential Faithfulness (Burzio, 2000), and Optimal Interleaving with Candidate Chains (Wolf, 2008). I show that each approach fails on at least one puzzle, as summarized in (10).

(10) Summary of the discussion in sections 3 and 4. ✓ indicates that the theory succeeds on the puzzle; ✗ indicates that it fails. Only the MSC theory succeeds on all three puzzles.

| Theory ↓ Puzzle → | Persistent blocking | Blocking within suffixes | Non-contrastive trigger |
|---|---|---|---|
| MSC theory of NDEB | ✓ | ✓ | ✓ |
| Strict Cycle Condition | ✗ | | |
| Coloured Containment | ✗ | | |
| Cyclic variant | | ✗ | |
| Sequential Faithfulness | | | ✗ |
| OT with Candidate Chains | | | ✗ |

# 2 Proposal

## 2.1 Architecture

This subsection describes the phonological architecture that will be used in subsection 2.2 for an account of NDEB. My claim in this paper is that NDEB supports a component that restricts possible URs in the lexicon. I will have nothing to say about the phonological formalism (e.g., rule-based or constraint-based) and I will not make commitments with respect to the nature of lexical representations (e.g., underspecified or fully specified). To make the proposal explicit, I will present it using a rule-based formalism and underspecification.[3] The architecture, which I now describe, is schematized in (1).

| | | | | | |
|---|---|---|---|---|---|
| Lexicon | $\{M_1\}$ | $\{M_2\}$ | $\{M_3\}$ | $\leftarrow$ | $\Sigma_L$ |
| | $\downarrow$ | $\downarrow$ | $\downarrow$ | | |
| | $/M_1/$ | $/M_2/$ | $/M_3/$ | | |
| Phonology | | $/[[M_1 - M_2] - M_3]/$ | | | |
| | | $\downarrow$ | | | |
| | | $[M_1 - M_2 - M_3]$ | | | |

Figure 1: A schema of the proposed architecture of grammar. The morpheme representations $\{M_1\}$, $\{M_2\}$, and $\{M_3\}$ are *initial representations* generated by concatenating elements from $\Sigma_L$, the alphabet of the language. Morpheme-structure rules apply to each morpheme individually to generate the URs $/M_1/$, $/M_2/$, and $/M_3/$. The phonology receives a structured UR created by the morphology and applies phonological rules to map this UR to the surface form $[M_1 - M_2 - M_3]$.

A central component of the architecture is the mapping from URs to surface forms, which is implemented here using ordered phonological rules as in SPE. I assume that a phonological grammar includes an alphabet – an inventory of feature bundles $\Sigma$ – the elements of which can be concatenated. For example, if $\mathrm{k, a, t} \in \Sigma$, then {kat} and {takta} are possible concatenations, among many others. I assume that individual languages can restrict $\Sigma$ to a proper subset, call it $\Sigma_L$. For a segment $\sigma \in \Sigma$, we can write $\sigma \notin \Sigma_L$, meaning that $\sigma$ cannot be used for concatenation in that language. For example, if English rules out x from its alphabet and we write x $\notin \Sigma_L$, then {bax} is not a possible concatenation in English. Negative statement such as x $\notin \Sigma_L$ are used for convenience and should not be taken to be grammatical constraints per se. What I mean by writing x $\notin \Sigma_L$ is that $\Sigma_L$, which could be positively stated in the grammar as a set of segments, does not include x. I will refer to representations created by concatenating elements from $\Sigma_L$ as *initial representations*, and I will mark them using curly brackets, as in {anta}. Morpheme structure rules map initial representations to URs. For example, if {anta} is an initial representation and post-nasal voicing (t $\rightarrow$ d / n __ )

[3] The choices of rules and underspecification are arbitrary: the mappings presented in this paper using a rule-based formalism can be reformulated using constraints, and one can think of variants of the proposal that do not make use of underspecification. For example, the distinction between unspecified and specified features can be replaced with a distinction between plain specified features and specified features with an exception diacritic that prevents a feature from being changed by a particular rule, as in SPE.

is the only morpheme structure rule in the grammar, the result of applying post-nasal voicing to {anta} would be the UR /anda/. Morpheme structure rules have the same format as ordinary rules, but they apply to isolated morphemes in the lexicon before the morphemes are combined (see Borowsky 1993, Baker 2005, and Bermúdez-Otero 2018 for alternatives to MSCs that allow phonological processes to apply to at least some individual morphemes before morphology applies). In this architecture, then, URs are created in two steps: first, elements from $\Sigma_L$ are concatenated to form an initial representation. Then, morpheme structure rules apply and map this representation to a UR. Later on, after the morphology, phonological rules map URs to surface forms. Whether a rule is a morpheme structure rule or a regular phonological rule is determined on a language-specific basis.

In addition, I assume that lexical representations may be underspecified, which means that segments in $\Sigma$ (and in $\Sigma_L$) may lack a specification for some of their features. See Kiparsky (1982), Archangeli (1988), and Steriade (1995) for relevant discussion. For example, a variant of the voiceless alveolar obstruent [t] in which the feature [continuant] is not specified may be in $\Sigma$. We can refer to this segment as [T] and write $\mathrm{T} \in \Sigma$. Underspecified features are filled in either by morpheme structure rules or by phonological rules (Halle 1959). Finally, both morpheme structure rules and phonological rules may be *feature filling*. This means that they can target segments underspecified for some feature $F$ and fill in the relevant value but, crucially, without affecting segments that are already specified for $F$. Example (11) demonstrates the property of feature filling using a version of Finnish' assibilation that applies to underspecified /T/.

(11) Assibilation: $\mathrm{T} \rightarrow \mathrm{s}\ /\ \_\_\ \mathrm{i}$ (feature-filling)

a. /Ti/ $\xrightarrow{\text{assibilation}}$ [si]

b. /ti/ $\xrightarrow{\text{assibilation}}$ [ti]

## 2.2 Analysis

In this subsection I provide an analysis of Finnish' assibilation using the architecture described in section 2.1. The basic pattern of Finnish' assibilation was presented above in (1a)-(1c) and is repeated here as (12a)-(12c). Following the convention in the literature, I use the term *morphologically-derived environment* to refer to an environment created through affixation, as in (12a), and *phonologically-derived environment* to refer to an environment created through the application of a phonological process, as in (12b).

(12) Assibilation in the artificial language Finnish' (repeated)

a. Assibilation applies across a morpheme boundary:
lut-a ∼ lus-i (/lut-i/)

b. Assibilation applies morpheme-internally when fed by final-vowel raising ($\mathrm{e} \rightarrow \mathrm{i}\ /\ \_\_\ \#$):
vete-pa ∼ vesi (vesi (/vete/ → ve**ti** → [vesi])

c. Otherwise, assibilation is blocked morpheme-internally:

- tila
- niti

The first ingredient in the analysis is the rule of assibilation in (13), which, following Kiparsky (1993), I take to be a feature-filling rule that specifies the voiceless alveolar /T/ as [+continuant].[4] The second ingredient is a rule that I refer to as *anti-assibilation* (14).

(13) Assibilation
$T \rightarrow s$ / __ i (feature-filling)

(14) Anti-assibilation
$T \rightarrow t$ / __ i (feature-filling)

Anti-assibilation is similar to the rule of assibilation: it is a feature-filling rule that applies in the same environment (/Ti/) and fills in a value for the feature [continuant]. The difference between the two is that anti-assibilation specifies that value as [-continuant] rather than [+continuant]. That is, anti-assibilation specifies /T/ as [t].

To see how assibilation and anti-assibilation interact, consider the UR /Ti/ and a hypothetical grammar in which anti-assibilation is ordered before assibilation. The derivation is provided in (15).

(15) Interaction between assibilation and anti-assibilation (hypothetical grammar)

| UR | /Ti/ |
|---|---|
| $T \rightarrow t$ / __ i | ti |
| $T \rightarrow s$ / __ i | - |
| SR | [ti] |

First, anti-assibilation applies and specifies /T/ as [t]. Then, assibilation does not apply since its structural description is not met: the rule is feature filling, but /t/ is not underspecified for continuancy. The result is the surface form [ti]. In short, anti-assibilation bleeds assibilation by destroying its environment of application.

My proposal is that in the grammar of Finnish', anti-assibilation is a morpheme structure rule and hence applies to isolated morphemes in the lexicon, whereas assibilation is a phonological rule that is part of the mapping from URs to surface forms. Fully-specified /t/ is not part of the Finnish' alphabet.

(16) Morpheme structure component:

a. $t \notin \Sigma_L$

b. $T \rightarrow t$ / __ i

The consequence for the form of URs in Finnish' is that /t/ and /T/ are in complementary distribution in the lexicon: /t/ occurs only before /i/ (following the application of anti-assibilation) and /T/ occurs elsewhere. Here are some examples:

(17) Distribution of /t/ and /T/ in URs

[4] For presentational ease, I ignore the feature [strident], which could be filled in by the assibilation rule itself or by a separate rule.

a. $\{$Tila$\}$ $\rightarrow$ /tila/

b. */lata/

c. /laTa/, /luT/

Example (17a) shows the derivation of the UR /tila/. Since t $\notin \Sigma_L$, any instance of /t/ in URs must be derived from /T/. The initial representation is therefore $\{$Tila$\}$, which anti-assibilation maps to /tila/. Example (17b) indicates that /lata/ is not a possible UR in Finnish': since t $\notin \Sigma_L$ and the environment for anti-assibilation is not met before /a/, /t/ cannot occur before /a/ in the UR.

In (17c), anti-assibilation does not apply, and /T/ remains underspecified. The value for [continuant] will be filled in by the mapping from URs to surface forms: the rule of assibilation turns /T/ into [s] before /i/; otherwise – that is, whenever assibilation does not apply – /T/ is specified as [t] through the default rule T $\rightarrow$ t. The two phonological rules are given in (18).

(18) Phonological rules:

a. T $\rightarrow$ s / __ i

b. T $\rightarrow$ t

Example (20) demonstrates the application of phonological rules in the derivation of the alternants in (19), assuming the UR /luT/ for the stem.

(19) lut-a $\sim$ lus-i

(20) Derivation of the alternants lut-a $\sim$ lus-i

| UR | /luT-i/ | /luT-a/ |
|---|---|---|
| T $\rightarrow$ s / __ i | lusi | - |
| T $\rightarrow$ t | - | luta |
| SR | [lusi] | [luta] |

Putting together the morpheme structure component and the phonological rules, this is the grammar of Finnish' we have so far:

(21) a. Morpheme structure component:
- t $\notin \Sigma_L$
- T $\rightarrow$ t / __ i

b. Phonological rules:
- T $\rightarrow$ s / __ i
- T $\rightarrow$ t

I will now show why this grammar applies assibilation in morphologically-derived environments but not in nonderived environments. Consider the derivation of hypothetical [timas-i], which alternates with [timat-a] and includes two potential environments for the application of assibilation: the first is morpheme-internal and the second spans the morpheme boundary. Assibilation only applies in the latter:

(22) timat-a $\sim$ timas-i

The derivation of [timas-i] is shown in (23).

(23) Derivation of [timas-i] (alternant: [timat-a])

a. Morpheme structure rules apply to each morpheme individually:
   - $\{$TimaT$\}$ $\rightarrow$ /timaT/
   - $\{$i$\}$ $\rightarrow$ /i/

b. Phonological rules apply:

| UR | /timaT-i/ |
|---|---|
| T $\rightarrow$ s / __ i | timas-i |
| T $\rightarrow$ t | - |
| SR | [timasi] |

First, morpheme structure rules apply to each morpheme individually (23a). Since t $\notin \Sigma_L$, the initial representation of the stem must be $\{$TimaT$\}$. Anti-assibilation applies to the first instance of /T/ but not to the second, because at this stage of the derivation the second /T/ is stem-final, so the environment for anti-assibilation is not met. The result is the UR /timaT/, where only the second /T/ remains underspecified for continuancy. In the mapping from URs to surface forms shown in (23b), assibilation successfully applies to the sequence /T-i/ which was created through affixation. It does not apply to the stem-initial /ti/, which at this point is already fully specified. The final surface form is [timasi].

The next step is to show why assibilation applies in phonologically-derived environments. Recall that final-vowel raising (24) raises a word-final /e/ to [i] (24a). Assibilation may apply morpheme-internally when fed by final-vowel raising (24b).

(24) e $\rightarrow$ i / __ #

a. juke-pa $\sim$ juki

b. vete-pa $\sim$ vesi

Here, nothing further has to be said. Final-vowel raising is ordered before assibilation (25). In words like [vesi], alternating /T/ precedes /e/ in the UR, so anti-assibilation does not get to apply. Therefore, /T/ remains underspecified, which means that assibilation will get to apply after affixation. The full derivation is provided in (26).

(25) a. Morpheme structure component:
   - t $\notin \Sigma_L$
   - T $\rightarrow$ t / __ i

b. Phonological rules:
   - e $\rightarrow$ i / __ #
   - T $\rightarrow$ s / __ i
   - T $\rightarrow$ t

(26) Derivation of [vesi]

a. Morpheme structure rules apply (vacuously):
$\{$veTe$\}$ $\rightarrow$ /veTe/

b. Phonological rules apply:

| UR | /veTe#/ |
|---|---|
| e → i / __ # | veTi# |
| T → s / __ i | vesi# |
| T → t | - |
| SR | [vesi] |

In sum, a process $P$ that is blocked in nonderived environments applies unless its input is made immune in an earlier stage of the derivation. Inputs can be made immune by a feature-filling rule *anti-P* that shares its structural description with $P$ and can apply to isolated morphemes in the lexicon. *Anti-P* thus induces the following partition on the set of environment of $P$:

(27) Partition into nonderived and derived environments

a. Environments present when *anti-P* applies (correspond to nonderived environments)

b. All other environments (correspond to derived environments)

As mentioned in the introduction, MSCs of the form *anti-P* are unusual. Their formulation seems arbitrary and their environment duplicates the environment of $P$, and at present I have nothing to say to make them more principled. Instead, I will focus on case studies that distinguish the predictions of (27) from the predictions of previous proposals and argue that MSCs should be adopted on empirical grounds. In the next section, I present three puzzles for theories of NDEB and show how the MSC theory accounts for them successfully.

# 3 Three puzzles for theories of NDEB

## 3.1 Puzzle I: Persistent blocking

The MSC theory determines the alternation status of a feature at the individual morpheme level. Consider again the blocking of Finnish' assibilation in morphologically-nonderived environments, using the example [timas-i] (which alternates with [timat-a]). For the MSC theory, blocking is exclusively determined according to the environment of each potential input for assibilation in the UR /timaT/: the first consonant – but not the stem-final consonant – precedes /i/ and therefore becomes immune to assibilation. Once a potential input to assibilation is made immune before suffixation, it is predicted to stay immune even if the environment for assibilation is destroyed and then re-created after suffixation, a scenario that I will refer to as *persistent blocking* (the same scenario was discussed in Kiparsky 1993 and Wolf 2008, section 4.2.5, under different names). To illustrate persistent blocking, suppose that Finnish' had a process of vowel deletion that deletes stem-final vowels pre-vocalically and may feed assibilation:

(28) /munte-i/ → [muns-i]

The prediction is that assibilation would not apply to suffixed forms if the stem-final vowel is /i/, even though the environment for assibilation is re-created through suffixation and spans the morpheme boundary:

(29) Prediction of the MSC theory for /munti-i/
/munti-i/ → [munt-i]

On the MSC theory, the morpheme structure component captures the distinction between the two verbs at the UR level: the URs are /munTe/ (with underspecified /T/) and /munti/ (with fully specified /t/). Assibilation can only apply to the first. /t/ in /munti/ remains immune to assibilation even after the deletion of the stem-final vowel and the addition of an /i/-initial suffix which creates the environment for assibilation. Note that the opposite of persistent blocking (a similar pattern but where /munti-i/ becomes [muns-i]) cannot be generated by the MSC theory given the assumptions adopted in this paper.

Other theories of NDEB that predict persistent blocking include those of Burzio (2000) and Wolf (2008). In contrast to these theories, much of the previous literature on NDEB has followed a different idea regarding the characterization of NDEB. The guiding intuition is that in both types of environments in which $P$ applies – across a morpheme boundary and when part of its environment is the result of another phonological process – part of the environment is 'new', or, stated differently, is introduced in the course of the derivation. In the Finnish' assibilation case, the environment in /lut-i/ is 'new' because it is formed through affixation, and the environment in /veti/ (derived from /vete/ through vowel raising) is 'new' because the high vowel is the result of vowel raising. Theories guided by this idea, like the Strict Cycle Condition (Mascaró, 1976) and Coloured Containment (van Oostendorp, 2007), incorporate a notion of 'new' or 'derived' environments into the grammar and often introduce a licensing condition to allow the application of $P$ only in such environments. I will refer to such theories as *derived-environment* theories, since they try to characterize the set of derived environments rather than the environments in which $P$ is blocked.

In derived-environment theories, the application of $P$ is determined based on the morphologically-complex form. For [timas-i], the relevant representation would be /timat-i/, the suffixed form before the application of assibilation. Assibilation applies in the second environment (/tima**t-i**/) but not in the first (/**ti**mat-i/) since only the second environment is 'derived' and spans a morpheme boundary. Below, I will discuss in more detail some of these theories and how they enforce application in derived environments. For now, what matters is that they all license application across a morpheme boundary:

(30) *Property of derived-environment theories*
Spanning a morpheme boundary is a sufficient condition for licensing.

Persistent blocking is puzzling for derived-environment theories. For the Finnish' example /munti-i/, the prediction is that assibilation would apply after vowel deletion, since an environment that triggers assibilation would span a morpheme boundary:

(31) Prediction of derived-environment theories for /munti-i/
/munti-i/ → [munsi]

In the remainder of this section I show that persistent blocking is attested in Romanian, supporting the prediction of the MSC theory and providing a direct argument against derived-environment theories. The case study is Romanian palatalization and the data are taken from unpublished notes by Donca Steriade (Steriade 2008b).

In Romanian, a palatalization rule turns a velar stop into a palatal before a front vowel or glide:

(32) Romanian palatalization

a. k → tʃ / __ {e, i, j}

b. g → dʒ / __ {e, i, j}

Palatalization obligatorily applies across morpheme boundaries (33) and is blocked morpheme-internally (34).[5]

(33) mak ∼ matʃ-j, *mak-j ‘poppy-SG.’∼‘poppy-PL.’

(34) a. unkj ‘uncle-SG.’

b. rokie ‘dress-SG.’

c. paket ‘package-SG.’

Vowels are deleted before the plural suffix /-i/, which is sometimes realized as a glide (35).

(35) a. metru ∼ metr-i ‘meter-SG.’∼‘meter-PL.’

b. bere ∼ ber-j ‘beer-SG.’∼‘beer-PL.’

c. popʌ ∼ pop-j ‘priest-SG.’∼‘priest-PL.’

The vowel-glide alternation is irrelevant for our current purposes, so without getting into details I will simply assume that deletion applies pre-vocalically and that a cover rule i → j, which is responsible for the glide-vowel alternation, applies after deletion (see Steriade 1984 and Chitoran 2002, section 5.3 for possible analyses of this process).

Crucially, palatalization is blocked exactly when the deleted vowel had been a palatalization trigger, creating an instance of persistent blocking. In (36a), the final vowel in the singular is a back vowel and palatalization applies in the plural. In (36b), the final vowel is a front vowel and palatalization in the plural is blocked.

(36) a. mɨnekʌ ∼ mɨnetʃ-j ‘sleeve-SG.’∼‘sleeve-PL.’

b. pʌduke ∼ pʌdu**k**-j ‘louse-SG.’∼‘louse-PL.’

This contrast is fully general. The following table demonstrates the behavior of palatalization in the plural form of every nominal declension class that takes the plural suffix /-i/. For each class, the two rightmost columns indicate the identity of the stem-final vowel and whether palatalization applies in the plural form.

(37) Palatalization in Romanian nouns that take the plural suffix /-i/[6]

[5]For presentational ease, I have omitted secondary palatalization from the Romanian examples. The distribution of secondary palatalization is irrelevant for our purposes.

[6]Romanian has a process that deletes /u/ word-finally (Steriade 1984; Chitoran 2002, section 5.3). This is why an [u] is missing on the surface in the singular form in line (a).

| | Noun-SG. | Noun-PL. | | Final vowel | Palatalzation? |
|---|---|---|---|---|---|
| MASC | | | | | |
| a. | mak | matʃ-j | ‘poppy’ | /u/ | ✓ |
| b. | pʌduke | pʌdu**k**-j | ‘louse’ | e | * |
| c. | dukʌ | dutʃ-j | ‘duke’ | ʌ | ✓ |
| d. | flamiŋgo | flamiŋdʒ-j | ‘flamingo’ | o | ✓ |
| FEM | | | | | |
| e. | fabrikʌ | fabritʃ-j | ‘factory’ | ʌ | ✓ |
| f. | pereke | pere**k**-j | ‘pair’ | e | * |

For the sake of concreteness, let us see why the MSC theory accounts for this pattern without modification. The grammar, given in (39), has anti-palatalization (38) as a morpheme-structure rule and palatalization as a phonological rule.[7]

(38) Anti-palatalization
$K \rightarrow k$ / __ {e, i, j}

(39) a. Morpheme structure component:
- $k \notin \Sigma_L$
- $K \rightarrow k$ / __ {e, i, j}

b. Phonological rules:
- $V \rightarrow \emptyset$ / __ $i_{[PL]}$[8]
- K → tʃ / __ {e, i, j}
- $K \rightarrow k$
- $i \rightarrow j$

Anti-palatalization applies in the lexicon and specifies K as /k/ in (36b) but not in (36a). This is shown in the following derivations.

(40) Derivation of [pʌduk-j] (singular: [pʌduke])

a. Morpheme structure rules apply:

1. {pʌduKe} → /pʌduke/
2. {i} → /i/

[7]The range of possible noun-final vowels in Romanian is restricted, suggesting that the final vowel should be regarded as an idiosyncratic theme vowel specified on a root-by-root basis. If this is indeed the case, a necessary assumption on the present account would be that the theme vowel is special in that it is present already in the morpheme-structure component. An alternative is that it is added later, in which case anti-palatalization would have to be relegated to the phonology. The latter option, where anti-$P$ is relegated to the phonology, possibly on a language-specific basis, is developed in section 3.2. I will not attempt to choose between the two options here.

[8]The facts of hiatus resolution in Romanian are significantly more complex than this rule suggests. See Chitoran (2002, chapter 4).

b. Phonological rules apply:

| UR | /pʌduke-i/ |
|---|---|
| V → ∅ / __ V | pʌduki |
| K → tʃ / __ {e, i, j} | - |
| K → k | - |
| i → j | pʌdukj |
| SR | [pʌdukj] |

(41) Derivation of [mɨnetʃ-j] (singular: [mɨnekʌ])

a. Morpheme structure rules apply (vacuously):

1. {mɨneKʌ} → /mɨneKʌ/
2. {i} → /i/

b. Phonological rules apply:

| UR | /mɨneKʌ-i/ |
|---|---|
| V → ∅ / __ V | mɨneKi |
| K → tʃ / __ {e, i, j} | mɨnetʃi |
| K → k | - |
| i → j | mɨnetʃj |
| SR | [mɨnetʃj] |

Romanian palatalization, then, shows persistent blocking and thus supports the prediction of the MSC theory. In section 4.1, I discuss specific derived-environment theories, including the Strict Cycle Condition and Coloured Containment, which license the application of $P$ across morpheme boundaries as in (30) and thus incorrectly predict the application of palatalization in /pʌduke-i/.

## 3.2 Puzzle II: Blocking within suffixes

The MSC-based analysis of Finnish' assibilation relies on MSCs to restrict the distribution of /t/ and /T/ in the lexicon: non-alternating /t/ occurs before /i/, and alternating /T/ occurs elsewhere. MSCs apply to each morpheme individually and ensure that stems like /timaT/ have the desired specification before suffixation. The same result could be alternatively achieved in a cyclic architecture where MSCs are replaced with first-cycle evaluation and the distributional restriction applies once, before suffixation. My goal in this section is to discuss the cyclic variant of the MSC-based proposal. While the cyclic variant can successfully capture the Finnish' assibilation pattern, I will show that it faces a challenge in accounting for cases of NDEB where application is blocked not only within stems, but also within suffixes. As we will see, accounting for blocking within suffixes will require a level of representation in which phonological restrictions apply to suffixes in isolation from the rest of the string – a level available in MSC-based architectures but crucially not in cyclic architectures that reject MSCs. Of course, the MSC theory is perfectly consistent with cyclic rule application after the morphology; the question addressed in this section is whether cyclicity without MSCs is *sufficient*.

Cyclic architectures allow phonological processes to be interleaved with affixation. Examples of cyclic architectures are Lexical Phonology and Morphology (Pesetsky, 1979, Kiparsky, 1982 et seq.), its implementation within OT known as Stratal OT (Kiparsky, 2000; Bermúdez-Otero, 2011, 2018), and Halle and Vergnaud's (1987) theory of the cycle. I will first show that a cyclic variant of the MSC-based analysis can account for Finnish' assibilation without MSCs (similarly, it can account for the behavior of Romanian palatalization discussed above). In this variant, there are no restrictions on the alphabet, which means that both /t/ and /T/ can be used for writing URs. Moreover, since anti-assibilation is not an MSC, /t/ and /T/ may occur anywhere within URs. Therefore, URs like /rat/ (with fully-specified /t/ in final position) and /Tila/ (with underspecified /T/ before /i/) can be generated. A cyclic grammar is provided in (42). It contains two rule blocks separated by suffixation. To keep the discussion general and compatible with various cyclic architectures, I will not name the rule blocks but will refer to them simply as Rule block A and Rule block B.[9] Rule block A contains two rules, which mirror the effects of MSCs in the MSC-based analysis. The first rule turns every /t/ into [T] and has a similar effect to the constraint $t \notin \Sigma_L$, which banned /t/ from initial representations in the MSC-based analysis. The second rule is the anti-assibilation rule. The remaining rules, including assibilation, are part of Rule block B.

(42) a. Rule block A:
$t \rightarrow T$
$T \rightarrow t$ / __ i
b. Add the suffix /-i/
c. Rule block B:
$T \rightarrow s$ / __ i
$T \rightarrow t$

The derivation of [timas-i] (which alternates with [timat-a]) using this grammar is given in (43). As there are no MSCs, multiple URs for the stem may lead to the same output. To see the rules in working, I have chosen the UR /Tilat/. Even given this UR, the correct output is derived. The analysis straightforwardly extends to phonologically-derived environments if final-vowel raising is placed in Rule Block B.

(43) Cyclic derivation of [timas-i]

| Rule block A | /Tilat/ |
|---|---|
| $t \rightarrow T$ | TimaT |
| $T \rightarrow t$ / __ i | timaT |
| Suffixation | /timaT-i/ |
| Rule block B | /timaT-i/ |
| $T \rightarrow s$ / __ i | timasi |
| $T \rightarrow t$ | - |
| | [timasi] |

A cyclic architecture, then, can capture the Finnish' assibilation pattern without using MSCs since it can impose the same distributional restriction on the stem before

[9]The distinction between Rule block A and Rule block B may correspond to the following distinctions made in cyclic theories: cyclic vs. post-cyclic, stem-level vs. word-level, lexical vs. post-lexical, etc.

suffixation. More generally, the cyclic architecture succeeds because every nonderived environment is introduced into the derivation before every derived environment. This allows anti-$P$ to be ordered at a stage in the derivation after every nonderived environment has been created and before any derived environment has been created, which, in turn, allows anti-$P$ to apply exclusively to nonderived environments and $P$ to apply later to derived environments.

The MSC architecture and the cyclic architecture without MSCs diverge in their predictions when the derivational precedence between nonderived and derived environments required by the cyclic theory breaks down. This may happen when a phonological process that is blocked in nonderived environments is also blocked within suffixes. Cases of such blocking mentioned in the literature are consonant gradation in Finnish (Kiparsky 1993, 2003), spirantization in Luganda (Wolf 2008, citing Odden 1990), and palatalization in Meskwaki (Wier, 2004). The challenge to the cyclic theory from Luganda spirantization was discussed by Wolf (2008, pp. 443-447), and I will present another version of the argument from Finnish consonant gradation.

Finnish consonant gradation (CG) de-geminates a double stop at the onset of a closed syllable and yields alternations as in (44), as illustrated in example (45) from Kiparsky 1993.[10]

(44) tten → ten, ttain → tain

(45) hattu ∼ hatu-n  'hat-NOM.SG.'∼'hat-GEN'

Example (46), also taken from Kiparsky (1993), is a single example that contains three environments for CG. CG is blocked in the first environment (underlined), which is derived, and applies in the other two environments (boldfaced), which are nonderived. The second geminate (/...to**tt**i.../) and the third geminate (/...**tt**oma.../) undergo CG since they are onsets of closed syllables at some level of representation.[11]

(46) /ho<u>tt</u>ento**tt**i-**tt**oma-ta/ → [ho<u>tt</u>ento**t**i-**t**on-ta]  'Hottentotless-PART.SG.'

CG is blocked when its environment is fully contained within the suffix -tten, an allomorph of the genitive plural (47).

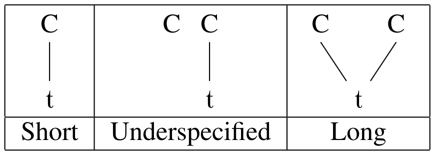
(47) maa-i-tten  *maa-i-ten  'country-PL.GEN'

The MSC theory can account for this blocking effect once a prosodic analog of segmental underspecification is adopted. An analog of this kind was proposed and defended by Kiparsky (1993) as part of his autosegmental underspecification-based account of Finnish CG. On this theory, a short consonant is associated with one autosegmental C (consonantal) slot, a long consonant is associated with two C slots, and a consonant underspecified for length is associated with one C slot that is preceded by another, unassociated C slot, as illustrated in (48).

[10]Kiparsky (2003) notes that the conditions of application of CG are more complicated. As far as I can tell those complications do not affect the argument of the paper, so I will ignore them for simplicity.

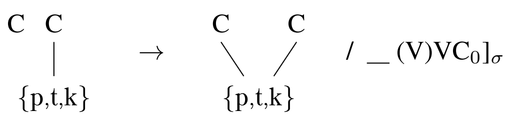
[11]For the second geminate, the syllable is closed by the third, suffix-initial geminate. For the third geminate, the syllable is closed after deletion of the suffix-final vowel /a/ triggered by the following suffix.

(48) Autosegmental underspecification




Given Kiparsky's analog, the usual MSC-based account will work as it did before. Considering the relevant alternating stops (p, t, k), the alphabet of Finnish will include only short consonants and consonants underspecified for length (the unassociated C is therefore not freely distributed in the language). An anti-CG MSC will turn the underspecified consonants into long consonants in the environment of CG (i.e., in the onset of a closed syllable) by associating the consonantal segment with its unassociated C slot (49). The application of anti-CG results in a long consonant which will later be immune to CG. CG itself, which applies after the morphemes have been combined, will only apply to underspecified consonants and turn them into short consonants in the environment of CG by deleting the unassociated C slot. All remaining underspecified consonants will be made long consonants by a default context-free rule.

(49) Anti-CG




Derivations illustrating this analysis are given in (50) and (51) below. Derivation (50) shows how CG applies across morphemes to the suffixed form [hatu-n] in (45).

(50) Derivation of [hatu-n]

a. Morpheme structure rules apply (syllabification, anti-CG):
   1. {haTTu} → /haT.Tu/
   2. {n} → /n/

b. Phonological processes apply:

| UR | /haT.Tu-n/ |
|---|---|
| CG | hatu-n |
| TT → tt | - |
| SR | [hatu-n] |

Since long consonants are banned from the alphabet of the language, the initial representation of the voiceless coronal stop is underspecified for length (the underspecified autosegmental representation in (48) is abbreviated as TT below and a long consonant as [tt]). To correctly identify the closed-syllable environment of CG, syllabification will have to apply as an additional MSC, in addition to its regular application as part of the mapping from URs to surface forms.[12,13] The first C slot of consonants underspecified for length is assumed to close the preceding syllable. On these assumptions,

[12] Applying syllabification as an MSC does not entail that syllabification can be contrastive. In the architecture proposed here, syllable structure will be universally banned from initial representations, so any syllable structure will have to be generated by general rules, whether MSCs or regular phonological rules.

[13] I leave open the question of whether MSCs are crucially ordered with respect to each other or whether

the environment for CG is not present in the morpheme-structure component in (50), so anti-CG does not apply, leaving the voiceless stop underspecified in the UR. Then, CG applies in the phonology, converting the underspecified consonant into a short consonant in a closed syllable.

The derivation in (51) shows how CG gets blocked in nonderived environments, using the word in (47) where the environment of CG is fully contained within a single morpheme.

(51) Derivation of [maa-i-tten]

a. Morpheme structure rules apply (syllabification, anti-CG):
   1. {maa} → /maa/
   2. {i} → /i/
   3. {TTen} → /tten/

b. Phonological rules apply:

| UR | /maa-i-tten/ |
|---|---|
| CG | - |
| TT → tt | - |
| SR | [maa-i-tten] |

As before, each morpheme is evaluated separately in the morpheme-structure component. Anti-CG applies to the final suffix, which is initially represented as {TTen}, and converts the underspecified voiceless stop into a long consonant. The UR of this morpheme is therefore /tten/, which means that it is immune to CG. As a result, CG fails to apply in the phonology.

This analysis raises the question of how the MSC-based theory deals with prosodic NDEB effects more generally. Since the answer depends on various theoretical choices that are orthogonal to those necessary for the argument of the current paper, I will not answer the question here but only present some possibilities. The analysis of Finnish CG illustrates the prediction that a prosodic process can show NDEB effects as long as the theory makes possible underspecified representations that can serve as its input. This means, for example, that epenthesis cannot show NDEB effects, on the assumption that epenthetic segments are not present in the lexicon to be protected by an MSC.[14] Deletion will be able to show NDEB effects as long as segments can be structurally protected in the lexicon from deletion. An immunity account of an NDEB

they should be implemented as if-then conditions rather than rewrite rules, since the choice is immaterial for the argument of the paper. See Chomsky & Halle (1968, p. 386) for an early argument that MSCs should not be ordered.

[14] The only reported NDEB pattern I know of that involves epenthesis is glottal-stop epenthesis in Makassarese, discussed by McCarthy (2003), where a glottal stop is inserted post-vocalically at the end of a word but only after epenthetic vowels (/...CV/ → [...CV] but /...C/ → [...Caʔ]). Since the triggering process of the allegedly NDEBed process is epenthesis, there are various ways of reanalyzing this pattern as a non-NDEBed process under the MSC-based theory. For example, the rule that inserts the vowel might insert the glottal component as well (∅ → aʔ), in which case there is no separate glottal-stop epenthesis process to begin with. Alternatively, vowel epenthesis might be assumed to proceed in two steps: an empty vocalic slot is inserted first (∅ → $V_0$) and is then filled by a vowel ($V_0$ → a). If glottal-stop epenthesis is assumed to only apply after empty vocalic slots ( / $V_0$ __ ) before the filling step, it would correctly apply only after epenthetic vowels.

effect involving deletion was proposed by Inkelas & Orgun (1995) for Turkish intervocalic velar deletion in derived environments (/sokak-a/ → [soka-a]). On their account, stem-final velars are exempt from syllabification, and velars syllabified before adding the suffix are later immune from deletion. The range of possible patterns of deletion in derived environments under the proposed theory will depend on the range of similar structural-immunity moves that the theory is allowed to make. I will leave the exploration of those possibilities for future work, and return now to the Finnish CG case.

Given the Finnish data we have seen so far, the cyclic variant of the theory would succeed as well without MSCs. In particular, non-application of CG in (47) is not yet a problem for the cyclic variant, since the suffix -tten might be added only after CG gets its last chance to apply. The crucial example that causes a problem for the cyclic variant is the contrast between (52a), where CG applies optionally, and (52b), where it applies obligatorily:

(52) a. /ulla**kk**o-i-hin/ → [ul.la.**k**oi.hin] ∼ [ul.la<u>k</u>.<u>k</u>oi.hin] 'attic-PL.ESS.'
b. /ulla**kk**o-i-tten/ → [ul.la.**k**oit.ten] ∼ *[ul.la<u>k</u>.<u>k</u>oit.ten] 'attic-PL.GEN.'

This contrast suggests that the suffix -tten itself creates an environment for the application of CG to a preceding geminate (/kk/). In (52a), with a suffix that begins with a non-geminate, CG applies optionally. On one account, optional application is due to two available structures for [oi], one of which triggers CG and the other does not (Keyser and Kiparsky 1984, Kiparsky 2003). Kiparsky (2003: 121) notes that when a diphthong like [oi] is followed by a geminate-initial suffix, CG applies obligatorily rather than optionally. That is, the geminate /tt/ of the suffix -tten plays a role in triggering CG and eliminating the ungrammatical option *[ul.la<u>k</u>.<u>k</u>oit.ten] in (52b). In order to eliminate the output with non-application in (52b) while still generating it in (52a), CG needs to apply again after -tten is added into the derivation. This leads to an ordering paradox for the cyclic theory. On the one hand, CG must be able to apply *after* the addition of -tten to make sure that /kk/ undergoes gradation. On the other hand, anti-CG could not have applied to -tten at any prior level of representation to protect it from undergoing CG, so, paradoxically, CG must not be able apply once -tten is added (otherwise, it would incorrectly apply to -tten).

The problem for the cyclic theory is that there is no level of representation in which phonological restrictions apply to suffixes in isolation from the rest of the string. Whenever the nonderived environment in -tten is present in the derivation, a derived environment (the hetero-morphemic closed syllable kko-i-t) is present as well. This is why anti-CG cannot apply to -tten without causing trouble elsewhere. MSCs address this problem directly: if anti-CG applies to individual morphemes in the lexicon before they are combined with other morphemes, it can apply to -tten before any derived environment is created.[15]

[15] A reviewer points out that this argument relies on the analysis of the allomorph of the genitive plural suffix as the mono-morphemic /-tten/. And while Kiparsky 1993 (p. 283) analyzes this suffix as mono-morphemic, in Kiparsky 2003 (p. 149) it is analyzed as the bi-morphemic /-t-ten/. Note that the possibility of a bi-morphemic analysis does not weaken the argument. If the suffix is underlyingly the mono-morphemic /-tten/, then the MSC variant of the theory provides a successful account of the blocking pattern while the cyclic variant fails. If, however, the suffix is the bi-morphemic /-t-ten/, then both variants fail. In other words, the combination of the MSC variant and the mono-morphemic analysis is the only combination that

One way out for the cyclic approach without MSCs is to mark the suffix -tten as an exception to CG. As Wolf (2008) notes, however, marking suffixes as exceptions will not work for a similar problem in Luganda, where a single suffix that behaves like -tten contains two potential inputs for the application of a process that shows NDEB. As we will now see, only one of the two inputs is included in an environment that is fully contained within the suffix, so marking the suffix as an exception will incorrectly block both inputs from undergoing the process (rather than just one).

Citing an example from Odden (1990), Wolf (2008) notes that morpheme-final consonants in Luganda undergo spirantization before /i/. Otherwise, consonants do not undergo spirantization before a tauto-morphemic /i/. This is illustrated using the example in (53), where the morpheme-final /k/ and /r/ become [s] and [z] respectively, but the first /r/ in /-irir/ does not change.

(53) /lamuk-irir-i/ → [lamus-iriz-i] 'greet without ceasing'

Since the the suffix /-irir/ triggers the application of spirantization to /k/ and fully contains an environment for spirantization itself, it poses the same problem for the cyclic approach as the Finnish suffix /-tten/: spirantization is predicted to apply within the suffix, producing the incorrect *[lamus-iziz-i]. In this case, marking /-irir/ as an exception to spirantization would incorrectly block spirantization from applying to the morpheme-final /r/, incorrectly producing *[lamus-irir-i]. The cyclic theory will be pressed to mark as exceptions precisely those suffix-internal inputs to spirantization that are a part of underlying morpheme-internal spirantization environments.

In sum, Finnish, Luganda, and the similar case of Meskwaki (Wier 2004) not discussed here are surprising for the cyclic variant of the present proposal, which predicts blocking within suffixes to be impossible. Those three languages at least suggest that blocking within suffixes is attested, and they receive a straightforward account by the MSC theory.

## 3.3 Puzzle III: Non-contrastive triggers

As mentioned in the introduction, the MSC theory follows the characterization of NDEB in (8), repeated here:

(54) *A characterization of NDEB*
$P$ is blocked in some environment when this environment is present in the UR of some morpheme.

Other theories of NDEB formalize the same characterization but link presence in the UR to blocking within architectures that adopt ROTB and thus reject MSCs. An example is Burzio's (2000) Sequential Faithfulness, a Parallel-OT theory where a faithfulness constraint can distinguish underlying tauto-morphemic from hetero-morphemic sequences. For example, to account for NDEB as in Finnish' assibilation, a faithfulness constraint would protect underlying tauto-morphemic but not hetero-morphemic

works. Since there was no empirical reason for Kiparsky's switch from /-tten/ to /-t-ten/ (Paul Kiparsky, personal communication), the argument for the MSC variant stands, and can be taken as evidence for a mono-morphemic analysis of the suffix.

/ti/ sequences from undergoing assibilation. Other theories that follow the same characterization (though in different ways) include those of Inkelas (2000) and Kula (2008).

The reliance of the characterization in (54) on presence in the UR faces a problem when the application of $P$ is triggered by derived non-contrastive phonological material, but only when (54) is combined with ROTB. Reported cases of NDEB with non-contrastive triggers are vowel raising in Romanian (Steriade 2008a) and vowel reduction in Armenian (Khanjian 2009). In both languages, the application of $P$ depends on the position of stress, but the distribution of stress is predictable (at least in some environments) and its surface position is determined by the grammar. Given ROTB, underlying stress can be generated liberally, creating URs where environments that should count as derived are underlyingly present or vice versa. As we will see, this results in a *Too-Many-URs* problem for theories following (54), because the abundance of URs causes the theory to fail to separate derived from nonderived environments. This problem does not arise in the MSC theory, which can filter the problematic URs using MSCs.

In the remainder of this section I describe the Romanian case from Steriade 2008a and show how the MSC theory accounts for it. I return to concrete theories that adopt (54) together with ROTB in section 4 and show that they fail to correctly describe the Romanian pattern following the reasoning just described. Section 4 also discusses Wolf's (2008) theory, which adopts a combination of ROTB with a variant of (54) that faces the same problem in accounting for the Romanian pattern.

In Romanian, unstressed /a/ obligatorily raises to [ə] in suffixed forms (55), but not if [a] is also unstressed in the unsuffixed form (56). Intuitively, /a/ raises only when its stresslessness is derived, in a sense that will become more precise below.

(55) Raising

| | | | | | |
|---|---|---|---|---|---|
| a. | bárbə | 'beard' | bərb-ós | *barb-ós | 'bearded-MASC' |
| b. | fáur | 'artisan' | fəur-í́ | *faur-í́ | 'to fashion' |
| c. | isprávə | 'brave deed' | isprav-nik | *ispr*ə́*v-nik | (nobleman's title) |

(56) No raising

| | | | | | |
|---|---|---|---|---|---|
| a. | mazí́l | 'deposed official' | mazil-í́ | *məzil-í́ | 'depose' |
| b. | kartóf | 'potato' | kartof-jór | *kərtof-jór | 'potato-DIM' |

Romanian stress largely obeys the following generalization: it falls on the final syllable if heavy and otherwise falls on the penultimate syllable (Steriade 1984, 2021; Chitoran 2002, p. 76). There are various exceptions to this pattern such as lexical items with antepenultimate stress, affixes with alternative stress requirements, further complications in verbs, and others, but the basic generalization has been reliably identified as the default stress pattern. Exceptions aside, this means that stress is not stored in the Romanian lexicon but is rather predictably derived by the grammar.

Stress in Romanian serves the role of the non-contrastive trigger discussed above: /a/ raises when unstressed, and the distribution of stress in stems is determined by the grammar. Without restricting the distribution of stress in URs, environments that must count as derived can be written into URs, leading to incorrect non-applications of vowel raising in such environments, as we will now see.

For the MSC theory, an account of blocking would require the following ingredients. First, a variant of /a/ that is underspecified for the feature [low] would be referred to as /A/. Raising would be stated as in (57) and anti-raising as in (58).

(57) Raising: $\text{A}_{\text{[-stress]}} \rightarrow$ ə

(58) Anti-raising: $\text{A}_{\text{[-stress]}} \rightarrow$ a

If we follow the same recipe as in previous sections, the basic grammar would be (59), with a cover stress rule preceding raising in the phonology. The cover stress rule assigns stress to the final syllable if heavy and otherwise to the penultimate syllable.[16]

(59) Grammar for Romanian raising (to be revised below)

a. Morpheme structure component:
   - $\text{a} \notin \Sigma_L$
   - $\text{A}_{\text{[-stress]}} \rightarrow$ a

b. Phonological rules:
   - STRESS
   - $\text{A}_{\text{[-stress]}} \rightarrow$ ə
   - A → a

The problem with (59) (before introducing additional MSCs) is that anti-raising refers to stress but the correct position of stress is determined by the grammar later. Anti-raising cannot make the necessary distinction between stressed and unstressed vowels and thus fails to capture the contrast in (60). For example, the grammar can generate the UR /barbə/, with an unstressed first vowel that has undergone anti-raising. From this UR, the grammar incorrectly derives *[barb-ós], with underapplication of raising in an environment that must count as derived.

| | | | | | | |
|---|---|---|---|---|---|---|
| (60) | a. | bárbə | 'beard' | bərb-ós | *barb-ós | 'bearded-MASC' |
| | b. | mazíl | 'deposed official' | mazil-í | *məzil-í | 'depose' |

MSCs can solve this problem by filtering out the problematic URs. In particular, (59) can be minimally modified so as to enforce the correct position of stem-level stress in the morpheme structure component. As before, if stress is assigned in the morpheme structure component, then every morpheme will be subject to stress assignment before morpheme combination.[17] On the assumption that different suffixes show idiosyncratic

[16]Exceptions to the default stress pattern in stems will be ignored in the analysis. There has been some debate about how such exceptions should be derived, with two prominent alternatives being extrametricality (Steriade 1984) and lexical listing or constraint indexation (Chitoran 2002; Steriade 2021). The argument here only relies on there being a default stress pattern and does not depend on the choice of analysis for exceptions, as far as I can tell.

[17]An alternative to assigning stress to every morpheme in the lexicon is to relegate the MSCs, including anti-raising and stress assignment, to a cyclic stem-level phonology, where they would apply once before suffixation. In the absence of empirical evidence for one option over the other, I will choose the MSC option to keep the analysis more similar to the analyses in the previous sections. Note that even though an alternative analysis without MSCs is available in the Romanian case, we will see below that avoiding MSCs is not an option within certain alternative theories of NDEB, which would still require MSCs to filter out the problematic URs. Thus, the importance of the Romanian case study is that it eliminates some alternatives to the MSC theory, contributing to the cumulative argument of the paper in favor of MSCs.

behavior with respect to stress, I will tentatively assume that suffixes are diacritically marked as to whether they are lexically stressed or resist stress assignment. In situations where two stresses arise through morpheme concatenation (as a result of combining a stressed stem with a stressed suffix), the cover stress rule deletes the first stress. Here is the final grammar, followed by derivations of the forms in (60) (for simplicity, I omit vowel-deletion rules from the grammar and do not indicate their application):

(61) Grammar for Romanian raising (final)

a. Morpheme structure component:
- $\text{a}, \acute{\sigma} \notin \Sigma_L$
- STRESS
- $\text{A}_{\text{[-stress]}} \rightarrow \text{a}$

b. Phonological rules:
- STRESS
- $\text{A}_{\text{[-stress]}} \rightarrow \text{ə}$
- $\text{A} \rightarrow \text{a}$

(62) Derivation of [bárbə]

a. Morpheme structure rules apply:

1. {bArbə} → /bÁrbə/

b. Phonological rules apply:

| UR | /bÁrbə/ |
|---|---|
| STRESS | - |
| $\text{A}_{\text{[-stress]}} \rightarrow \text{ə}$ | - |
| A → a | bárbə |
| SR | [bárbə] |

(63) Derivation of [bərb-ós] (unsuffixed form: [bárbə])

a. Morpheme structure rules apply:

1. {bArbə} → /bÁrbə/
2. {os} → /ós/

b. Phonological rules apply:

| UR | /bÁrbə-ós/ |
|---|---|
| STRESS | bArbós |
| $\text{A}_{\text{[-stress]}} \rightarrow \text{ə}$ | bərbós |
| A → a | - |
| SR | [bərbós] |

(64) Derivation of [mazíl]

a. Morpheme structure rules apply:[18]

[18] As mentioned in section 6, Romanian has a process of u-deletion (Steriade 1984; Chitoran 2002, section 5.3), which seems to also be operative here. There is distributional evidence for an underlying /u/ in [mazíl] and other consonant-final nouns: this /u/ surfaces before some suffixes, i.e., in some environments where it is

1. {mAzilu} → /mazílu/

b. Phonological processes apply:

| UR | /mazílu/ |
|---|---|
| STRESS | - |
| $A_{[\text{-stress}]}$ → ə | - |
| A → a | - |
| SR | [mazíl] |

(65) Derivation of [mazil-í] (unsuffixed form: [mazíl])

a. Morpheme structure rules apply:

1. {mAzilu} → /mazílu/
2. {í} → /í/

b. Phonological processes apply:

| UR | /mazílu-í/ |
|---|---|
| STRESS | mazilí |
| $A_{[\text{-stress}]}$ → ə | - |
| A → a | - |
| SR | [mazilí] |

# 4 Previous theories

In this section I use the three puzzles from the previous section to take a critical look at alternative theories of NDEB. Competing theories have been proposed by Mascaró (1976), Kiparsky (1993), Burzio (2000) Inkelas (2000), Łubowicz (2002), McCarthy (2003), van Oostendorp (2007), Kula (2008), Wolf (2008), and Anttila (2009), among others. As the literature on NDEB is quite vast, I will not be able to do justice to all of the relevant theories. Instead, I will discuss what I take to be a representative sample of the literature, focusing on theories that are either still pursued or have been abandoned without argument, and refer the reader to critical reviews of theories not discussed here directly. See, in particular, Kiparsky (1993) for a review of the literature prior to 1993, Inkelas (2000) for a review of the early OT literature (1993-2000), and Wolf (2008) for a more comprehensive critical review of the literature, including the literature following Inkelas (2000). As we will see, none of the theories I discuss are able to account for all three puzzles.

Note that theories not discussed below do not necessarily fail on any of the puzzles. As an example, I believe that Comparative Markedness (McCarthy 2003) succeeds on all three puzzles. However, differently from the theories I discuss here in detail, Comparative Markedness has been argued against in the literature (see, in particular,

not word-final; singular nouns may only end in a surface [u] when this [u] follows an otherwise impermissible complex coda (as in the word [metru] 'meter'), suggesting that deletion does not apply in these cases; and except for those singular nouns that end in a consonant and show a surface [u] elsewhere, singular nouns must end in a vowel. I will take these observations to mean that in consonant-final singular nouns that have undergone vowel deletion, final stress can be analyzed as penultimate stress that is assigned before deletion, e.g., both [bárbə] and [mazíl], which differ with respect to the position of stress on the surface, receive penultimate stress at an abstract level of representation.

an argument from Russian in Blumenfeld 2003 and an argument from Ndjébbana in Wolf 2008), and I have nothing further to contribute to its evaluation.

## 4.1 Derived-environment theories

### 4.1.1 Strict cycle condition

Mascaró (1976) argues that NDEB provides evidence for a phonological analog of Chomsky's Strict Cycle Condition (SCC; Chomsky, 1973). The phonological version is given in (66).[19]

(66) **Strict Cycle Condition.** For a cyclic rule $R$ to apply properly in any given cycle $j$, it must make specific use of information proper to (i.e., introduced by virtue of) cycle $j$.
This situation obtains if either of the following conditions is met:

1. The rule makes crucial reference to the information in the representation that spans the boundary between the current cycle and the preceding one.
2. The rule applies solely within the domain of the previous cycle but crucially refers to information supplied by a rule operating on the current cycle.

Application of a cyclic rule is licensed in morphologically-derived environments by the first condition and in phonologically-derived environments by the second condition. The table in (67) illustrates the analysis of Finnish' assibilation using the SCC. Final-vowel raising and assibilation are both assumed to be cyclic rules, and cyclic rules cannot apply in the first cycle by stipulation. A word boundary is inserted in the final cycle. The leftmost column demonstrates application in a morphologically-derived environment, the middle column application in a phonologically-derived environment, and the rightmost column blocking in a nonderived environment. The number of the SCC condition that licenses each rule application is given in brackets next to the outcome of the rule.

(67) Finnish' assibilation using the SCC

| First cycle | [lut] | [vete] | [tila] |
|---|---|---|---|
| e → i / __ # | - | - | - |
| t → s / __ i | - | - | blocked |
| Second cycle | [#[lut]i#] | [#[vete]#] | [#[tila]#] |
| e → i / __ # | - | veti (1) | - |
| t → s / __ i | lusi (1) | vesi (2) | blocked |
| | [lusi] | [vesi] | [tila] |

Hualde (1989) and Kiparsky (1993) have argued against the SCC as a theory of NDEB on the grounds that there is no correlation between the cyclicity of a rule and whether it shows NDEB effects. The argument I discuss here, which is based on persistent

[19] The presentation of the SCC in this section follows the presentation in Kenstowicz 1994, pp. 208–209, which in turn is based on Halle (1979).

blocking (Puzzle I), is different in nature. It suggests that the SCC fails to properly characterize the behavior of NDEB processes regardless of their cyclicity status. In particular, condition (1) of the SCC dictates that spanning a morpheme boundary is a sufficient property for licensing. This condition fails to predict persistent blocking: it wrongly predicts obligatory application of Romanian palatalization after vowel deletion, as shown in (68).

(68) Prediction for Romanian [pʌduk-j] (incorrect):
/pʌduke-i/ → *[pʌdutʃ-j]

### 4.1.2 Coloured Containment

van Oostendorp (2007) proposes an account of NDEB that makes use of a mechanism of morpheme indexing called "colouring". The assumption is that every morpheme is annotated with its own "color" – a morpheme-specific index which is distributed over all segments and other material (features, moras, etc) which make up the morpheme. For example, in the representation of Finnish' /timat-i/, the first morpheme would be associated with the color $\alpha$ and the second morpheme with the color $\beta$, as shown in (69) using a simplified linear representation.

(69) /$t_\alpha i_\alpha l_\alpha a_\alpha t_\alpha$-$i_\beta$/

Blocking in nonderived environments arises from a proposed constraint against monochromatic feature spreading, which I have simplified using the following statement (see the original paper for more details about the mechanics of colouring and spreading):

(70) Do not associate a feature and a segment of the same colour.

Finnish' assibilation would presumably involve spreading of the feature [continuant] from /i/ to /t/, but only if /i/ and /t/ are not of the same color:

(71) a. $t_\alpha i_\alpha \rightarrow t_\alpha i_\alpha$
b. $t_\alpha i_\beta \rightarrow s_\alpha i_\beta$

This account makes the right prediction that /timat-i/ should become [timas-i] in Finnish', but it does not predict persistent blocking (Puzzle I) and thus fails on Romanian palatalization. This is demonstrated in (72), where spreading is incorrectly licensed across a morpheme boundary.

(72) Prediction for Romanian [pʌduk-j] (incorrect):
/$p_\alpha ʌ_\alpha d_\alpha u_\alpha k_\alpha e_\alpha$-$i_\beta$/ → $p_\alpha ʌ_\alpha d_\alpha u_\alpha k_\alpha$-$i_\beta$ → *[pʌdutʃ-j]

## 4.2 Burzio (2000): Sequential faithfulness

The intuition behind Burzio's 2000 approach to NDEB is that sequences that are present in the UR of a single morpheme have a privileged status in terms of faithfulness compared to other sequences. He proposes a new type of faithfulness constraints to account for NDEB. As opposed to traditional faithfulness constraints which typically protect individual features, Burzio's constraints penalize modifications of sequences or combinations of features. I will refer to these constraints as Sequential Faithfulness constraints. An example of a Sequential Faithfulness constraint is FAITH[ti], which penalizes any output deviation from the input sequence /ti/. Burzio assumes that such constraints do not protect sequences that are separated by morpheme boundaries, presumably because morphemes are not concatenated in the input. This assumption creates a distinction between the two /ti/ sequences in the Finnish' /timat-i/: modifying the first sequence (for instance, by applying assibilation) would incur a violation of FAITH[ti], but modifying the second sequence will not. The following tableau shows how Sequential Faithfulness successfully accounts for the derivation /timat-i/ → [timasi]:

(73) Tableau for [timasi]

| | | /timat-i/ | FAITH[ti] | *ti | ID[cont] |
|---|---|---|---|---|---|
| a. | | timat-i | | **! | |
| b. | | silat-i | *! | * | * |
| c. | | silas-i | *! | | ** |
| d. | ☞ | timas-i | | * | * |

Non-contrastive triggers (Puzzle III), as in Romanian vowel raising, pose a Too-Many-URs problem for Sequential Faithfulness. Repeating the reasoning from section 3.3, the problem comes from ROTB, which generates URs with environments that trigger vowel raising and must count as derived, in this case unstressed low vowels. Since such environments are present in the UR, a Sequential Faithfulness constraint protects them and incorrectly blocks raising. The following simplified tableau demonstrates the (correct) derivation of Romanian [bərb-ós], assuming the UR /bárbə/ for the stem. In this tableau, the constraint FAITH[$a_{[\text{-stress}]}$] is not violated since /a/ is stressed in the input.

(74) Tableau for [bərb-ós], UR: /bárbə-ós/

| | | /bárbə-ós/ | FAITH[$a_{[\text{-stress}]}$] | *$a_{[\text{-stress}]}$ | ID[low] |
|---|---|---|---|---|---|
| a. | | barb-ós | | *! | |
| b. | ☞ | bərb-ós | | | * |

Given ROTB, other possible URs for the stem are /barbə/ and /barbə́/, with an underlying unstressed /a/. For these URs, vowel raising incurs a violation of FAITH[$a_{[\text{-stress}]}$], since /a/ is underlyingly unstressed. The result is that raising is incorrectly blocked in the suffixed form:

(75) Tableau for the desired [bərb-ós], UR: /barbə/

| | | /barbə-ós/ | FAITH[$a_{[\text{-stress}]}$] | *$a_{[\text{-stress}]}$ | ID[low] |
|---|---|---|---|---|---|
| a. | ☞ | barb-ós | | * | |
| b. | ☹ | bərb-ós | *! | | * |

(76) Tableau for the desired [bərb-ós], UR: /barbə́/

| | | /barbə́-ós/ | FAITH[$a_{[\text{-stress}]}$] | *$a_{[\text{-stress}]}$ | ID[low] |
|---|---|---|---|---|---|
| a. | ☞ | barb-ós | | * | |
| b. | ☹ | bərb-ós | *! | | * |

Sequential Faithfulness, then, generates both the grammatical [bərb-ós] and the ungrammatical *[barb-ós] as the suffixed forms of [bárbə]. More generally, for every hypothetical surface form in which vowel raising applies, the grammar also generates a variant in which raising does not apply. We can conclude that if $P$ has a non-contrastive trigger, Sequential Faithfulness requires $P$ to apply as an optional process, contrary to what we find in Romanian, where raising is obligatory.

## 4.3 Wolf (2008): Optimal Interleaving with Candidate Chains

Wolf's (2008) architecture is a cyclic implementation of Optimality Theory with Candidate Chains, a serial variant of OT (OT-CC; McCarthy, 2007). In this theory, a surface form is constructed in several atomic steps, such as changing a feature, adding a suffix, etc. Since the details of the theory are complex, it will be useful to first consider an informal version of Wolf's approach to NDEB, focusing on morphologically-derived environments. The intuition guiding Wolf's approach is the following:

(77) $P$ is blocked in some environment when this environment is present in the derivation both before and after suffixation.

For example, to derive [timasi] from /timat-i/ in Finnish', we need to ask, for each potential environment for assibilation in the suffixed form, whether that environment is also present at some point in the derivation before suffixation. When the answer is yes, assibilation is blocked:

(78) Prediction for /timat-i/ (correct)

| Environment | Env. present before suffixation? | Assi. blocked? | Output |
|---|---|---|---|
| /**ti**mat-i/ | Yes: **ti**mat | Yes | **ti** |
| /tima**t-i**/ | No | No | **si** |

⇒ [timas-i]

In Wolf's theory, morphemes are added to the derivation cyclically. Significantly, the UR of the stem is inserted before suffixation. Non-contrastive triggers (Puzzle III), as in Romanian vowel raising, pose a Too-Many-URs problem for Wolf's theory following the very same reasoning as before. ROTB generates URs for the stem in which the environment for raising is present as well as URs in which it is not. The former lead to an incorrect blocking of raising. Consider again the word [bərb-ós] and the UR /barbə/. For this UR, raising is incorrectly blocked by Wolf's theory (the theory makes the same incorrect prediction for the UR /barbə́/):

(79) Prediction for /barb-ós/ (incorrect)

| Environment | Env. present before suffixation? | Assi. blocked? | Output |
|---|---|---|---|
| /b**a**rbə-ós/ | Yes (UR of the stem): b**a**rbə | Yes | **a** |

⇒ *[barb-ós]

The remainder of this section repeats the problem just discussed for Wolf's theory, considering the formal details of the theory and showing more precisely how the problem arises.

In Wolf's version of OT-CC, each candidate consists of a chain that starts with a morphosyntactic input, such as /ROOT-AF/, and ends with the candidate output. Each link in the chain must differ from the preceding link by one atomic change. For example, the candidate [timasi] in Finnish' would be represented using the chain in (80). The first step in the chain is the insertion of the exponent of the root, which is followed by insertion of the suffix, and then assibilation.

(80) Candidate chain for [timasi]
<ROOT-AF, timat-AF, timati, timasi>

Occasionally, more than one chain of atomic changes leads to the same output. In such cases, an output candidate consists of a collection of all relevant chains. For example, the candidate [simasi] consists of two chains, one where assibilation of the first consonant of the word applies after suffix insertion (81a), and one where it applies before suffix insertion (81b).[20]

(81) Representation of the candidate [simasi] as a collection of chains
a. <timat-AF, timati, simati, simasi>
b. <timat-AF, simat-AF, simati, simasi>

PRECEDENCE constraints are special constraints within OT-CC that impose restrictions on the order of operations. They do so by penalizing the order of violations of faithfulness constraints incurred by changes between members of a candidate chain. For example, the constraint PREC(INSERT-AFFIX,IDENT[cont]) penalizes a violation of the faithfulness constraint IDENT[cont]) (incurred by assibilation), in case that violation is not crucially preceded by insertion of an affix (that is, by violation of the faithfulness constraint INSERT-AFFIX, which penalizes affix insertions). This PREC constraint is the key component of Wolf's theory of NDEB. In [simasi], assibilation of the first segment incurs a violation of IDENT[cont] that is not crucially preceded by violation of INSERT-AFFIX: IDENT[cont] is violated (considering just the first segment) after affix insertion in (81a) but before affix insertion in (81b). The result is that [simasi] violates the PREC contsraint.[21] In contrast, for the candidate [timasi], the violation of IDENT[cont] can

[20]The theory prohibits other conceivable chains in which the stem-final /t/ becomes /s/ prior to suffixation, on the assumption that each link in the chain must be more harmonic than the preceding link. Since the environment for assibilation is not met before suffixation, a change from /timat/ to /timas/ would incur a violation of faithfulness without improving on any high-ranked markedness constraint. Such chains are therefore blocked.

[21]To ensure that multiple applications of assibilation are distinguished from one another, precedence is not directly evaluated on the candidates themselves, but rather on tuples of faithfulness violations that the candidates induce, called LUMSeqs (1), and each violation is indexed with respect to the position in the word which is the source of the violation. In both LUMSeqs for candidate (b), the violation $\text{ID[cont]}_{@5}$ (which corresponds to the application of assibilation to the fifth segment of the word) follows INSERT-AF (which corresponds to affixation), which means that this instance of assibilation is crucially preceded by affixation

only happen after affix insertion, because the environment of assibilation is not present earlier.

The grammar of Finnish' would have the three constraints in (82) with the ranking in (83). According to this ranking, *ti triggers assibilation, but only when the higher ranked precedence constraint is satisfied.

(82) a. *ti

b. IDENT[cont]

c. PREC(INSERT-AFFIX,IDENT[cont]): assign a violation mark for each time that:

- A process that violates IDENT[cont] applies without having been preceded by a process that violates INSERT-AFFIX
- A process that violates IDENT[cont] applies and is followed by a process that violates INSERT-AFFIX

(83) PREC(INSERT-AFFIX,IDENT[cont]) $\gg$ *ti $\gg$ IDENT[cont]

The following tableau shows how the theory correctly derives [timasi] (the first link in each chain, ROOT-AF, is omitted for reasons of space). Since candidate (b) violates the highest ranked PREC constraint and candidate (a) does not, candidate (a) is the winner.

(84) Tableau for [timasi]

| | | /ROOT-AF/ | PREC(INSERT-AF,ID[cont]) | *ti | ID[cont] |
|---|---|---|---|---|---|
| a. | ☞ | <timat-AF, timati, timasi> | | * | * |
| b. | | <timat-AF, timati, simati, simasi><br><timat-AF, simat-AF, simati, simasi> | *! | | ** |

Here is how an analysis of Romanian raising would work in this architecture. A straightforward ranking, given in (85), is of the following constraints: a cover constraint STRESS stands for whatever constraints enforce correct surface stress in Romanian; the constraint PREC(INSERT-AFFIX,IDENT[low]) requires that raising be crucially preceded by affixation; and the markedness constraint $*a_{[\text{-stress}]}$ is responsible for triggering raising, in violation of the faithfulness constraint IDENT[low].

(85) STRESS , PREC(INSERT-AFFIX,IDENT[low]) $\gg$ $*a_{[\text{-stress}]}$ $\gg$ IDENT[low]

The tableau in (86) demonstrates the correct derivation of [bərb-ós] assuming the UR /bárbə/ for the root (for simplicity, the tableau ignores the deletion of stem-final /ə/). Notice that, crucially, stress is underlyingly penultimate: /a/ is stressed from the outset, so raising cannot apply before suffixation and there is no PREC violation. Hence, raising is (correctly) not blocked.

---

and so does not incur a PREC violation.

(1) a. LUMSeq: <INSERT-AF, $\text{ID[cont]}_{@5}$>

b. LUMSeq: <INSERT-AF, $\text{ID[cont]}_{@1}$, $\text{ID[cont]}_{@5}$>
LUMSeq: < $\text{ID[cont]}_{@1}$, INSERT-AF, $\text{ID[cont]}_{@5}$>

(86) Correct derivation of [bərb-ós] (assuming the UR /bárbə/)

| | | /ROOT-AF/ | STRESS | PREC(AF,IDENT[low]) | $*a_{[\text{-stress}]}$ | ID[low] |
|---|---|---|---|---|---|---|
| a. | ☞ | <bárbə-AF, bárbə-ós, barbə-ós, bərbə-ós> | | | | * |
| b. | | <bárbə-AF, bárbə-ós, barbə-ós> | | | *! | |

Given ROTB, other URs are possible in which stress does not occur in penultimate position, such as /barbə́/ as /barbə/. Markedness constraints can still enforce the correct penultimate stress on the surface. The problem is that for such URs, the environment for raising is met before or after suffixation, so raising is not crucially preceded by suffixation. Therefore, raising is incorrectly blocked (87a).

(87) a. /barbə́-ós/ → *[barb-ós]

b. /barbə-ós/ → *[barb-ós]

The tableau in (88) is a concrete tableau for (87a), where stress is underlyingly final. As with Burzio's theory, the conclusion is that ROTB generates too many URs, which prevent Wolf's theory from representing obligatory NDEB processes with non-contrastive triggers like Romanian vowel raising.

(88) Incorrect derivation of *[barb-ós] (assuming the UR /barbə́/):

| | | /ROOT-AF/ | STRESS | PREC(AF,IDENT[low]) | $*a_{[\text{-stress}]}$ | ID[low] |
|---|---|---|---|---|---|---|
| a. | ☹ | <barbə́-AF, bárbə-ós, barbə-ós, bərbə-ós><br><barbə́-AF, bərbə́-AF, bərbə́-ós, bərbə-ós> | | *! | | * |
| b. | ☞ | <barbə́-AF, barbə́-ós, barbə-ós> | | | * | |

# 5 Conclusion

The three puzzles discussed in this paper collectively pose a challenge to previous theories of NDEB. The proposed MSC theory was able to deal with all three puzzles. This result is summarized again in the following table.

(89) Summary of the discussion in sections 3 and 4

| Theory ↓ Puzzle → | Persistent blocking | Blocking within suffixes | Non-contrastive trigger |
|---|---|---|---|
| MSC theory of NDEB | ✓ | ✓ | ✓ |
| Strict Cycle Condition | ✗ | | |
| Coloured Containment | ✗ | | |
| Cyclic variant | | ✗ | |
| Sequential Faithfulness | | | ✗ |
| OT with Candidate Chains | | | ✗ |

The proposed theory adopted various assumptions that were not necessary for a successful account of the three puzzles, such as a rule-based formalism and underspecification. Rather, the success of the theory was due to two properties of MSCs. The ability of MSCs to capture generalizations over single morphemes in the lexicon enabled the

MSC theory to deal with the puzzles of persistent blocking and blocking within suffixes. The ability of MSCs to filter URs with non-contrastive phonological features played a central role in the account of blocked processes that have non-contrastive triggers. Without these properties, previous theories of NDEB were not able to generate existing blocking patterns.

The MSC theory did not rely on any dedicated grammatical mechanisms to account for NDEB. Instead, according to this theory, NDEB is an epiphenomenon of a somewhat arbitrary interaction between MSCs and the mapping from URs to surface forms. In this respect, the theory departs from the intuition that NDEB reflects some fundamental architectural property, such as the Strict Cycle Condition from Mascaró's early work on NDEB or Wolf's PREC constraints. Of course, the idea that NDEB is epiphenomenal is not new: it was proposed by Kiparsky (1993), which is the direct predecessor of the present proposal, and was adopted in later theories of NDEB such as those of Inkelas (2000), Burzio (2000), and Kula (2008). Whether NDEB is the result of dedicated grammatical mechanisms is an empirical question, and this paper has shown that previous theories of NDEB fail to generate existing blocking patterns whether they adopt such mechanisms or not. However, this paper has not shown that the mere existence of mechanisms dedicated to NDEB is necessarily the reason for the failure. To see why, one can imagine a rather complicated combination of the MSC architecture with Wolf's (2008) OT-CC theory of NDEB: Wolf's PREC constraints could be responsible for most blocking in a non-epiphenomenal way, and MSCs could filter the URs that posed a Too-Many-URs problem for the OT-CC theory. In addition, while a principle like the Strict Cycle Condition did not provide a successful account of every attested NDEB pattern, it can still co-exist in principle with mechanisms that do.

The MSC theory raises multiple important questions that this paper has left open, such as: what are the broader typological consequences of allowing language-specific MSCs – in particular, anti-rules – to interact with phonological processes? How would a constraint-based implementation of the architecture differ from a rule-based implementation in terms of its empirical predictions? How can the MSC theory be distinguished from a cyclic theory that allows phonological processes to apply to isolated morphemes (along the lines of the affix cycles of Borowsky 1993, Baker 2005, and Bermúdez-Otero 2018)? Can theories of phonological learning that have been shown to acquire simple MSCs (Rasin and Katzir 2016, Rasin and Katzir 2020) also acquire the MSCs proposed here to account for NDEB?

Answering these questions could strengthen or weaken the main conclusion of the paper: that capturing generalizations over the URs of single morphemes is needed for a successful account of NDEB. The consequences of answering these questions for phonological architecture would therefore be significant: the dual-component architecture of early generative phonology was replaced by architectures that adopt ROTB, but it has been hard to find distinguishing empirical evidence. If this paper is on the right track, then the phenomenon of NDEB provides novel empirical evidence in favor of a dual-component architecture of phonology and against the principle of ROTB.

# References


Anttila, Arto. 2006. Variation and opacity. *Natural Language and Linguistic Theory* 24:893–944.

Anttila, Arto. 2009. Derived environment effects in colloquial Helsinki Finnish. In *The nature of the word: Essays in honor of paul kiparsky*, ed. Kristin Hanson and Sharon Inkelas, 433–460. Cambridge: MIT Press.

Archangeli, Diana. 1988. Aspects of underspecification theory. *Phonology* 5:183–207.

Baker, Brett. 2005. The domain of phonological processes. In *Proceedings of the 2004 Conference of the Australian Linguistics Society*, ed. Ilana Mushin. Available online at http://hdl.handle.net/2123/112.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In *The Blackwell companion to phonology*, ed. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, volume 4, 2019–2048. Malden, MA: Wiley-Blackwell.

Bermúdez-Otero, Ricardo. 2018. Stratal phonology. In *The Routledge handbook of phonological theory*, ed. S.J. Hannahs and Anna R. K. Bosch, 100–134. Abingdon: Routledge.

Blumenfeld, Lev. 2003. Counterfeeding, derived environment effects, and comparative markedness. *Theoretical Linguistics* 29:89–99.

Booij, Geert. 2011. Morpheme structure constraints. In *The blackwell companion to phonology*, ed. Elizabeth Hume Marc van Oostendorp, Colin Ewen and Keren Rice, 2049—2069. Malden, MA: Wiley-Blackwell.

Borowsky, Toni. 1993. On the word level. In *Studies in lexical phonology (phonetics and phonology 4)*, ed. Sharon Hargus and Ellen M. Kaisse, 199–234. San Diego: Academic Press.

Burzio, Luigi. 2000. Cycles, non-derived-environment-blocking, and correspondence. In *Optimality theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer. Cambridge: Cambridge University Press.

Chitoran, Ioana. 2002. The phonology and morphology of Romanian diphthongization. *Probus* 14:205–246.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky. Holt, Rinehart and Winston, Inc.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Gouskova, Maria. to appear. MSCs in positional voicing neutralization: The problem of gapped inventories. Manuscript, NYU, New York, NY. To appear in *Phonology*.

Halle, Morris. 1959. *The sound pattern of Russian*. Walter de Gruyter.

Halle, Morris. 1962. Phonology in generative grammar. *Word* 18:54–72.

Halle, Morris. 1979. Formal vs. functional considerations in phonology. In *Studies in diachronic, synchronic, and typological linguistics: Festschrift for Oswald Szemerényi*, ed. B. Brogyanyi, 325–341. Amsterdam: Benjamins.

Halle, Morris, and Jean-Roger Vergnaud. 1987. Stress and the cycle. *Linguistic Inquiry* 18:45–84.

Hammond, Michael. 1992. Deriving the strict cycle condition. *CLS* 28:126–140.

Hualde, Jose Ignacio. 1989. The strict cycle condition and noncyclic rules. *Linguistic Inquiry* 20:675–680.

Inkelas, Sharon. 2000. Phonotactic blocking through structural immunity. In *Lexicon in focus*, ed. Barbara Stiebels and Dieter Wunderlich, Studia Grammatica, 7–40. Akademie Verlag.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Blackwell.

Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology: Description and theory*. Academic Press.

Keyser, Samuel Jay, and Paul Kiparsky. 1984. Syllable structure in Finnish phonology. In *Language sound structure*, ed. Mark Aronoff and Richard T. Oehrle, 7–31. Cambridge, Mass: MIT Press.

Khanjian, Hrayr. 2009. Stress dependent vowel reduction. In *Annual Meeting of the Berkeley Linguistics Society*, volume 35, 178–189.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the Morning Calm*, 3–91. Linguistics Society of Korea, Seoul, Korea: Hanshin Publishing Company.

Kiparsky, Paul. 1993. Blocking in nonderived environments. In *Phonetics and phonology 4: Studies in lexical phonology*, ed. Ellen M. Kaisse and Sharon Hargus, 277–313. San Diego: Academic Press.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17:351–367.

Kiparsky, Paul. 2003. Finnish noun inflection. In *Generative approaches to finnic linguistics*, 109–161. CSLI.

Kisseberth, Charles W. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1:291–306.

Kula, Nancy C. 2008. Derived environment effects: a representational approach. *Lingua* 118:1328–1343.

Łubowicz, Anna. 2002. Derived environment effects in optimality theory. *Lingua* 112:243–280.

Mascaró, Joan. 1976. Catalan phonology and the phonological cycle. Doctoral Dissertation, MIT.

McCarthy, John. 2003. Comparative Markedness. *Theoretical Linguistics* 29:1–51.

McCarthy, John. 2007. *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox.

Odden, David. 1990. Phonology and its interaction with syntax and morphology. *Studies in the Linguistic Sciences* 20.2:69–108.

van Oostendorp, M. 2007. Derived environment effects and consistency of exponence. In *Freedom of analysis*, ed. Sylvia Blaho, Patrik Bye, and Martin Krämer, 123–149. Berlin: Mouton de Gruyter.

Pesetsky, David. 1979. Russian morphology and lexical theory. Ms., MIT.

Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical report, Rutgers University, Center for Cognitive Science.

Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Wiley-Blackwell: Oxford.

Rasin, Ezer, and Roni Katzir. 2016. On evaluation metrics in Optimality Theory. *Linguistic Inquiry* 47(2):235–282.

Rasin, Ezer, and Roni Katzir. 2020. A conditional learnability argument for constraints on underlying representations. *Journal of Linguistics* 56:745–773.

Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27:720–731.

Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43:393–436.

Steriade, Donca. 1984. Glides and vowels in romanian. In *Annual Meeting of the Berkeley Linguistics Society*, volume 10, 47–64.

Steriade, Donca. 1995. Underspecification and markedness. In *The handbook of phonological theory*, ed. John Goldsmith, 114–174. Oxford: Blackwell.

Steriade, Donca. 2008a. A pseudo-cyclic effect in Romanian morphophonology. In *Inflectional identity*, ed. Asaf Bachrach and Andrew Nevins, 313–359. Oxford: Oxford University Press.

Steriade, Donca. 2008b. Unpublished lecture notes, MIT.

Steriade, Donca. 2021. Cyclicity without containment in Romanian perfects. In *Romance languages and linguistic theory 2017*, 312–334. John Benjamins.

Wier, Thomas R. 2004. Comparative markedness and opacity in Meskwaki palatalization. In *Proceedings of the Thirtieth Annual Meeting of the Berkeley Linguistics Society*, ed. Marc Ettlinger, Nicholas Fleisher, and Mischa Park-Doob, 426–434. Berkeley: BLS.

Wolf, Matthew. 2008. Optimal interleaving: serial phonology-morphology interaction in a constraint-based model. Doctoral Dissertation, University of Massachusetts, Amherst.