

Phonological Evidence for Movement in Inflectional Morphology An Argument for Harmonic Serialism

Daniel Gleim, Gereon Müller, Mariia Privizentseva & Sören Tebay
Universität Leipzig

Abstract

Seemingly non-local phonological operations triggered by inflectional exponents have been observed in a number of languages. Focussing on *ruki* rule application in Sanskrit (Kiparsky (1982b)), de-spirantization in Barwar Aramaic (Khan (2008)), *ni*-insertion in Quechua (Myler (2013)), vowel harmony in Kazakh (Bowman & Lokshin (2014)), and accent shift in Lithuanian (Kushnir (2018)), we argue that these phenomena should be analyzed as strictly local phonological reflexes of movement in a pre-syntactic autonomous morphological component. Such morphological movement is shown to arise without further assumptions under the approach to inflectional morphology based on harmonic serialism (McCarthy (2016)) that has recently been developed in Müller (2018). Here, each morphological operation immediately triggers an optimization procedure; morphological structure-building is subject to simple alignment constraints; and counter-cyclic operations are precluded. Against this background, phonological reflexes of movement are predicted to show up when a potentially complete word triggers a first phonological cycle, which is then followed by morphological movement (and a second phonological cycle). Finally, we argue that constraint-driven morphological movement is superior to alternative accounts based on non-local phonology, interfixation (Hyman (2003), Kiparsky (1982b; 2017)), lowering, local dislocation (Embick & Noyer (2001)), or syntactic movement (Myler (2017)).

1. Introduction

There are many cases where it looks as though an inflectional exponent can trigger non-local phonological changes. In the present paper, we will consider de-spirantization apparently triggered by non-local agreement exponents in Barwar Aramaic (see Khan (2008)), Sausurean accent shift apparently triggered by non-local person/number exponents in Lithuanian (see Kushnir (2018)), *ni*-insertion apparently triggered by non-local possessive markers in Quechua (see Myler (2013)), *ruki* rule application apparently triggered on verb roots by non-local prefixes in Sanskrit (see Kiparsky (1982b)), and vowel harmony apparently applying with non-local Q morphemes in Kazakh (see Bowman & Lokshin (2014)). These kinds of phenomena have standardly been analyzed either in terms of non-local phonological operations, or by resort to counter-cyclic morphological operations like interfixation. In contrast, we will argue that phonological reflexes involving seemingly non-local morphological exponents can and should be captured in a strictly local way, by postulating movement operations in an autonomous morphological component of grammar. More specifically, we contend that morphological movement of exponents basically comes for free in the approach to inflectional morphology that is developed in Müller (2018). This approach relies on harmonic serialism (see McCarthy (2016)), a strictly derivational version of optimality theory (see Prince & Smolensky (2004)). We will show that morphological movement of exponents follows without further ado from the interaction of independently motivated constraints on structure-building (centered around morphological arrays) and linearization of morphological exponents in harmonic serialism, given (i) a certain pattern of ranking of these constraints, and (ii) the general restriction in harmonic serialism that competing outputs can differ from

lay.eggs⟨PRS⟩-PL
 ‘They lay eggs.’

lay.eggs⟨PAST⟩-3SG.F.SUBJ
 ‘She laid eggs.’

(4) shows the same alternation, albeit with an epenthetic vowel: Three-consonant (CCC) clusters must be broken up by a vowel.³ In (4-a), the vowel is inserted between first and the second members of the cluster, which leaves the adjacency of the dental fricative and the sonorant intact. This again derives the right context for de-spirantization to apply. In the feminine form for the same root, concatenation creates a four consonant cluster; cf. (4-b). In four-consonant clusters, the locus of epenthesis is no longer optional: It must be between the second and third members of the cluster, in this case between /ð/ and /n/. Here, epenthesis bleeds the application of de-spirantization.

- | | | | | |
|-----|----|---------------------------------------------|----|---------------------------------------------------|
| (4) | a. | [kawədna]
/kawðn-a/
ass-M.SG
‘ass’ | b. | [kawðənta]
/kawðn-ta/
ass-F.SG
‘she-ass’ |
|-----|----|---------------------------------------------|----|---------------------------------------------------|

The instances of de-spirantization that apply across morpheme boundaries behave exactly like the ones in roots that we have seen so far. Consider, for instance, the so-called l-suffixes, which indicate subject agreement in the perfective aspect and object agreement in imperfective environments. These exponents are triggers of de-spirantization (however, in the imperfective aspect these affixes are rarely directly adjacent to the root). In (5), the last radical of the verb roots (/θ/ and /ð/, respectively) is adjacent to the l-exponent, and accordingly undergoes de-spirantization.

- | | | | | |
|-----|----|-------------------------------------------------------------------|----|-------------------------------------------------------------------|
| (5) | a. | [tpitle]
/tp(i)θ-le/
sneeze⟨PST⟩-3SG.M.SBJ
‘He sneezed.’ | b. | [gridle]
/gr(i)ð-le/
scrub⟨PST⟩-3SG.M.SBJ
‘He scrubbed.’ |
|-----|----|-------------------------------------------------------------------|----|-------------------------------------------------------------------|

If an object agreement suffix intervenes between the root and the l-suffix, de-spirantization is transparently blocked; see (6). The intervening exponent blocks the process exactly like the vowels encoding perfectivity and imperfectivity in (3) and the epenthetic vowel in (4) do.

- (6) [tʲriðale] *[tʲridale]
 /tʲr(i)ð-a-le/
 chase.away⟨PAST⟩-F.OBJ-3SG.M.SBJ
 ‘He chased her away.’

However, there is an instance in which de-spirantization unexpectedly applies non-locally across an intervening morphological exponent. The tense/aspect exponent *wa* (‘REMOTE’) is linearized between the root and the l-suffixes. As (2) has shown for /θ/, and as (7) shows for /ð/, a root-final dental fricative is nonetheless (optionally) de-spirantized in this environment. Since *wa* intervenes, the phonological context for de-spirantization is not met on the surface. The process thus over-applies non-locally in the context of a remoteness exponent *wa*.

- (7) [tʲridwale] ~ [tʲriðwale]
 /tʲr(i)ð-wa-le/
 chase.away⟨PAST⟩-REMOTE-3SG.SBJ
 ‘He had chased away.’

³ As noted by Khan (2008, 110), the site of vowel insertion is variable.

There are no other potential triggers of de-spirantization in (2) and (7) beyond the non-local l-suffix. In particular, the operation does not otherwise take place if a dental fricative precedes a bilabial glide – neither as a general phonological process (see (8-a)), nor as a phonological process specific to the tense/aspect marker *wa* (see (8-b)).⁴

- | | | | | |
|-----|----|------------------------------------------------------------|----|--------------------------------------------------------|
| (8) | a. | [kaθwa]
/k⟨a⟩θw-a/
write⟨PRS⟩-3SG.F
‘She writes.’ | b. | [ʔiθwa]
/ʔiθ-wa/
be.there-REMOTE
‘There was.’ |
|-----|----|------------------------------------------------------------|----|--------------------------------------------------------|

2.2. Non-Local Saussurean Accent Shift and Theme Vowels in Lithuanian

A similar phenomenon shows up with accent placement in Lithuanian; see Kushnir (2018), on which the following discussion is based. Phonological words in Lithuanian have one main stress; stress can be assigned to any mora in a word. Any morphological exponent can bear an inherent lexical accent. Normally, if a stem and an inflectional exponent are both pre-specified for accent, it is the accent on the stem that prevails; i.e., the left-most accent wins (the Basic Accentuation Principle). There is a systematic exception, though, which can be referred to as (the synchronic version of) Saussure’s Law (see Halle & Vergnaud (1987)): Whenever two underlying accents coincide on two subsequent moras word-finally, the surface accent is aligned with the right edge of the word. The effect of Saussure’s Law is illustrated in (9). In (9-a), neither mora on the accusative singular exponent /a_μa_μ/ is pre-specified for accent, and main stress ends up on the stem. In (9-b), a mora of the locative singular exponent /é_μ/ does bear inherent accent, but since the two moras that are associated with stress are separated by an intervening mora that is not, Saussure’s Law does not apply, and main stress goes on the stem again. In (9-c), there is a mora pre-specified for stress on the dative plural exponent /ó_μo_μms/, but still Saussure’s Law does not spring into action because the clash does not occur word-finally (there is yet another mora further to the right). However, in (9-d), with the Saussurean instrumental singular exponent /á/, there are two word-final adjacent moras (/k/ is not moraic here), and Saussure’s Law ensures that the main stress ends up on the inflectional exponent, rather than on the stem.

- | | | | |
|-----|----|---------------------------------------------------------------------------------|------------------|
| (9) | a. | ra _μ ń _μ k + a _μ a _μ → rańk-aa | (hand-ACC.SG) |
| | b. | ká _μ i _μ m + é _μ → káim-ε | (village-LOC.SG) |
| | c. | ra _μ ń _μ k + ó _μ o _μ ms → rańk-ooms | (hand-DAT.PL) |
| | d. | ra _μ ń _μ k + á _μ → ranká | (hand-INSTR.SG.) |

The verbal agreement markers /-u/ (1.SG) and /-i/ (2.SG) are also Saussurean affixes, like the instrumental singular case marker for nominals in (9-d). However, with verbs, there is an interesting complication: Between the root and the inflectional exponent encoding ϕ -features, there is a theme vowel; cf., e.g., /a/ in (10-ab). The short theme vowel /a/ is regularly deleted in front of a following exponent starting with a vowel, as in (10-cd).

⁴ Note that there are no data in Khan’s study that would show whether de-spirantization occurs if the l-suffix is used in the present and thus encodes object (rather than subject) agreement. (i) might be expected to be a morphologically well-formed form in Barwar Aramaic. However, Khan (2008) does not give such forms for θ/δ -final roots.

- | | |
|-----|-------------------------------------------------------------------------------------|
| (i) | qəm-tʰ⟨a⟩rð-wa-le
PAST-chase.away⟨PRS⟩-REMOTE-3SG.M.OBJ
‘He chases him away.’ |
|-----|-------------------------------------------------------------------------------------|

- (10) a. kás + a + mε → kaásame (dig-1.PL.PRES)
 b. áug + a + mε → áug+a+mε (grow-1.PL.PRES)
 c. kás + [a→Ø] + ú → kasú (dig-1.SG.PRES)
 d. áug + [a→Ø] + ú → augú (grow-1.SG.PRES)

In (10-c), it may look at first sight as though the context for application of Saussure’s Law can be created by theme vowel deletion (whereas no such issue arises in (10-d), where the two accent-bearing moras are too far away from one another). However, closer inspection reveals that this cannot be the whole story. In contrast to short theme vowels, long theme vowels in Lithuanian verb inflection are not deleted in front of vowels; they are merely reduced. Thus, /ee/ becomes [ʲɛ] in front of the first and second person singular exponents, and /oo/ becomes [a]. Crucially, the Saussurean effect persists; “the accent shift takes place, disregarding the theme vowel entirely” (Kushnir (2018)). This is shown by the examples in (11-ab).

- (11) a. kás + [ee→ʲa] + ú → kasʲaú (dig-1.SG.PAST)
 b. žín + [oo→a] + ú → žinaú (know-1.SG.PAST)

Thus, just as with de-spirantization in Barwar Aramaic, Saussurean accent shift in Lithuanian verbs appears to be able to apply non-locally; given that the operation can independently be shown to depend on strict locality of the moras involved (see (9-b)), the conclusion suggests itself that intervening theme vowels behave as if they were not part of the structure at the point where the Saussurean shift takes place.

2.3. Non-Local Ni-Insertion in Quechua

Another piece of evidence which seems to suggest non-local application of a phonological operation comes from Bolivian and Huallaga Quechua (see Bills et al. (1969), Weber (1989), and Myler (2013)). In these languages, illegitimate sequences of phonological segments are repaired by epenthesis of CV. In some cases, an epenthetic syllable appears even though conditions on its insertion are not met in the surface form and two morphological exponents that would give rise to a banned sequence are separated by another exponents. The resulting configuration is of the type $Affix_1-CV-Affix_2-Affix_3$, where the sequence $Affix_1-Affix_3$ is illegitimate and triggers epenthesis, but $Affix_1-Affix_2$ is not.⁵

Let us start with conditions on epenthesis. The syllable *ni* is inserted between two morphological exponents to avoid a creation of super-heavy syllables (see Myler (2013, 191)). A syllable is super-heavy if its nucleus and coda consist of three or more moras. Short vowels and simple codas correspond to one mora; long vowels, diphthongs and complex codas correspond to two moras. *Ni*-insertion in Huallaga Quechua is illustrated in (12) and (13). If a third person possessive suffix *n* is added to a stem ending in a short vowel, the rhyme consists of two moras and epenthesis does not apply; see (12-a). If this exponent is attached to a stem ending in a consonant, as in (12-b), or a long vowel, as in (12-c), it triggers the insertion of *ni*.

- (12) a. uma-n b. mayur-ni-n c. papa:-ni-n
 head-3.POSS older-NI-3.POSS father-NI-3.POSS
 ‘his head’ ‘my older (sibling)’ ‘his father’

The same effect is caused by suffixes with the shape CCV(...); see (13).⁶ Two consonants at the beginning of an affix cannot be syllabified as the onset of a next syllable due to an

⁵ Unless indicated otherwise, all the data in this section are taken from Weber (1989).

⁶ Note that *ch* in (13) stands for a voiceless lamino-alveolar affricate (cf. Weber (1989, 452)).

additional restriction on complex onsets in Huallaga Quechua (cf. Weber (1989, 455)). This means that the first of the two consonants in the suffix has to belong to the previous syllable, which is unproblematic in (13-a). However, in (13-b), this would result in a trimoraic rhyme (consisting of **urn*) unless *ni*-insertion takes place.

- | | | | | |
|------|----|------------------|----|------------------------|
| (13) | a. | uma-nchi | b. | mayur-ni-nchi |
| | | head-1.POSS.INCL | | older-NI-3.POSS.INCL |
| | | ‘our heads’ | | ‘our sibling/siblings’ |

Weber (1976; 1989) analyzes *ni*-insertion as phonologically conditioned allomorphy of certain affixes. As an argument for this view he provides suffixes that do not follow the phonological restrictions described above. First, if the suffix *yoq* is attached to a root ending in a consonant or a long vowel, it triggers insertion of *ni* even though there is no trimoraic rhyme in these cases. Second, *ni* never appears before the verbal suffixes *yka* and *yku*, and the unacceptable sequence of segments is resolved differently. Despite these facts we continue to analyze *ni*-insertion as a regular phonological process that prevents formation of impossible syllables, because the number of morphemes with exceptional behavior seems to be smaller under this view and because it is based on independently attested restrictions on syllable structures in Quechua.⁷

Syllabic *ni* can also appear when two affixes that create the environment for its insertion are separated by the exponent *lla* ‘just’. The examples in (14) (from Myler (2013), citing Bills et al. (1969)) and (15) show this for Bolivian Quechua and Huallaga Quechua, respectively. The presence of *ni* is obligatory in Bolivian Quechua and optional in Huallaga Quechua. Note that the second person possessive used in (15) starts with two consonants /yki/, but the glide is deleted after /i/ in (15-b).

- | | | | | | |
|------|----|-----------------------|----|-------------------------|----|
| (14) | a. | *wawa-s-lla-y | b. | wawa-s-ni-lla-y | |
| | | child-PL-just-1.POSS | | child-PL-NI-just-1.POSS | |
| | | ‘just my children’ | | | |
| (15) | a. | *kikish-yki | b. | kikish-ni-ki | |
| | | armpit-2.POSS | | armpit-NI-2.POSS | |
| | | ‘your armpit’ | | | |
| | c. | *kikish-lla-ni-ki | d. | kikish-ni-lla-yki | e. |
| | | armpit-just-NI-2.POSS | | armpit-NI-just-2.POSS | |
| | | kikish-lla-yki | | | |
| | | armpit-just-2.POSS | | | |
| | | ‘just your armpit’ | | | |

Note also that /ll/ in the suffix signifies a lamino-palatal lateral, which does not cause insertion of *ni* on its own; see (16) from Huallaga Quechua.

- | | |
|------|-----------------|
| (16) | chay-yaq-lla |
| | that-LIM-just |
| | ‘just to there’ |

The intervening suffix *lla* can be in different positions. In Huallaga Quechua, it generally precedes the possessive affix; cf. (17-a). Examples where *lla* follows it are significantly degraded, but not completely ungrammatical; cf. (17-b). If a noun is marked for both plural

⁷ Adelaar & Muysken (2004:206) also analyze *ni*-insertion as a phonological process, but attribute it to resolution of otherwise illegitimate consonant clusters.

2.5. Non-Local Vowel Harmony in Kazakh

As a final example of phonological processes triggered by inflectional exponents which are non-local in surface representations, this section introduces data from non-local vowel harmony in Kazakh. We will consider the pattern of non-local application across the comitative (instrumental) suffix *men* before explaining the regular vowel harmony in further detail. Vowel harmony in Kazakh inflectional morphology usually applies locally. In (22) (from Muhamedova (2015)), the vowel of the interrogative suffix *bA* varies between front and back, depending on the backness of the preceding stem vowel.

- (22) a. Germanija-ya bar-duŋ-ba?
 Germanija-γA bar-dIŋ-bA?
 Germany-DAT go-PAST-2SG-Q
 ‘Have you been to Germany?’
- b. Ajqun yj alup ber-di-me?
 Ajqun yj al-ɪp ber-dI-bA?
 Aykhin house buy-CONV AUX.PST-3-Q
 ‘Did Aykhin buy the house?’

The instrumental case suffix *men* is an exception, since it is transparent to vowel harmony. Vowel harmony applies across this suffix, but the suffix itself does not undergo vowel harmony; see (23) (from Bowman & Lokshin (2014)).⁸ The interrogative suffix again varies, but the instrumental suffix stays invariant. Vowel harmony thus applies non-locally.

- (23) a. bʊl ʃal nan-men-ba
 bʊl ʃal nan-men-bA
 this old.man bread-INSTR-Q
 ‘Is this an old man with some bread?’
- b. bʊl ʃal bʊbek-men-be
 bʊl ʃal bʊbek-men-bA
 this old.man baby-INSTR-Q
 ‘Is this an old man with a baby?’

Vowel harmony is a regular phonological process and usually affects all suffixes. These suffix vowels assimilate in backness and rounding to a preceding root vowel (cf. Krippes (1996)). The height distinction between high and non-high vowels is kept constant; see the inventory in (24). Back rounded /o/ does not undergo vowel harmony, since it cannot occur in suffixes. Relatedly, the low vowel /a/ never assimilates in rounding. The other low vowel /æ/ mostly occurs in loanwords.

(24) Kazakh vowel phoneme inventory

	front		back	
	unrounded	rounded	unrounded	rounded
high	i	↔ y	ɯ	↔ u
mid	e	↔ ø		o
low	(æ)		a	

⁸ Muhamedova (2015) notes that for some speakers the interrogative suffix *bA* does not undergo vowel harmony at all; see also McCollum (2018) for a detailed discussion.

Vowel harmony is regular and normally applies to case suffixes. As shown in (25), the genitive case suffix *nIŋ* and the locative suffix *dA* undergo harmony; they occur with front vowels /i/ and /e/ after front vowel stems and with back vowels /u/ and /a/ after back vowels.⁹ They also allow vowel harmony to apply to their left and right. Vowel harmony is restricted to the word domain, since, e.g., the postposition *bojuu* in (25-e) does not undergo harmony. Vowel harmony is thus a regular and local phonological process in Kazakh.

- | | | | | | |
|---------|--------------------|----|---------------------|----|-------------------|
| (25) a. | kitab-tar-um-nuŋ | b. | ekew-iŋ-niŋ | | |
| | kitab-tAr-Im-nIŋ | | ekew-Iŋ-nIŋ | | |
| | book-PL-POSS-GEN | | both-POSS.2SG-GEN | | |
| | ‘of his/her books’ | | ‘of you two’ | | |
| c. | pæter-de-me? | d. | Qazaqstan-da | e. | tyni bojuu |
| | pæter-dA-bA? | | Qazaqstan-dA | | tyni bojuu |
| | flat-LOC-Q | | Kazakhstan-LOC | | night during |
| | ‘... in a flat?’ | | ‘... in Kazakhstan’ | | ‘the whole night’ |

The process applying to the question suffix after the comitative suffix shows the same alternation that we see in other cases of vowel harmony; compare (25-c) with (22-b). Similarly, this is an argument against a phonological underspecification solution with structure preservation (see van der Hulst (2016)), as /a/ generally has a harmonic counterpart. For now, we can conclude that idiosyncratic transparency poses a challenge for any purely phonological analysis of the data (cf. Mahanta (2012), Bowman & Lokshin (2014)); we will come back to this issue in section 5.

To sum up this section, we have seen five different pieces of evidence showing that a morphological exponent can be involved in a regular, well-behaved and otherwise strictly local phonological process even though it shows up in a non-local position, separated by an intervening inflectional exponent. In what follows, we will pursue what strikes us as the classical hypothesis in grammatical theory to account for paradoxes of this type: The non-local exponent is in fact a local exponent at the point where the phonological operations discussed above take place, but ends up in a non-local position as a consequence of subsequent movement. This presupposes a systematic theory of movement of morphological exponents. As we will show in the next section, the approach to inflectional morphology in terms of harmonic serialism developed in Müller (2018) predicts the existence of exponent movement in an autonomous morphological component of grammar without any further assumptions specifically designed to bring about displacement.

3. Inflectional Morphology in Harmonic Serialism

3.1. Background Assumptions

Harmonic serialism is a derivational version of optimality theory which has been envisaged as an alternative to standard parallel optimality theory from the very beginning (cf. Prince & Smolensky (2004)). In harmonic serialism, the generator (GEN) and harmony evaluation (H-EVAL) components of the grammar are not discretely ordered, but alternate constantly. Basically, the theory works as shown in (26) (see McCarthy (2008), Heck & Müller (2007)).

(26) *Harmonic serialism:*

- a. Given some input I_i , the candidate set $CS_i = \{O_{i1}, O_{i2}, \dots, O_{in}\}$ is generated by

⁹ The first example is from Bowman & Lokshin (2014), the others are from Muhamedova (2015).

- applying at most *one operation* to I_i .
- b. The output O_{ij} with the best constraint profile is selected as optimal.
 - c. O_{ij} forms the input I_{ij} for the next generation step producing a new candidate set $CS_j = \{O_{ij1}, O_{ij2}, \dots, O_{ijn}\}$.
 - d. The output O_{ijk} with the best constraint profile is selected as optimal.
 - e. Candidate set generation stops (i.e., the derivation converges) when the output of an optimization procedure is identical to the input (i.e., when the constraint profile cannot be improved anymore).

Thus, given some input, at most one basic operation can apply, which generates a finite set of output candidates. At this point, optimization selects the optimal output on the basis of a set of ranked and violable constraints. This optimal output is then subject to further application of at most one basic operation, which is immediately followed by further optimization, and so on, until the optimal output is identical to the current input, and convergence is reached. Harmonic serialism has been actively pursued over the last decade both for phonology (see, e.g., McCarthy (2010; 2016), Torres-Tamarit (2016), and Elfner (2016)) and for syntax (see Heck & Müller (2013; 2016) and Assmann et al. (2015), among others).¹⁰ In contrast, until very recently there had basically been no work in morphology.¹¹ This situation has changed with Müller (2018), where the outlines of a harmonic serialist approach to inflectional morphology are developed.

A basic assumption is that exponence in inflectional morphology involves structure-building via Merge (see Alexiadou & Müller (2008) and Bruening (2017) for earlier proposals along these lines), rather than substitution transformations applying to terminal nodes (as in Halle & Marantz (1993)), or entire subtrees (see Ackema & Neeleman (2003; 2004) and Caha (2013), among others), or spans (see Merchant (2015), Svenonius (2016), and Ermolaeva & Kobele (2019)). Such structure-building takes place in a pre-syntactic autonomous morphological component. An exponent realizes morpho-syntactic features associated with a stem by merging with it.

More specifically, morphological exponence is assumed to proceed as follows. Initially, a stem is taken from the lexicon with its inherent features, and enriched by non-inherent features, yielding a fully specified feature matrix.¹² Together, these features provide the context for underspecified inflection markers, comparable to the fully specified paradigm cell in Paradigm Function Morphology (see Stump (2001)), or a fully specified syntactic insertion context in, e.g., Distributed Morphology (see Halle & Marantz (1993)). Next, triggered by high-ranked Merge Conditions (MCs) which demand a discharge of structure-building features [$\bullet\alpha\bullet$], [$\bullet\beta\bullet$], etc., on stems, inflectional exponents of type [α], [β], ... are then successively merged with the stem, thereby eventually generating whole words.¹³ All of an inflectional exponent's features are inherent; but they are often underspecified. The

¹⁰ In syntax, the model is often referred to as 'extremely local optimization'.

¹¹ See, however, Wunderlich (1997) and Caballero & Inkelas (2013) for serial affixation in Minimalist Morphology and Optimal Construction Grammar, respectively.

¹² This corresponds to Stump's (2001) concept of a set of well-formed morpho-syntactic properties. Note that the assumption that it is stems that start morphological structure-building is a simplification. In Müller (2018, ch. 5), it is argued (in line with much recent literature) that in order to account for morpho-syntactically determined stem allomorphy, such a role should be attributed to an abstract categorizing head; the categorizing head is then equipped with the fully specified feature matrix in need of realization, and is merged with a stem that (partially) satisfies it. This complication need not concern us in the present context, though.

¹³ The underlying assumption is that the same kinds of structure-building features are also active in syntax, where they trigger syntactic Merge operations. See Adger (2003), Kobele (2006), Pesetsky & Torrego (2006), Heck & Müller (2007), Abels (2012), Georgi (2014), Collins & Stabler (2016), and many more.

morphological categories $[\alpha]$, $[\beta]$, involved here may or may not correspond directly to syntactic categories (i.e., they can be *morphomic*; Aronoff (1994)); they are determined by *morphological arrays*, i.e., sets of exponents connected by shared morpho-syntactic features. Morphological arrays are defined in (27).

(27) *Morphological Array_X*:

An exponent α is in the morphological array for a grammatical category X (MA_X) in the domain of a syntactic category W iff (i), (ii) or (iii) holds:

- (i) α realizes a grammatical category X in the domain of W by a morpho-syntactic feature that is a (possibly underspecified) instantiation of X.
- (ii) α realizes a grammatical category Y in the domain of W by a morpho-syntactic feature that is a (possibly underspecified) instantiation of Y, and there is an exponent in MA_X that realizes Y.
- (iii) α is a unique radically underspecified exponent for X in the domain of W.

In addition to MCs, there are IDENT and MAX constraints deriving the compatibility and specificity requirements for underspecified morphological exponents that have to be stipulated in approaches such as Paradigm Function Morphology (cf. Panini's Principle) and Distributed Morphology (cf. the Subset Principle). Thereby, standard underspecification-based approaches to syncretism in optimality theory can be implemented in harmonic serialism; see Grimshaw (2001), Trommer (2001), Stiebels (2006), Wolf (2008), and Müller (2011), among others, on this approach to syncretism in optimality theory. Furthermore, there are alignment (and other) constraints determining the order of exponents, primarily by demanding left-alignment ($L \leftarrow \alpha$) or right-alignment ($\alpha \Rightarrow R$) of exponents with certain types of features α (see Trommer (2001; 2008)). Finally, the fully inflected word is transferred to the syntactic component, which cannot see the internal structure of the word generated in the morphological component but can access all the morpho-syntactic features associated with the stem, and carry out Agree operations with them (cf. Chomsky (2001), Bruening (2017)).

A final important general assumption of this harmonic serialist approach to inflectional morphology is that, like structure-building in the syntax, morphological structure-building is subject to the *Strict Cycle Condition* (see Chomsky (1973; 1995; 2000; 2019), Perlmutter & Soames (1979), Freidin (1992), Pullum (1992)), according to which an exponent can only be merged with an (extended) stem at the root; cf. (28).¹⁴

(28) *Strict Cycle Condition*:

Within the current domain δ , no operation may affect solely a proper subdomain γ that is dominated by δ .

Unlike the MCs, the faithfulness constraints (IDENT, MAX), and the alignment constraints, which are part of the H-EVAL component and therefore violable and ranked, the Strict Cycle Condition is part of the GEN component, and thus inviolable – outputs that would violate it cannot be generated in the first place.

As shown in Müller (2018), an approach to inflectional morphology based on harmonic serialism offers new perspectives on some core phenomena in inflectional morphology, viz., affix order, extended exponence (see Matthews (1972) and Caballero & Harris (2012)), disjunctive blocking (i.e., compatibility/specificity effects with underspecified exponents), locality of allomorphy, and *ABA patterns (see Bobaljik (2012) on the latter). However, per-

¹⁴ A version of (28) that is exclusively concerned with structure-building operations is referred to as the Extension Condition in Chomsky (1995; 2000).

haps the most striking property of the harmonic serialist approach to inflectional morphology is that, unlike virtually all established approaches recognizing a separate morphological component of the grammar, it automatically predicts the existence of *movement* of morphological exponents in words.

3.2. *Movement in Morphology*

Here is one simple scenario (among many others) which will almost invariably give rise to movement. Suppose first that initially there is a stem A with two structure-building features [$\bullet\alpha\bullet$], [$\bullet\beta\bullet$] that trigger Merge operations with morphological exponents from the morphological arrays identified by [α] and [β], via MC_α and MC_β . Next, for the sake of simplicity, let us further assume that the two items [${}_\alpha B$] and [${}_\beta C$], respectively, are the (most likely underspecified) exponents in these morphological arrays that best satisfy the faithfulness constraints demanding compatibility and specificity. Third, there are two alignment constraints pulling the exponents in the same direction, e.g., to the right edge: $\alpha\Rightarrow R$, $\beta\Rightarrow R$. Fourth, there is a higher-ranked alignment constraint ensuring that the stem A is left-aligned: $L\Leftarrow A$. And fifth and finally, the MCs for α and β outrank the alignment constraints for α and β . Consider now what happens under this scenario (a) when the ranking among the two MCs is the opposite of the ranking among the respective alignment constraints, and (b) when the ranking among the two MCs is identical to the ranking among the respective alignment constraints, as in (29-a) and (29-b), respectively.

- (29) a. $MC(\alpha) \gg MC(\beta) \gg L\Leftarrow A \gg \beta\Rightarrow R \gg \alpha\Rightarrow R$
 b. $MC(\alpha) \gg MC(\beta) \gg L\Leftarrow A \gg \alpha\Rightarrow R \gg \beta\Rightarrow R$

Under the ranking in (29-a), [${}_\alpha B$] merges with A first, and shows up as a suffix. Then [${}_\beta C$] merges with A, also as a suffix. It will have to attach to the extended stem outside, at the root (because of the Strict Cycle Condition). This gives rise to a violation of $\alpha\Rightarrow R$, but there is no way to improve the constraint profile further by additional optimizations, so convergence is reached; see (30-a). However, things are different under the ranking in (29-b). Again, [${}_\alpha B$] merges with A first, and it shows up as a suffix. Again, [${}_\beta C$] merges with A next, also as a suffix. And as before, [${}_\beta C$] needs to attach at the root, due to the Strict Cycle Condition. Still, in this case the violation of $\alpha\Rightarrow R$ that this Merge operation creates can be undone by an optimal output in a next optimization step, by morphological movement: [${}_\alpha B$] moves around [${}_\beta C$], which trades in the more severe violation of $\alpha\Rightarrow R$ for a less severe violation of $\beta\Rightarrow R$; see (30-b).

- (30) a. [${}_A [{}_A A B_\alpha] C_\beta$] b. [${}_A [{}_A [{}_A A] C_\beta] B_\alpha$]

Essentially the same conclusions will hold under the reverse settings of the left/right specifications in alignment constraints, as in (31-ab).

- (31) a. $MC(\alpha) \gg MC(\beta) \gg A\Rightarrow R \gg L\Leftarrow\beta \gg L\Leftarrow\alpha$
 b. $MC(\alpha) \gg MC(\beta) \gg A\Rightarrow R \gg L\Leftarrow\alpha \gg L\Leftarrow\beta$

The scenario in (31-a) will not involve movement of the α -exponent to the left edge; but the one in (31-b) will bring about such movement; see (32-ab).

- (32) a. [${}_A C_\beta [{}_A B_\alpha A]$] b. [${}_A B_\alpha [{}_A C_\beta [{}_A A]]$]

To see this in more detail, let us look at the sequence of optimization procedures involved here. Let us also replace the abstract discussion based on variables like A, B, C, α , β with real examples, focussing on (simplified versions of) German verb inflection first (this will

instantiate a non-movement pattern, as in (29-a)), and Berber verb inflection afterwards (this will illustrate a movement pattern, as in (31-b)).¹⁵

Thus, suppose that A = a German V stem like *kauf* (‘buy’) that is equipped with morphological structure-building features [**•T•**] (= [**•α•**]) and [**•Agr•**] (= [**•β•**]) identifying tense and subject agreement exponents, respectively. Two relevant exponents for, say, PAST.2.SG environments are [_T te] and [_{Agr} st], and the eventual output that needs to be derived is *kauf-te-st* (‘buy-PAST-2.SG.’, ‘You bought.’). The initial optimization step is shown in (33).¹⁶

(33) *German verb inflection* (harmonic serialism, step 1):

I ₁ : [_V kauf]: [•T•], [•Agr•], [2], [SG], [PAST], {[_T /te/↔[PAST]], ... }, {[_{Agr} /st/↔[2.SG]], ... }	MC _T	MC _{Agr}	L↔V	Agr⇒R	T⇒R
O ₁₁ : [_V kauf]: [•T•], [•Agr•]	*!	*			
☞O ₁₂ : [_V kauf-te]: [•Agr•]		*			
O ₁₃ : [_V kauf-st]: [•T•]	*!				
O ₁₄ : [_V te-kauf]: [•Agr•]		*	*!		*
O ₁₅ : [_V st-kauf]: [•T•]	*!		*	*	

Since all competing outputs O₁₁–O₁₅ can be separated from the input I₁ by at most one operation, the candidate set is necessarily finite.¹⁷ Output O₁₁ leaves the input intact and is filtered out because of violations of MC_T and MC_{Agr}. O₁₂ discharges V’s [**•T•**] feature by merging /te/ as a suffix; this satisfies all constraints, except for MC_{Agr}. However, if MC_{Agr} is satisfied by discharging [**•Agr•**] on V and merging /st/, as in O₁₃, the higher-ranked MC_T is fatally violated. Note that the confinement to at most one operation between input and output ensures that no candidate can be generated in the first step that satisfies both MCs simultaneously, by merging both exponents. Next, O₁₄ is excluded because it fatally violates both the alignment constraint for the stem (L↔V) and the alignment constraint for the inflectional exponent (T⇒R), by realizing the tense exponent as a prefix. Finally, O₁₅’s constraint profile is hopeless. It fatally violates MC_T, like O₁₃; in addition, like O₁₄, it has gratuitous alignment violations (of L↔V and Agr⇒R).

In the next optimization step, the optimal output O₁₂ of (33) is used as the input; see (34).¹⁸

¹⁵ We hasten to add that the two examples we discuss primarily serve the purpose of highlighting the basic mechanics of the theory and its predictions for morphological movement, and are not meant as comprehensive analyses. Eventually, issues are more complex than we assume here.

¹⁶ Some notational conventions: I₁ is the initial input; here, the V stem and its structure-building features. The second line in the top-left box list the fully specified feature matrix associated with this stem that is in need of morphological realization (and that can subsequently be accessed in the syntax). The third line in this box illustrates the morphological arrays that will be used; for each of the two morphological arrays, only one morphological exponent is given here (viz., the one that will be selected by the faithfulness constraints from which this optimization abstracts away.) The morphological exponents pair (↔) a phonological form / / with a (often underspecified) set of morpho-syntactic features [], as in Distributed Morphology. Finally, immediate output descendants of a given input (e.g., O₁₃) extend the input’s last digit (e.g., O₁).

¹⁷ This presupposes that morphological arrays are also finite. This is the case. Choice from a given morphological array is governed by IDENT and MAX constraints, which accounts for syncretism via underspecified exponents. Although underspecification is ultimately highly relevant for most of the morphological systems addressed here, we will abstract away from this issue throughout this article since it is orthogonal to our main concerns.

¹⁸ It is postulated in Müller (2018) that once an exponent has been taken from a morphological array, it is gone permanently. This assumption is adopted here even though it is not actually important for our present purposes.

(34) *German verb inflection* (harmonic serialism, step 2):

I ₁₂ : [V kauf-te]: [●Agr●], [2], [SG], [PAST], { ... }, { [Agr /st/↔[2.SG]], ... }	MC _T	MC _{Agr}	L←V	Agr⇒R	T⇒R
O ₁₂₁ : [V kauf-te]: [●Agr●]		*!			
☞O ₁₂₂ : [V [V kauf-te]-st]					*
O ₁₂₃ : [V st-[V kauf-te]]			*!	**	

As before, the first competing output O₁₂₁ leaves the input intact. It is filtered out because of the MC_{Agr} violation that was tolerable in the previous optimization step but has become fatal now. In contrast, O₁₂₂ discharges the [●Agr●] feature by merging /st/ as a suffix; this candidate is optimal. This illustrates a typical property of harmonic serialism: There is a continuous, gradual improvement of the constraint profile. Still, O₁₂₂ does not respect all constraints: By satisfying high-ranked MC_{Agr}, it invariably introduces a new violation of lower-ranked T⇒R: Now /te/ is separated from the right edge of the word by an intervening exponent (viz., /st/). This problem could in principle be solved by merging /st/ as a prefix, as in O₁₂₃; but this fatally violates the two higher-ranked alignment constraints.¹⁹ Note that an output which would merge the /st/ exponent counter-cyclically, between the V stem and /te/, and which would (fatally) violate Agr⇒R but not L←V, cannot be generated, given the Strict Cycle Condition.

In the third and final optimization step, convergence is reached; see (35).

(35) *German verb inflection* (harmonic serialism, step 3):

I ₁₂₂ : [V [V kauf-te]-st] [2], [SG], [PAST], { ... }, { ... }	MC _T	MC _{Agr}	L←V	Agr⇒R	T⇒R
☞O ₁₂₂₁ : [V [V kauf-te]-st]					*
O ₁₂₂₂ : [V st-[V kauf-te]]			*!	**	
O ₁₂₂₃ : [V te-[V [V kauf]-st]]			*!		**
O ₁₂₂₃ : [V [V [V kauf]-st]-te]				*!	

O₁₂₂₁ does not change the input, and thus maintains the previous optimal candidate's T⇒R violation, but at this point there is no way to improve the constraint profile further by carrying out some operation. In particular, if /st/ shows up in front of the V stem, as in O₁₂₂₂, two higher-ranked alignment constraints are violated, as with O₁₂₃ in (34); and similar consequences result if /te/ were to be merged in either a prefix or suffix position, as in O₁₂₂₃ and O₁₂₂₄. However, it is worth noting at this point already that O₁₂₂₂–O₁₂₂₄ are outputs that carry out morphological movement: E.g., in O₁₂₂₂, /st/ is moved to the left edge from the right-peripheral position it shows up in in I₁₂₂.²⁰

Let us next look at a scenario that involves movement because the two MCs for features α and β show up in the same order as the two (unidirectional) alignment constraints for α and β , as in (31-b), where left-alignment constraints bring about a prefix status of inflectional exponents.²¹ Consider, e.g., an inflected Berber verb like *ad-y-seg* ('FUT-3.MASC.SG-buy'; 'He will buy.'). cf. Ouhalla (1991). Suppose that the order of MC_T and MC_{Agr} is the same as in German; in fact, let us generalize this assumption and propose that the order of

¹⁹ We assume a gradient interpretation of alignment constraints: O₁₂₃ violates Agr⇒R twice because /st/ is separated from the right edge by two intervening items.

²⁰ See Müller (2018, ch. 2) for slightly more intricate cases, based on Trommer (2008), where an initial suffix can in fact legitimately become a prefix in the course of repeated optimization.

²¹ Inflected verb forms instantiating the suffixal pattern in (29-b) for tense and subject agreement marking are not unattested, but somewhat rarer in the world's languages; see Julien (2002), Trommer (2001).

these two MCs is cross-linguistically invariant, and follows from an independently motivated functional sequence (f-seq; see Starke (2001)) of inflectional categories that also predicts a uniform order of functional projections in the syntax. This proposal then derives Mirror Principle effects (see Baker (1985)), albeit at an abstract level: Basic structure-building takes place in the same way, determined by f-seq, in morphology and syntax; but the Mirror Principle does not have to be surface-true because, as will be shown momentarily, morphological movement can apply after initial structure-building.²² Returning to the constraints that play a role, in addition to the two MCs, there are two alignment constraints $L \Leftarrow T$, $L \Leftarrow Agr$, and a higher-ranked constraint $V \Rightarrow R$ ensuring prefix status of the inflectional exponents.²³

The first optimization step is shown in (36).

(36) *Berber verb inflection* (harmonic serialism, step 1):

I ₁ : [v seg]: [●T●], [●Agr●], [3], [SG], [MASC], [FUT], { [T /ad/↔[FUT]], ... }, { [Agr /y/↔[3.SG.MASC]], ... }	MC _T	MC _{Agr}	V⇒R	L⇐T	L⇐Agr
O ₁₁ : [v seg]: [●T●], [●Agr●]	*!	*			
O ₁₂ : [v seg-y]: [●Agr●]	*!		*	*	
O ₁₃ : [v seg-ad]: [●T●]		*	*!		*
O ₁₄ : [v y-seg]: [●Agr●]	*!				
☞O ₁₅ : [v ad-seg]: [●T●]		*			

In (36), the inert output O₁₁ that leaves the input intact violates both MCs; and the outputs O₁₂ and O₁₄ that merge the agreement exponent first fatally violate higher-ranked MC_T. The optimal output O₁₅ merges the tense exponent /ad/ as a prefix, rather than as a suffix, as in O₁₃.

In the following step, based on O₁₅ as the new input, the optimal output merges the agreement exponent /y/ as a prefix; see (37). Here, O₁₅₃ emerges as optimal even though it introduces a violation of L⇐T (since /ad/ is now not at the left edge anymore); however, if /y/ is merged as a suffix, as in O₁₅₂, this will give rise not only to (tolerable) violations of lower-ranked L⇐Agr, but also to a fatal violation of higher-ranked V⇒R. An output that would squeeze in /y/ between /ad/ and the V stem /seg/ would yield the best constraint profile (violating only L⇐Agr) but cannot be generated, as a consequence of the Strict Cycle Condition.

(37) *Berber verb inflection* (harmonic serialism, step 2):

I ₁₅ : [v ad-seg]: [●T●], [3], [SG], [MASC], [FUT], { ... }, { [Agr /y/↔[3.SG.MASC]], ... }	MC _T	MC _{Agr}	V⇒R	L⇐T	L⇐Agr
O ₁₅₁ : [v ad-seg]: [●T●]		*!			
O ₁₅₂ : [v [v ad-seg]-y]			*!		**
☞O ₁₅₃ : [v y-[v ad-seg]]				*	

In this scenario, the constraint profile can and must be further improved by moving the tense exponent /ad/ to the front in the next optimization step; see (38).

²² We will address the question later of how morphomic features enter the picture.

²³ This is a huge simplification since several agreement exponents are actually realized as suffixes in Berber; see Noyer (1992) and Frampton (2002). A comprehensive analysis of the paradigm would have to address this by exploiting the basic violability of V⇒R for the optimal realization of certain ϕ -feature combinations. For present purposes, we will abstract away from these complications.

(38) *Berber verb inflection* (harmonic serialism, step 3):

I ₁₅₃ : [V y-[V ad-seg]], [3], [SG], [MASC], [FUT], { ... }, { ... }	MC _T	MC _{Agr}	V⇒R	L⇐T	L⇐Agr
O ₁₅₃₁ : [V y-[V ad-seg]]				*!	
⇨O ₁₅₃₂ : [V ad-[V y-[V seg]]]					*
O ₁₅₃₃ : [V [V [V ad-seg]-y]]			*!		**
O ₁₅₃₄ : [V [V y-[V seg]-ad]]			*!	**	

Here, O₁₅₃₂ trades in the input's L⇐T violation for a less severe L⇐Agr violation, and is therefore optimal. Other outputs carrying out rightward movement also compete but are filtered out as suboptimal because they incur fatal V⇒R violations: O₁₅₃₃, O₁₅₃₄. It is clear that any further movement operation will deteriorate the constraint profile; hence, convergence is reached in the next step.

To sum up: In the harmonic serialist approach to inflectional morphology outlined in this section, movement of exponents is a natural characteristic that arises under various scenarios, one of them consisting of two MCs and two alignment constraints for two features showing up in the same order (this is the scenario that will also be relevant for the account of phonological operations triggered by apparently non-local exponents described in section 2). Two properties of the approach are worth emphasizing in this context. First, the present approach does not rely on any special assumptions, of whatever type, to bring about movement of exponents: The constraints that trigger such movement are independently motivated, by considerations completely unrelated to movement. Something comparable does not hold for other theories of morphology that envisage displacement operations or various kinds (e.g., Distributed Morphology); here these operations must be stipulated on a case-by-case basis (we will come back to this issue in section 5). And second, the result that morphological movement comes for free is inherently tied to harmonic serialism, and cannot be attained in a standard parallel optimality-theoretic approach. What is crucial here is the restriction that outputs differ from inputs by applying maximally one operation; for this reason, operations may be locally (i.e., at an intermediate stage of the derivation) optimal that can never be optimal from a global perspective. Accordingly, in an optimality-theoretic analysis where the restriction is not in place, a representation like O₁₅₃ in (37), where the agreement exponent is outside of the tense exponent (/y/-/ad/-/seg/), will ceteris paribus never emerge; as illustrated in (39), the surface order will be determined directly, without movement.

(39) *Berber verb inflection* (standard parallel optimality theory):

I ₁ : [V seg]: [●T●], [●Agr●], [3], [SG], [MASC], [FUT], { [T /ad/↔[FUT]], ... }, { [Agr /y/↔[3.SG.MASC]], ... }	MC _T	MC _{Agr}	V⇒R	L⇐T	L⇐Agr
O ₁₁ : [V seg]: [●T●], [●Agr●]	*!	*			
O ₁₂ : [V seg-y]: [●Agr●]	*!		*		
O ₁₃ : [V seg-ad]: [●T●]		*	*!		
O ₁₄ : [V y-seg]: [●Agr●]	*!				
O ₁₅ : [V ad-seg]: [●T●]		*			
⇨O ₁₆ : [V ad-[V y-seg]]					*
O ₁₇ : [V y-[V ad-seg]]				*!	
O ₁₈ : [V [V y-seg]-ad]			*!	**	
O ₁₉ : [V [V ad-seg]-y]			*!		**
O ₂₀ : [V [V seg-ad]-y]			*!*	*	**
O ₂₁ : [V [V seg-y]-ad]			*!*	**	*

Thus, the correct surface order of exponents can be generated under the same constraints, and under an identical ranking, in the harmonic serialist and the standard optimality-theoretic approaches. However, there are conceptual and empirical considerations distinguishing the two alternatives. A possible conceptual argument has been given above: The harmonic serialist approach makes it possible to maintain some version of the Mirror Principle. As for empirical arguments, three pieces of evidence for movement of exponents in the morphological component, hence (in the present context), for harmonic serialism, are presented in Müller (2018).

First, the phenomenon of *discontinuous exponence* (see Harris (1945)), including circumfixation, can be accounted for by assuming that the “discontinuous” exponent is actually a discrete item initially, part of which gets separated by morphological movement.

Second, the phenomenon of *partially superfluous extended exponence*, where the morpho-syntactic features of two exponents in a single word stand in a proper subset relation, poses a problem for many theories of morphology because the more general (i.e., subset) exponent would seem to be globally unmotivated. Many solutions to this problem predict that the more general exponent has to be closer to the initial stem than the more specific exponent (see Caballero & Harris (2012), Stiebels (2015), and Müller (2018)). However, in some cases (e.g., with extended exponence of negation on the verb in Swahili; cf. Stump (2001)) the surface order is at variance with this requirement. Morphological movement in harmonic serialism solves this problem: The more general exponent is merged closer to the stem first, the more specific exponent is merged later, and finally the former exponent moves across the latter.

Third, the phenomenon of *non-local stem allomorphy* (i.e., stem allomorphy conditioned by features on exponents that are not next to the stem on the surface) has motivated a number of special devices in the literature (see Merchant (2015), Moskal & Smith (2016), and Božič (2019)) but can be reinterpreted in the present approach as extremely local stem allomorphy accompanied by subsequent exponent movement.

Against this background, it should be clear that the cases of seemingly non-local phonological operations in words discussed in section 2 qualify as a further argument for movement in morphology.

4. Phonological Evidence for Movement in Morphology

4.1. Cycles

The hypothesis we want to pursue in what follows is that what may at first sight look like a non-local application of a phonological operation can upon closer inspection be shown to be a strictly local phonological operation conditioned by an exponent that eventually shows up in a non-local position as a consequence of morphological movement. At this point, the necessary assumptions to implement this for the five case studies are already in place, except for one: It needs to be determined when phonological operations can take place in harmonic serialist derivations.

We would like to suggest that phonological operations apply to the output of a morphological cycle, and that there are *two* morphological cycles in the present approach to inflectional morphology based on harmonic serialism: The first morphological cycle is finished when all MC-triggered Merge operations have applied, and the morphological arrays are exhausted.²⁴ Intuitively, this is the stage where the word is potentially complete for the first time. The second morphological cycle is finished when the derivation has converged on

²⁴ This will be minimally modified in the following section in view of morphomic exponents.

a final output. In the above example illustrating German verb inflection, the two morphological cycles converge on O_{122} in (34). However, in the example illustrating Berber verb inflection, the two morphological cycles are not identical: O_{153} in (37) completes a first morphological cycle, and O_{1532} in (38) ends a second one.^{25,26} This opens up the possibility that phonological operations may apply after the first cycle already, before morphological movement takes place; and as we will show in the next five subsections, this is what underlies the effects discussed in section 2.

4.2. *De-Spirantization in Barwar Aramaic*

Consider first seemingly non-local de-spirantization in Barwar Aramaic. Recall from (2) and (7) that the tense/aspect exponent $/wa/\leftrightarrow[\text{REMOTE}]$ can intervene between the stem and an agreement exponent like $/le/\leftrightarrow[3.\text{SG}.\text{MASC}]$ on the surface but does not block de-spirantization of the final stem consonant triggered by $/le/$. Given the reasoning in section 3, this implies that $/le/$ is merged before $/wa/$ (and the right-alignment constraint for $/le/$ outranks that for $/wa/$). At first sight, this might seem to necessitate deviating from the assumption that the order of MCs in the morphological component reflects f-seq. However, closer scrutiny reveals that this is not the case. Although $/wa/$ has been classified as a “tense/aspect” exponent above, its actual function in the system is not the basic realization of this grammatical category. Rather, the real tense exponent in Barwar Aramaic is the root itself, in the form of non-concatenative root-and-pattern morphology typical of Semitic languages. The exponent $/wa/$ has a more adverb-like role; it freely attaches to present or past tense stems to add information of anteriority, or remoteness. Historically, it derives from a separate past tense verb **hwa* (cf. Khan (2008, 177)). We take this to be a clear indication that $/wa/$ lies outside the clausal f-seq, which produces a ranking $\text{MC}_T \gg \text{MC}_{\text{Agr}}$ in Barwar Aramaic, as in other languages; *wa* is introduced into words by a $\text{MC}_{\text{Adv(erb):T(emp)}}$ that is lower-ranked. Under these assumptions, and zooming in at a stage in the derivation where non-concatenative root morphology has been optimized (in response to MC_T and other constraints), the first optimization step for a form with de-spirantization like *t^yridwale* (‘chase.away(PAST)-REMOTE-3SG.SBJ’, ‘He had chased away’; cf. (7)) is shown in (40). The optimal output O_{12} merges $/le/$ as a suffix.²⁷

²⁵ The characterization of the first morphological cycle as a “potentially complete” word where all structure-building features have been discharged bears an obvious resemblance to Chomsky’s (2000) characterization of phases (especially the subcase of “a verb phrase in which all θ -roles are assigned”); an analogous reasoning could be provided for the second morphological cycle (“full clause”). Following much recent literature (e.g., Marvin (2002), Embick (2010)), one could in principle therefore refer to the endpoints of the two morphological cycles as “phases” – but it should be kept in mind that in the present approach, morphological and syntactic phases would be distinct kinds of objects.

An alternative way of looking at the two morphological cycles would be to assimilate them to D-structure and S-structure in classic government and binding theory (see Chomsky (1981)): The first cycle (“D-structure”) is defined by basic structure-building; the second cycle (“S-structure”) is defined by the application of all movement operations.

²⁶ Note that this view is in principle fully compatible with the existence of prior stem, and also root, cycles in a stratal approach; see Kiparsky (1982b), Bermúdez-Otero (2008), and Trommer (2011).

²⁷ We assume here that the V stem is *optionally* enriched with a structure-building feature for the remoteness exponent $/wa/$, as it is standardly assumed for optional pieces of structure in approaches where all structure-building must be feature-driven. Note also that the morphological array for $/wa/$ is assumed to be a non-singleton set; as observed by Khan (2008), there are a few other exponents with similar types of functions.

(40) *De-spirantization in Barwar Aramaic* (harmonic serialism, step 1):

I ₁ : [v tʰr(i)ð]: [●Agr●], [●Adv:T●], [3], [SG], [MASC] [PAST], { [Agr /le/↔[3.SG.MASC]], ... }, { [Adv:T /wa/↔[REMOTE]], ... }	MC _{Agr}	MC _{Adv:T}	L←V	Agr⇒R	Adv:T⇒R
O ₁₁ : [v tʰr(i)ð]: [●Agr●], [●Adv:T●]	*!	*			
⇨O ₁₂ : [v [v tʰr(i)ð]-le]: [●Adv:T●]		*			
O ₁₃ : [v [v tʰr(i)ð]-wa]: [●Agr●]	*!				
O ₁₄ : [v le-[v tʰr(i)ð]]: [●Adv:T●]		*	*!	*	
O ₁₅ : [v wa-[v tʰr(i)ð]]: [●Agr●]	*!		*		*

In the next optimization step documented in (41), O₁₂₂ is optimal: /wa/ is added, and it is also added as a suffix (because of the ranking L←V ≫ Agr⇒R). As before, placing /wa/ in a position between V and /le/ would in principle give rise to the best constraint profile (because L←V and Agr⇒R would be satisfied, and Adv:T⇒R would only be violated once), but this option is not available because of the Strict Cycle Condition.

(41) *De-spirantization in Barwar Aramaic* (harmonic serialism, step 2):

I ₁₂ : [v [v tʰr(i)ð]-le]: [●Adv:T●], [3], [SG], [MASC] [PAST], { ... }, { [Adv:T /wa/↔[REMOTE]], ... }	MC _{Agr}	MC _{Adv:T}	L←V	Agr⇒R	Adv:T⇒R
O ₁₂₁ : [v [v tʰr(i)ð]-le]: [●Adv:T●]		*!			
⇨O ₁₂₂ : [v [v [v tʰr(i)ð]-le]-wa]				*	
O ₁₂₃ : [v wa-[v [v tʰr(i)ð]-le]]			*!		**

At this point, the V stem has discharged all its structure-building features, so a morphological cycle is completed, and phonological operations can be triggered. Thus, de-spirantization applies to O₁₂₂, under strictly local conditions, turning /tʰr(i)ð-le-wa/ into /tʰr(i)d-le-wa/. After this, the next morphological cycle starts; and, as shown in (42), the constraint profile of the word can indeed be further improved by carrying out morphological movement of /le/ to the right edge: O₁₂₂₂ trades in the violation of higher-ranked Agr⇒R incurred by O₁₂₂₁, which leaves the input intact, for a violation of lower-ranked Adv:T⇒R.

(42) *De-spirantization in Barwar Aramaic* (harmonic serialism, step 3):

I ₁₂₂ : [v [v [v tʰr(i)d]-le]-wa] [3], [SG], [MASC] [PAST], { ... }, { ... }	MC _{Agr}	MC _{Adv:T}	L←V	Agr⇒R	Adv:T⇒R
O ₁₂₂₁ : [v [v [v tʰr(i)d]-le]-wa]				*!	
⇨O ₁₂₂₂ : [v [v [v [v tʰr(i)ð]]-wa]-le]					*
O ₁₂₂₃ : [v wa-[v [v tʰr(i)d]-le]]			*!		**
O ₁₂₂₄ : [v le-[v [v tʰr(i)d]]-wa]			*!	**	

The next step yields convergence, and a second phonological cycle can start which, however, does not affect de-spirantization anymore: Morphological movement counter-bleeds de-spirantization.

An immediate prediction of this approach is that exponents which are introduced earlier than the subject agreement marker /le/ in this configuration according to f-seq should ceteris paribus not give rise to de-spirantization. This is indeed the case. For instance, as shown in (6) above, an object agreement exponent intervening between the stem and a subject agreement exponent, as in /tʰr(i)ð-a-le/ ('chase.away(PAST)-F.OBJ-3SG.M.SBJ'), blocks de-spirantization. Finally, a remark is due to the optionality of the process; recall from (7) that [tʰridwale] and [tʰriðwale] co-exist. At least for present purposes, it may suffice to

assume that de-spirantization applies without qualification after the second morphological cycle but may or may not apply after the first one (this could, e.g., be modelled by postulating a tie of constraints forcing and counter-acting de-spirantization in one phonological stratum which is resolved into a fixed order in the second one).

4.3. *Saussurean Accent Shift in Lithuanian*

The problem posed by theme vowels in Lithuanian is that they do not seem to block Saussurean accent shift (such that main stress is placed on the inflectional exponent pre-specified for accent) even though they do seem intervene on the surface, are associated with a mora themselves, and should thus block this phonological operation; thus, /kás + [ee→^ja]/ + /ú/ (‘dig-1.SG.PAST’) becomes *kas^jau* (see (11)). In what follows, we will show that this phenomenon lends itself to essentially the same kind of analysis as de-spirantization in Barwar Aramaic, with one minor qualification concerning the definition of the end of the first morphological cycle.

To begin with, we follow Kushnir (2018) in assuming that (abstracting away from pre-verbs) the finite Lithuanian verb consists of a stem followed by three inflectional exponents: first, a tense exponent, which is \emptyset in present and past contexts and realized as /s/ in future contexts; second, a theme vowel; and third, a subject agreement exponent encoding person and number: V-T-Th-Agr. Since the language does not have overt tense exponents for present and past environments, the difference has to be encoded in some other way. This is accomplished by (a) stem alternations, and (b) theme vowel choice – as for the latter, see, e.g., /dírb- \emptyset -a-u/ (‘work-1.SG.PRES’) vs. /dírb- \emptyset -oo-u/ (‘work-1.SG.PAST’). However, there is no direct implicational relation between tenses and theme vowels: On the one hand, a given tense can co-occur with various theme vowels (e.g., present tense can co-occur with /a/, /ia/, /i/, and /oo/, depending on the V stem); and on the other hand, a given theme vowel can co-occur with more than one tense (so, /oo/ shows up in the present tense with some verbs, and in the past tense with many other verbs; /i/ occurs in the present tense with some verbs, and in the future tense with all verbs). In view of this state of affairs, we conclude that the theme vowels in Lithuanian are best analyzed as a morphomic category [Th] that plays no role in syntax, and is only indirectly related to morpho-syntactic (i.e., non-morphomic) features. In sum, this yields three morphological arrays for the Lithuanian verb, which are accordingly targetted by three structure-building features on the V stem: [\bullet T \bullet], [\bullet Agr \bullet], and [\bullet Th \bullet].²⁸ Given f-seq, we expect MC_T to outrank MC_{Agr}; crucially, MC_{Th}, being concerned with a morphomic feature, will be lowest-ranked, in the same way that non-f-seq-related MC_{Adv:T} was argued to be in the previous section. This implies that theme vowels are merged last. The ranking of the respective alignment constraints that produces movement of the agreement exponent across the theme vowel, so that the latter ends up in its surface position between the T exponent and the Agr exponent, then is Agr⇒R ≫ Th⇒R ≫ T⇒R.

Based on these assumptions, the first optimization step in the derivation of a form like *kas^jau* (‘dig-1.SG.PAST’) yields a winning candidate where the T exponent (\emptyset) has been merged as a suffix; cf. O₁₂ in (43).²⁹

²⁸ Note that this may or may not require a minimal modification of the concept of morphological array in (27), to ensure that theme vowels do not enter a morphological array together with T exponents. The issue here is whether or not the theme vowel can be said to *realize* a tense feature if it is partially related to a tense environment.

²⁹ To simplify exposition, we will disregard all outputs that merge a T, Agr, or Th exponent in prefix position, or that carry out movement to prefix position. As argued before, this is never an option for exponents in the presence of a high-ranked constraint L⇐V.

(43) *Saussurean accent shift in Lithuanian* (harmonic serialism, step 1):

I ₁ : [v kás]: [●T●], [●Agr●], [●Th●], [1], [SG], [PAST], { [T /Ø/↔[PAST]], ... }, { [Agr /ú/↔[1.SG]], ... }, { [Th /ee/↔[kás:PAST]], ... }	MC _T	MC _{Agr}	MC _{Th}	AGR⇒R	Th⇒R	T⇒R
O ₁₁ : kás: [●T●], [●Agr●], [●Th●]	*!	*	*			
⇨O ₁₂ : kás-Ø: [●Agr●], [●Th●]		*	*			
O ₁₃ : kás-ú: [●T●], [●Th●]	*!		*			
O ₁₄ : kás-ee: [●T●], [●Agr●]	*!	*				

In the next step, the Agr exponent /ú/ is merged; see (44).

(44) *Saussurean accent shift in Lithuanian* (harmonic serialism, step 2):

I ₁₂ : kás-Ø: [●Agr●], [●Th●], [1], [SG], [PAST], { ... }, { [Agr /ú/↔[1.SG]], ... }, { [Th /ee/↔[kás:PAST]], ... }	MC _T	MC _{Agr}	MC _{Th}	AGR⇒R	Th⇒R	T⇒R
O ₁₂₁ : kás-Ø: [●Agr●], [●Th●]		*!	*			
⇨O ₁₂₂ : kás-Ø-ú: [●Th●]			*			*
O ₁₂₃ : kás-Ø-ee: [●Agr●]		*!				*

Importantly, it has to be at this point, before merging the theme vowel in the next step, that the first morphological cycle is concluded, and phonological operations can be triggered: As soon as the theme vowel is part of the representation, the context for Saussurean accent shift is not given anymore, unless additional assumptions are made; recall the example in (9-c) above where a mora following the stressed mora in the dative plural exponent /ó_μo_μms/ blocks the process. Hence, the definition of completion of the first morphological cycle can be slightly changed: It is not captured by the point where all MC-triggered Merge operations have applied, but rather by the point where all MC-triggered Merge operations involving *morpho-syntactic* (i.e., non-morphomic) features have taken place. This maintains the original hypothesis that it is the first presence of a morpho-syntactically *complete* word that defines the option to start phonological operations; morphomic exponents like theme vowels are irrelevant from this perspective.³⁰ Given this assumption, Saussurean shift applies to O₁₂₂ in (44); this fixes the main stress of the word, which cannot subsequently be undone. In the next step, the theme vowel is merged; cf. (45).³¹

(45) *Saussurean accent shift in Lithuanian* (harmonic serialism, step 3):

I ₁₂₂ : kas-Ø-ú: [●Th●], [1], [SG], [PAST], { ... }, { ... }, { [Th /ee/↔[kás:PAST]], ... }	MC _T	MC _{Agr}	MC _{Th}	AGR⇒R	Th⇒R	T⇒R
O ₁₂₂₁ : kas-Ø-ú: [●Th●]			*!			**
⇨O ₁₂₂₂ : kas-Ø-ú-ee				*		**

Since Agr⇒R outranks Th⇒R, morphological movement of /ú/ to the right edge takes place in the next optimization round: The violation of Agr⇒R that was tolerable in the previous optimization because there was no better alternative becomes fatal as soon as there is one; see (46).

³⁰ Also see work like Oltra Massuet & Arregi (2005), where theme vowels are inserted post-syntactically in a Distributed Morphology approach.

³¹ Note that late introduction of the theme vowel is also the solution proposed in Kushnir (2018); however, Kushnir employs the counter-cyclic operation of interfixation to achieve the intended result. See section 5.

(46) *Saussurean accent shift in Lithuanian* (harmonic serialism, step 4):

I ₁₂₂₂ : kas-Ø-ú-ee, [1], [SG], [PAST], { ... }, { ... }, { ... }	MC _T	MC _{AgT}	MC _{Th}	AGR⇒R	Th⇒R	T⇒R
O ₁₂₂₂₁ : kas-Ø-ú-ee				*!		**
☞O ₁₂₂₂₂ : kas-Ø-ee-ú					*	**

The next and final optimization step produces convergence. At this point, the second phonological cycle starts, which turns O₁₂₂₂₂, i.e., /kas-Ø-ee-ú/, into the surface representation *kas^j aú*.

4.4. *Ni-Insertion in Quechua*

Recall that the phonological operation of *ni*-epenthesis breaking up super-heavy syllables seems to apply across an intervening exponent *lla* (‘just’) that would normally be expected to destroy the environment for *ni*-insertion both in Bolivian Quechua, where this process is obligatory (cf. (14): *wawa-s-ni-lla-y* (‘child-PL-NI-just-1.POSS’; ‘just my children’)), and in Huallaga Quechua, where it is optional (cf. (15-a): *kikish-ni-lla-yki* (‘armpit-just-2.POSS’; ‘just your armpit’)). The basic pattern straightforwardly lends itself to the kind of analysis in terms of morphological movement we have given in the previous two subsections. In fact, the main difference to the analysis proposed for Barwar Aramaic is that the stem that is subject to inflection is of type N now, rather than of type V. The morphological array for exponents bearing the feature [Poss] is clearly part of the f-seq of nominal categories (possessive marking is a property of nouns in the language); and the morphological array for exponents like *lla*, which is completely optional, is a marker denoting “limitation”, and “typically means ‘just’ or ‘only’, but also has a range of more subtle uses which are much harder to translate” (Myler (2013, 4.2)), clearly is not. By the same reasoning as in 4.2. above, we can therefore conclude that there are two MCs, MC_{Poss} and MC_{Lim(itiation)}, and that the former outranks the latter. Given, furthermore, a parallel ranking for the alignment constraints Poss⇒R and Lim⇒R and a high-ranked L⇐N, morphological movement of the possessor exponent will be triggered after both MCs have been satisfied. The first optimization step for Bolivian Quechua *wawa-s-ni-lla-y* (‘child-PL-NI-just-1.POSS’) is shown in (47).³²

(47) *Ni-epenthesis in Bolivian Quechua* (harmonic serialism, step 1):

I ₁ : [N wawa-s]: [●Poss●], [●Lim●], [PL], [1.POSS], { [Poss /y/↔[1.POSS]], ... }, { [Lim /lla/↔[JUST]], ... }	MC _{Poss}	MC _{Lim}	L⇐N	Poss⇒R	Lim⇒R
O ₁₁ : [N wawa-s]: [●Poss●], [●Lim●]	*!	*			
☞O ₁₂ : [N [N wawa-s]-y]: [●Lim●]		*			
O ₁₃ : [N [N wawa-s]-lla]: [●Poss●]	*!				
O ₁₄ : [N y-[N wawa-s]]: [●Lim●]		*	*!	*	
O ₁₅ : [N lla-[N wawa-s]]: [●Poss●]	*!		*		*

After merging the possessor exponent /y/ in (47) because of highest-ranked MC_{Poss}, the derivation gets rid of the next severe constraint violation (viz., that of MC_{Lim}) by merging /lla/; cf. (48).

³² More precisely, this is the first optimization step that is relevant in the present context – we zoom in on the derivation after the plural exponent has been merged. The right-alignment constraint for the plural exponent must be ranked low, so that movement to the right periphery of the word is never triggered for this exponent.

(48) *Ni-epenthesis in Bolivian Quechua* (harmonic serialism, step 2):

I ₁₂ : [N [N wawa-s]-y]: [●Lim●], [PL], [1.POSS], { ... }, { [Lim /lla/↔[JUST]], ... }	MC _{Poss}	MC _{Lim}	L←N	Poss⇒R	Lim⇒R
O ₁₂₁ : [N [N wawa-s]-y]: [●Lim●]		*!			
☞O ₁₂₂ : [N [N [N wawa-s]-y]-lla]				*	
O ₁₂₃ : [N lla-[N [N wawa-s]-y]]			*!		**

Given the high ranking of L←N, /lla/ must be merged as a suffix in (48); and given the Strict Cycle Condition, it can only be added at the right edge, in violation of Poss⇒R. At this point, the first morphological cycle is concluded (all non-morphomic MCs have been satisfied), and the phonological operation of *ni-epenthesis* applies to O₁₂₂, breaking up the super-heavy syllable with a rhyme consisting of *a*, *s*, and *y*: /wawa-s-y-lla/ becomes /wawa-s-ni-y-lla/; /lla/ is already part of the representation but does not intervene because it shows up at the right edge at this stage. After this, the morphological derivation continues; since Poss⇒R outranks Lim⇒R, morphological movement of /y/ across /lla/ is triggered; the counter-bleeding effect for *ni-epenthesis* is illustrated in (49).³³ Finally, convergence is reached.

(49) *Ni-epenthesis in Bolivian Quechua* (harmonic serialism, step 3):

I ₁₂₂ : [N [N [N wawa-s]-ni-y]-lla], [PL], [1.POSS], { ... }, { ... }	MC _{Poss}	MC _{Lim}	L←N	Poss⇒R	Lim⇒R
O ₁₂₁ : [N [N [N wawa-s]-ni-y]-lla]				*!	
☞O ₁₂₂ : [N [N [N [N wawa-s]-ni-]-lla]-y]					*

With respect to the optionality of the process in Huallaga Quechua, it can be assumed (essentially as postulated for the optionality of de-spirantization in subsection 4.2.) that *ni-epenthesis* always applies after the morphological derivation is complete; but it may or may not apply after the first morphological cycle (which can be implemented via the concept of constraint tie).

Finally, as regards the contextually determined systematic variation in /lla/ placement visible in (17) in Huallaga Quechua (where /lla/ follows the possessor exponent if the latter co-occurs with a plural exponent, but not otherwise), we contend that the effect is due to a constraint COH(ERENCE) that is independently motivated in Trommer (2008), and ranked higher in the variety at hand; COH demands that two exponents that “belong together” in some sense must show up next to one another. See Müller (2018, ch. 2) for illustration of how COH can trigger and block movement of inflectional exponents in the harmonic serialist approach adopted here.

4.5. Ruki Rule Application in Sanskrit

The core observation here is that an inflected Sanskrit past tense verb shows application of the ruki rule with prefixed verbs where the final segment of the prefix acts as a trigger, turning stem-initial *s* into retroflex *ṣ* after *r*, *u*, *k*, and *i*, even though a past tense (non- ϕ) exponent /a/ intervenes on the surface; cf. (20-c): /abhy-a-ṣiñc-at/ (‘on-PAST-pour-, 3.SG’). The general logic of the analyses presented so far implies that the prefix /abhi/ (‘unto’) combines with the stem /siñc/ (‘sprinkle’) before the past tense exponent /a/ is added. It turns out that it suffices to make either of the two following traditional assumptions to derive this: (i) Compounding and derivation precede inflection; (ii) Prefix verbs are stored as such in the lexicon. In what

³³ Outputs carrying out suboptimal movement operations to the left edge are left out here; see (42).

follows, we will remain neutral as to the choice between (i) and (ii). Either way, there is a stage of the derivation that looks as in (50), with a complex verb /abhi-siñc/ consisting of prefix and stem as the input. Furthermore, we assume that the past tense prefix /a/↔[PAST] and the agreement suffix /at/↔[3.SG] are part of two separate morphological arrays, MC_T and MC_{Agr} .³⁴ Since three non-stem exponents and one stem exponent are involved, there need to be four alignment constraints; of these, $Agr \Rightarrow R$ must be highest-ranked because agreement markers show up as suffixes, in violation of $V \Rightarrow R$; and $L \Leftarrow P$ (for the prefix) and $L \Leftarrow T$ (for the T exponent) are ranked lower, in this order.

(50) *Ruki rule application in Sanskrit* (harmonic serialism, step 1):

I_1 : [v [P abhi-[V siñc]]]: [●T●], [●Agr●] [3], [SG], [PAST] { [T /a/↔[PAST]], ... }, { [Agr /at/↔[3.SG]], ... }	MC_T	MC_{Agr}	$Agr \Rightarrow R$	$V \Rightarrow R$	$L \Leftarrow P$	$L \Leftarrow T$
O_{11} : [v [P abhi-[V siñc]]]: [●T●], [●Agr●]	*!	*				
O_{12} : [v [v [P abhi-[V siñc]]]-a]: [●Agr●]		*		*!		**
O_{13} : [v a-[v [P abhi-[V siñc]]]]: [●Agr●]		*			*	
O_{14} : [v [v [P abhi-[V siñc]]]-at]: [●T●]	*			*		
O_{15} : [v at-[v [P abhi-[V siñc]]]]: [●T●]	*!		*		*	

The optimal output in (50) is O_{13} , which merges the T exponent /a/ in a prefix position, at the cost of a violation of the constraint $L \Leftarrow P$, which requires the verbal prefix to be at the left edge, and which, as such, is ranked higher than $L \Leftarrow T$ but cannot be satisfied at this point because of $V \Rightarrow R$ and the Strict Cycle Condition. In the next step, the Agr exponent /at/ is merged; see (51).

(51) *Ruki rule application in Sanskrit* (harmonic serialism, step 2):

I_{13} : [v a-[v [P abhi-[V siñc]]]]: [●Agr●] [3], [SG], [PAST] { ... }, { [Agr /at/↔[3.SG]], ... }	MC_T	MC_{Agr}	$Agr \Rightarrow R$	$V \Rightarrow R$	$L \Leftarrow P$	$L \Leftarrow T$
O_{131} : [v a-[v [P abhi-[V siñc]]]]: [●Agr●]		*!			*	
O_{132} : [v [v a-[v [P abhi-[V siñc]]]]-at]				*	*	
O_{133} : [v at-[v a-[v [P abhi-[V siñc]]]]]			*!***		**	*
O_{134} : [v abhi-[v a-[v [P [V siñc]]]]]: [●Agr●]		*!				*

The optimal output O_{132} in (51) violates $V \Rightarrow R$ so as to satisfy higher-ranked $Agr \Rightarrow R$; and it inherits the input's violation of $L \Leftarrow P$ because MC_{Agr} is the higher-ranked constraint; see O_{134} . However, at this point the word is structurally complete; consequently, a first phonological cycle is initiated, and the ruki rule applies under adjacency of /abhi/ and /siñc/.

Next, the prefix /abhi/ moves to the left edge, because of $L \Leftarrow P \gg L \Leftarrow T$; cf. (52).³⁵

³⁴ This might be a simplification. One might argue that the two exponents instantiate a case of overlapping extended exponence (in Caballero & Harris's (2012) terminology), with /at/ also being specified for PAST tense information. If so, the two exponents would then belong to a single morphological array, given the definition in (27), and could therefore only be targeted by a single MC ($MC_{T/Agr}$). Under this analysis, the second of the two exponents is merged with the stem not because of a MC, but so as to satisfy MAX faithfulness constraints (see Müller (2018, ch. 3)). This difference is orthogonal to the issue of morphological movement as such. However, the first morphological cycle would ceteris paribus be completed directly after step 1 in (50) (rather than after step 2 in (51), as we will see momentarily). While it is conceivable that this might give rise to different empirical predictions, we are not aware of any phonological effects that might bear on the issue whether /at/ is present or not at the stage when the ruki rule applies. Hence, in what follows we maintain the simpler version in terms of two MCs for T and Agr.

³⁵ Again, outputs employing other kinds of movement are ignored in the tableau.

(52) *Ruki rule application in Sanskrit* (harmonic serialism, step 3):

I ₁₃₃ : [V [V a-[V [P abhi-[V šĩnc]]]]-at] [3], [SG], [PAST] { ... }, { ... }	MC _T	MC _{Agf}	Agr⇒R	V⇒R	L⇐P	L⇐T
O ₁₃₃₁ : [V [V a-[V [P abhi-[V šĩnc]]]]-at]				*	*	
⇨O ₁₃₃₂ : [V abhi-[V [V a-[V -[V šĩnc]]]]-at]						*

Thus, morphological movement of exponents counter-bleeds a phonological operation, as before: /a/ does not intervene at the relevant stage of the derivation even though it does intervene on the surface.³⁶ Finally, convergence is reached, and the second set of phonological operations can apply.

4.6. *Vowel Harmony in Kazakh*

As shown above (cf. (23)), vowel harmony in Kazakh affects a final Q exponent /bA/ across an intervening instrumental case exponent /men/ that does not participate in the process and should be expected to block spreading of the stem's backness value to the Q marker; cf. *bʉl fal nan-men-ba* ('this old.man bread-INSTR-Q') vs. *bʉl fal bʉbek-men-be* ('this old.man baby-INSTR-Q'). The first thing to note is that interrogativity is a grammatical category for which nouns can be inflected in Kazakh, just like declarativity; the feature [C] of which Q is a possible value is part of the f-seq of nominal projections.³⁷ Next, we propose that the instrumental case exponent /men/ does not share a morphological array with the other case exponents in the language, and is introduced into words by a special, designated MC_{Instr} constraint that is outside of the regular, f-seq-determined order. In line with this special morpho-syntactic behaviour, /men/ also exhibits unusual phonological properties: It cannot be stressed and, more importantly in the present context, it resists vowel harmony. Under these assumptions, the apparently non-local instance of vowel harmony can be derived in a strictly local fashion in the same way as the phonological processes discussed in the previous sections: The order of the right-alignment constraints for the C exponent /bA/ and the instrumental exponent /men/ is the same as the order of the two MCs for these items.

Thus, to generate a word like *nan-men-ba* ('bread-INSTR-Q'), the noun stem /nan/ is first merged with the the Q exponent /bA/, as in O₁₂, which wins the competition in (53).

³⁶ This conclusion is also reached in the analyses of the phenomenon developed by Kiparsky (1982b) and Myler (2017). Kiparsky's analysis presupposes a counter-cyclic operation integrating /a/; Myler's approach actually also employs movement of /abhi/, like the present one (albeit of a very different type). Cf. section 5.

³⁷ The fact that nouns can be inflected for interrogativity only in clause-final position in Kazakh does not call into question the validity of this assumption; it is often the case that words which realize a given grammatical category are confined to certain syntactic positions. A particularly obvious example illustrating this is the choice between strong and weak adjective declension in German: Whether a (case/φ) exponent from the strong paradigm or a (case/φ) exponent from the weak paradigm is selected is exclusively determined by the syntactic context of the adjective (viz., presence vs. absence of a certain type of D item). For the case at hand, it can be assumed that [C:Q] on N that is realized pre-syntactically by /bA/ requires local Agree with a C head in the syntax, giving rise to adjacency.

(53) *Vowel harmony in Kazakh* (harmonic serialism, step 1):

I ₁ : [N nan]: [●C●], [●Instr●] [SG], [Q], [INSTR] {[C /bA/↔[Q]], ... }, {[Instr /men/↔[INSTR]], ... }					
	MC _C	MC _{Instr}	L←N	C⇒R	Instr⇒R
O ₁₁ : [N nan]: [●C●], [●Instr●]	*!	*			
⇨O ₁₂ : [N [N nan]-bA]: [●Instr●]		*			
O ₁₃ : [N bA-[N nan]]: [●Instr●]		*	*!	*	
O ₁₄ : [N [N nan]-men]: [●C●]	*!				
O ₁₅ : [N men-[N nan]]: [●C●]	*!		*		*

In the second step, /men/ is merged as a suffix; cf. (54).

(54) *Vowel harmony in Kazakh* (harmonic serialism, step 2):

O ₁₂ : [N [N nan]-bA]: [●Instr●] [SG], [Q], [INSTR] { ... }, {[Instr /men/↔[INSTR]], ... }					
	MC _C	MC _{Instr}	L←N	C⇒R	Instr⇒R
O ₁₂₁ : [N [N nan]-bA]: [●Instr●]		*!			
⇨O ₁₂₂ : [N [N [N nan]-bA]-men]				*	
O ₁₂₃ : [N men-[N [N nan]-bA]]			*!		*

At this point, the first morphological cycle is finished because the N stem has exhausted its structure-building features. Vowel harmony applies, spreading backness from the stem to the adjacent Q suffix /bA/; /men/ is pre-specified for resisting vowel harmony, and so does not change. The output of this round of phonological spell-out is then subjected to the next morphological cycle. Now, movement of the Q exponent /ba/ to the right edge takes place because C⇒R outranks Instr⇒R; see (55) (where outputs with suboptimal movements are ignored, as before).

(55) *Vowel harmony in Kazakh* (harmonic serialism, step 3):

O ₁₂₂ : [N [N [N nan]-ba]-men] [SG], [Q], [INSTR] { ... }, { ... }					
	MC _C	MC _{Instr}	L←N	C⇒R	Instr⇒R
O ₁₂₂₁ : [N [N [N nan]-ba]-men]				*!	
⇨O ₁₂₂₂ : [N [N [N [N nan]]-men]-ba]					*

The next step is the final step; O₁₂₂₂ as the input is mapped to the same output. As before, a seemingly non-local phonological process (vowel harmony) emerges as a local process that is counter-bleed by morphological movement.

4.7. *Conclusion*

To conclude so far, we have looked at five different seemingly non-local phonological operations and have argued that they can all be conceived of as strictly local operations in the harmonic serialist approach to inflectional morphology developed in Müller (2018). In this approach, the defining property of harmonic serialism that competing outputs can differ from the input by application of at most one elementary operation, together with the Strict Cycle Condition, automatically yields the consequence that alignment constraints that need to be postulated for reasons entirely unrelated to movement may force exponents to leave the position they are first merged in, and move to an edge position later in the derivation. Given the assumption that a first set of phonological operations can in principle apply as soon as a morphological word is complete (i.e., has all the non-morphomic exponents required by the

structure-building constraints), the phonological operations we have investigated here can all be shown to be counter-bleed by subsequent morphological movement.³⁸

In this paper, we have focussed on inflectional morphology. We suspect that there may turn out to be many more cases in this area once one starts looking in earnest for the relevant patterns.³⁹ Furthermore, it can be noted that similar phenomena involving apparently non-local phonological effects have also been observed for derivational morphology; see in particular Hyman (1994; 2002; 2003) on frication that seems to be triggered by non-local causative exponent in Bemba, and Myler (2017) on reduplication involving a non-local passive exponent in Ndebele. Assuming structure-building and linearization in derivational morphology to proceed in the same way as in inflectional morphology, these kinds of data would seem to lend themselves to the same kinds of analyses.

More generally, then, we take it that the phenomenon under discussion can be viewed as further support for a principled approach to morphological movement of exponents. As noted above, there are other arguments for this concept detailed in Müller (2018), based on discontinuous exponence, partially superfluous extended exponence, and non-local stem allomorphy. An additional conceptual argument for morphological movement of the type assumed here is that it permits stating, via shared, uniform f-seqs, a version of Mirror Principle that is not necessarily falsified by deviating surface orders: On this view, what counts for the Mirror Principle in both syntax and morphology is the initial order in which items are merged; this may subsequently then be undone by movement.

Needless to say, there are many features of the approach to inflectional morphology adopted here that would be in need of further discussion and justification. We cannot possibly do this in the present paper; see Müller (2018) for extensive discussion. However, in the final part of the paper, we would like to consider alternatives to morphological movement of inflectional exponents in harmonic serialism, as they have been proposed for some of the above phenomena in the literature.

5. Alternative Approaches

Several alternative approaches to non-local phonological operations have been proposed in the literature. They can be conveniently divided into three groups, based on their assumptions. Some approaches generally allow for non-local application of phonological processes; thus no problem occurs in the morphosyntactic component. Other approaches maintain the assumption that phonological processes apply only locally. They attempt to solve this problem by morphological operations (interfixation, lowering or dislocation), or by syntactic movement, accompanied by a modified version of the spell-out mechanism. These three groups of approaches are in turn discussed in this final section.

5.1. *Non-Local Phonology*

One possible approach to the phenomenon of exceptional non-locality in phonological processes is to take the evidence at face value. If phonological processes are potentially non-local, a theory could just allow for all phonological processes to be non-local. Such approaches are rare in generative phonology for reasons that will be discussed below. One such non-local approach to phonology within a more general optimality-theoretic frame-

³⁸ It can be noted that the present approach to the interleaving of phonology and morphology is incompatible with any notion of bracket erasure (cf. Chomsky & Halle (1968), Pesetsky (1979), and Kiparsky (1982a)). However, all the alternatives discussed in section 5 below, with the exception of non-local phonology, share this property; so we would like to contend that this incompatibility is strongly suggested by the data.

³⁹ See, e.g., Embick (2010) on verb marker allomorphy in Palauan.

work is Agreement by Correspondence (ABC; cf. Hansson (2001), Rose & Walker (2004), and Rhodes (2012)).

Bowman & Lokshin (2014) propose an account for exceptionally transparent vowels in Kazakh vowel harmony based on ABC. The gist of this analysis is that vowel harmony is always a non-local phenomenon that affects all vowels in a word. If the constraint now applies to all vowels except the comitative case suffix, we generate the existing pattern. In this subsection, we will first introduce the specific approach and then discuss its application to other cases of exceptionally non-local phonological processes.

An ABC analysis consists of two different constraint types. One constraint type CORR requires all segments of a certain type to be in a correspondence class. Correspondence classes are marked with a subscript index. In the present case, this is a CORR-VV constraint that requires all vowels to be in a correspondence class; cf. (57-a). A second constraint type places a certain condition on this correspondence class. In our case, the vowels should agree in backness, so the constraint is ID-VV(back), as given in (57-b).

(56) *Kazakh vowel harmony constraints by ABC* (Bowman & Lokshin (2014, 7))

- a. CORR-VV: Let S be an output string of segments and let X and Y be segments specified [-consonantal,+sonorant]. X and Y correspond if $X, Y \in S$.
- b. IDENT-VV(back): Let X be a segment in the output and let Y be a correspondent of X in the output. If X is [α back], then Y is [α back].

Ranking these constraints above a general ID(back) constraint already derives the harmony. The directionality of vowel harmony is derived by a high ranked faithfulness for the initial syllable INITID(back). The analysis is illustrated in the tableau in (57). The CORR-VV constraint excludes all candidates with non-corresponding vowels, e.g. the candidate in (57-a). IDENT-VV(back) on the other hand penalizes candidates with non-agreeing vowels, as shown by (57-c). Finally, the INITID(back) constraint derives directionality by protecting the first vowel, cf. (57-d).

(57) *Regular vowel harmony in Kazakh by ABC* (Bowman & Lokshin (2014, 7))

I ₀ : ij-tar	INITID(back)	ID-VV(back)	CORR-VV	ID(back)
O ₁ ij-tar			*!*	
☞ O ₂ i _x j-te _x r				*
O ₃ i _x j-ta _x r		*!		
O ₄ u _x j-ta _x r	*!			

In order to derive the exceptional transparency of the comitative case suffix, Bowman & Lokshin (2014) assume a modified system of phonological strength. In their case, strength is lexical/morphological instead of contrastivity-based, as in the original proposal by Rhodes (2012). Vowels can be marked as strong or weak in the underlying representation of a word. In Kazakh, all vowels but the exceptionally transparent one are marked as strong with respect to backness harmony. The CORR-VV constraint is then modified to only apply to strongly specified vowels CORR-V_SV_S. Since this idea of strength is stripped of its phonological meaning in earlier literature, it now only applies lexically to a certain class of morphemes. Thus, it can be seen as a special case of lexical constraint indexation (see Pater (2007)),⁴⁰ dividing the morphemes into undergoers and non-undergoers for vowel harmony, but with a twist. The constraint only applies if both targets of the constraint bear an index. This

⁴⁰ Bowman & Lokshin (2014) have to assume constraint indexation for independent reasons anyway.

is different from the usual locality conditions in indexed constraint approaches (see Pater (2009)). Weak vowels are marked with a superscript obelisk in the following tableau in (58). Since the general version of this constraint is now ranked below the IDENT(back) constraint, the comitative suffix vowel cannot undergo vowel harmony; cf. candidate (58-e).

(58) *Exceptional transparency in Kazakh by ABC* (Bowman & Lokshin (2014, 8))

I ₀ : nan-men [†] -be	INITID(back)	ID-VV(back)	CORR-V _S V _S	ID(back)	CORR-VV
O ₁ nan-men [†] -be			*!*		
☞ O ₂ nan _x -men [†] -ba _x				*	*
O ₃ nan _x -men [†] -be _x		*!		*	*
O ₄ nan _x -men [†] _x -ba _x		**!		*	*
O ₅ nan _x -man [†] _x -ba _x				**!	

This account works for Kazakh vowel harmony because its domain is large enough. Vowel harmony affects all vowels in a word and ignores intervening consonants. Other processes discussed so far usually operate only under segmental adjacency (Aramaic, Sanskrit) or syllable adjacency (Quechua, Lithuanian). In order to account for these processes, an additional rankable locality constraint would have to be introduced. One version of this constraint would have to be ranked above, and one below the faithfulness constraint. This constraint would have to penalize intervening segments/syllables, with the higher ranked version of this constraint referring only to morphemes marked as strong and the lower ranked version of this constraint referring to all morphemes. In his early work, Hansson (2001, 237) explicitly denies the possibility of such a constraint: “Intervening segments do not themselves enter into the agreement relation holding between the trigger-target pair, and therefore they must be irrelevant to that relation [...]”. Even the locality constraints later introduced by Rose & Walker (2004) and formalized by Bennett (2015a) do not count intervening segments.⁴¹ Instead, they are violated for each pair of corresponding segments/syllables as soon as one non-corresponding element intervenes (cf. also Shih & Inkelas (2018, 141)). An account of more local phonological processes would thus require either a reformulation of this constraint or the addition of a new constraint type.

ABC has been mostly applied to consonantal and vocalic harmony or dissimilation processes (see Hansson (2001), Rose & Walker (2004), Rhodes (2012), and Bennett (2015a)). Therefore, segments in a correspondence class here mostly need to agree with respect to a certain feature, and such constraints do already exist in the system. The processes discussed in the data above however also include other processes, such as epenthesis and accent shift. It is at least not immediately clear how an ABC account would handle these.

There are further conceptual issues raised by this approach. For one thing, the notion of modularity has to be given up or considerably weakened. Constraints in the phonological part of the grammar have to make reference to the identity of certain morphemes. Phonology has to make access to morphological information. If strength of specification is really a phonological feature, it acts as a diacritic second-order feature. It does not have any other purpose or phonological substance apart from distinguishing undergoers and non-undergoers of vowel harmony. The account thus arguably comes close to merely restating the facts.

Second, non-locality in phonology has to be assumed as a default. Local processes only occur if some locality constraint is ranked high enough. This contrasts with the empirical evidence that can be gained from a crosslinguistic perspective. Most phonological processes

⁴¹ However, see Bennett (2015b, 61) for a proposal where constraints are supposed to count intervening syllables.

are local: non-local processes like consonant harmony are vanishingly rare.

5.2. *Counter-Cyclic Movement in Morphology and Interfixation*

In Distributed Morphology, there are two mechanisms of post-syntactic movement in morphology – viz., Lowering and Local Dislocation – that could be employed to account for the data in section 2 without resorting to non-local phonology. Interfixation, on the other hand, is a pre-theoretical concept that bears some resemblance to Lowering. Let us start with the latter.

Interfixation is an obvious alternative to our proposal. It stands for a specific morphological process that introduces an exponent between two other exponents that have already merged in the structure. As noted above, such an approach has been pursued by Hyman (1994; 2002; 2003) for seemingly non-local phonological operations in the derivational morphology of Bantu languages; also see Kiparsky (1982b; 2017), Kushnir (2018), and, to some extent, Myler (2013) for analyses of this type. If the Y model of grammar is adopted together with post-syntactic morphology, and if it is assumed that derivational morphology is not realizational (in Stump’s (2001) terms), interfixation must take place in the narrow syntax – if it was to apply at PF, the interfix would not be interpreted semantically as it was not part of the structure before. Interfixing incurs a clear violation of the Strict Cycle Condition (see (28) above), because it extends a tree below the root projection. Since interfixation as a grammatical operation has (to the best of our knowledge) so far not been formally made precise, further discussion of the concept is not entirely straightforward; however, we take the problem raised by the inherent violation of the Strict Cycle Condition to be a severe one.

The avatar of interfixation in the framework of Distributed Morphology (cf. Halle & Marantz (1993), Embick & Noyer (2001)) is Lowering. In contrast to interfixation, where the interfix comes out of the lexicon or the workspace, a lowered morpheme is first merged cyclically in narrow syntax. At PF, this morpheme undergoes an operation of Lowering which moves the head of a higher phrase downward, onto the head of its complement (see Embick & Noyer (2001, 561)). Lowering faces the same problem of counter-cyclicity as interfixation – downward movement violates strict cyclicity because it alters a proper subdomain of the highest cyclic domain.⁴² However, Lowering is not only problematic because of the conceptual issue of strict cyclicity, it also fails to derive the patterns discussed in this paper. Lowering is generally assumed to apply before vocabulary insertion, i.e., morphological exponence, in Distributed Morphology (see Embick & Noyer (2001)). As a consequence, at the point in the derivation when Lowering applies, there are only abstract morphemes consisting of nothing more than bundles of morpho-syntactic features – there is not yet a phonological side to them. This entails that the morphemes in their underlying non-lowered order cannot interact phonologically. Lowering is hence able to account for counter-scopal or otherwise unexpected orders of exponents, but fails to derive the non-local phonological patterns for systematic reasons.

Another mechanism proposed in Distributed Morphology to account for deviant exponent order is Local Dislocation (LD; cf. Embick & Noyer (2001)). Unlike Lowering, LD does not operate on hierarchical structures, but on strings of exponents after vocabulary insertion and linearization. It applies at a stage of the derivation where the morphemes have their full morpho-syntactic and phonological features accessible (see Embick (2007)). LD is conceived of as a rule that transforms a relation of adjacency into a relation of adjunction,

⁴² This conclusion should be uncontroversial as such from a more general point of view (see, e.g., Pullum (1992)); and it follows directly from the interpretation of strict cyclicity in terms of the Extension Condition; see footnote 14. Given the actual wording in (28), lowering to an embedded position must be assumed to “affect solely” a proper subdomain of the current domain.

which can result in a change of the order of the two elements. The following discussion is confined to the notion of LD in Embick & Noyer (2001), Embick (2007), and Embick & Noyer (2007).

Compared to our approach, a LD-based analysis faces three problems. First, like Lowering or interfixation, it is not compatible with strict cyclicity; however, at least at first sight, the problems here might be somewhat less severe: If one follows Embick & Noyer's (2001) claim that there is no hierarchical structure at the point of LD, the Strict Cycle Condition as defined in (28) (which talks about domains, subdomains, and, most importantly, domination) does not strictly speaking apply. However, Embick and Noyer's contention seems debatable: Their approach to LD relies on (i) the concept of embedding, and (ii) the concept of a head of a morphological constituent during the application of LD. This entails that there is hierarchical structure at this point after all. Embick (2007, 322) explicitly states that different representations, among them hierarchical syntactic structure, exist in parallel at the point of the derivation where LD applies. This makes the Strict Cycle Condition in its original version applicable, as the structure is not yet flattened. If one were to modify LD so that it does indeed exclusively apply on a flattened string, it would not violate the Strict Cycle Condition in the exact form adopted so far, since after flattening there are no domains left that could stand in a structural domination relation. However, if a slightly more general definition of the Strict Cycle Condition is adopted (as in, e.g., Perlmutter & Soames (1979), Freidin (1992), or, for that matter, Pullum (1992)), where no outside material can be put into a discrete morpho-syntactic object, the Strict Cycle Condition is violated by any type of LD.

The second problem is the lack of a well-defined trigger for dislocating exponents. The trigger for LD to a position between two words is a constraint that demands that a certain word should morpho-phonologically pattern as an affix. The trigger for LD to a position between affixes is rather unclear (see Embick & Noyer (2007)). Embick (2007, 327) contemplates the idea that LD between affixes might be a different operation altogether. An overarching assumption is that arbitrary and/or unmotivated operations can be tolerated post-syntactically in Distributed Morphology (cf. Embick & Noyer (2007, 321)); then again, Embick & Noyer (2007, 320), propose that LD in Huave might be triggered by a linearization constraint that is a bit more complex than, but in fact to some extent comparable to, the alignment constraints employed in the analyses in sections 3 and 4 above. More generally, we would like to contend that a systematic formalization of the triggers in an LD-based approach may lead to mechanisms producing effects that are indeed similar to the constraints that we employ.⁴³

The third and most important problem is an empirical one: LD is ultimately not able to account for the phonological long-distance dependencies highlighted in section 2. On the one hand, the scope of LD in Embick's (2007) version does not extend to the data under discussion. Embick (2007, 325) introduces the *Consistency* condition, which demands that adjacency requirements may not contradict each other. This condition excludes non-string vacuous LD inside morphological words, or the reordering of affixes (see Embick (2007, 327)). This version of LD is thus not able to account for the data because what is needed is clearly a re-ordering of affixes. On the other hand, the version of LD in Embick & Noyer (2001; 2007) does not include Consistency, but it nonetheless fails to derive the data. Its problem lies in the way the morphology-phonology interactions regarding LD are envisioned. To see this, let us consider, as a representative case study of LD between affixes, the positioning of the reflexive exponent *si* in Lithuanian (cf. Embick & Noyer (2001; 2007), Embick (2007)). This affix appears in second position, between the verb stem and

⁴³ See Trommer (2001) and Rolle (2019) for proposals that modify Distributed Morphology along these lines.

a first prefix if there is one (*iš-si-laik-aũ*, ‘I hold my stand’)). However, if there is no other prefix, *si* is a suffix to the verb and all other verbal suffixes (*laik-aũ-si*, ‘I get along’). Prima facie, in the latter case LD is expected to move *si* from its underlying position preceding the verb to a position between the verb and its inflectional affixes, since this is the available second position. However, this is not what happens: *si* follows the inflectional affixes in this case (**laik-si-aũ*). Embick & Noyer (2001) assume that the reason that LD does not target an intermediate position here is that there has been previous string-vacuous LD of suffixal material onto the verb, so that suffix(es) and stem form a morphologically unbreakable unit. This string-vacuous LD is justified by the ‘closer phonological affinity to the stem’ (Embick & Noyer (2001, 580)) of suffixes, as opposed to prefixes, in Lithuanian. Embick (2007, 324) defines a close phonological relationship as exhibiting word-level phonology. Crucially, phonological processes like de-spirantization in Aramaic or vowel harmony in Kazakh are clearly word level phonology. For them to apply locally, a Distributed Morphology account parallel to the derivation of Lithuanian *si* must include an operation of string vacuous LD of the underlying inner suffix onto the stem. The exponent cluster created by LD should then be inseparable; it cannot be broken up by later applications of LD (see Embick & Noyer (2001, 580)). The conclusion then is that the LD operation that is responsible for the application of the phonological process should in turn block the application of LD that would be responsible for the attested affix order. This is exemplified for Kazakh data like (23) (*nan-men-ba*, ‘bread-INSTR-Q’) with the derivation in (59).

- (59) (i) structure after vocabulary insertion [[[nan]bA]men]
(ii) string-vacuous LD and vowel harmony [[nanba]men]
(iii) re-ordering LD of *men* *[[mennanba]

As the first relevant step, vocabulary items are inserted at their respective positions. As a second step, the affix *-bA* dislocates string-vacuously, so that it forms a close morpho-phonological unit with the root *nan* and can undergo vowel harmony. For further morphological operations, *nanba* now forms a single unit. Finally, LD of the outer exponent *men* can now only invert *men* with this complex unit *nanba*, yielding the incorrect **men-nanba* instead of the required *nan-men-ba*.

Notwithstanding the three potentially problematic issues just discussed, it can be noted that a LD-based analysis and the analysis developed in this paper make different empirical predictions. Since LD has access to both phonology and morphology, it can in principle be triggered by both morphological and phonological requirements. In the present approach, morphology and phonology are interleaved, but strict modularity is maintained. Phonological information is thus not expected to be able to trigger movement in morphology.

5.3. Movement in Syntax

Another account of non-local phonological processes is developed in Myler (2017). It shares with the present approach the idea that seemingly non-local phonological processes may in fact apply locally at some point of the derivation. This is derived by a combination of (i) syntactic movement and (ii) a novel algorithm of vocabulary insertion that makes it possible to spell out a specifier before the head of the projection is spelled out. In this subsection, we will argue that this innovative and elegant proposal nevertheless suffers from two significant problems: First, it predicts phonological interactions between non-adjacent words, which to the best of our knowledge is not attested; second, it employs syntactic movement without a clearly identifiable trigger.

Myler (2017) starts with the observation that seemingly non-local phonological processes can arise if an exponent order violates the Mirror Principle and suggests that morphemes are

first merged in accordance with the Mirror Principle, but the order is then obscured by movement; in this respect, Myler’s (2017) approach is thus a predecessor of the present analysis. The spell-out algorithm is set up in such a way that the resulting structure makes local application of seemingly non-local phonological processes possible. In particular, spell-out proceeds from the bottom to the top of a tree, and depth of embedding is counted by a number of categorially distinct maximal projections rather than just nodes. Here is the exact rule (see Myler (2017, 102)):

(60) *Temporal order of vocabulary insertion*

For a pair of terminal nodes x and y :

If x is the head of a maximal projection M such that M is categorially distinct from y and M dominates y , then $y \} x$.

If $y \} x$, then y undergoes vocabulary insertion prior to x .

This algorithm generates $Z \} Y$ and $X \} Y$ for the structure in (61). This means that vocabulary insertion applies to nodes X and Z first and only then to Y . X and Z are thus adjacent on the vocabulary item level before insertion of Y , and phonological operations applying between them can qualify as local in some sense.

(61) [_{YP} [_{ZP} ... Z] [_{Y'} Y [_{XP} X ...]]]

If Z is the root, X and Y are affixes, and ZP is moved from the complement of X to the specifier of Y , then (60) derives a possible phonological interaction between the root Z and the linearly non-adjacent affix X merged closer to Z in the latter’s base position. So far, so good. However, the spell-out algorithm applies in the same way if the structure is not derived by movement. It predicts phonological interactions between a specifier of Y and a complement of Y to be always possible across the linearly intervening head Y . In addition, Z , Y and X may be separate words, not just parts of one word. The model thus would seem to wrongly predict non-local interactions between non-adjacent words in configurations that are not derived by movement.

Turning to the second potential problem, Myler (2017) shows that (60) correctly derives non-local application of the *ruki* rule in Sanskrit, where the final vowel of the prefix *abhi* turns the initial consonant of the root into a retroflex ṣ , across the intervening past tense prefix *a*. Since Sanskrit data have also been discussed in the present paper (cf. sections 2.4 & 4.5), let us compare the two analyses. Myler assumes the structure in (62): A past tense augment prefix is viewed as the head of a designated AugP, and the first prefix *abhi* is moved to its specifier. According to the spell-out algorithm, *abhi* undergoes vocabulary insertion before *a*. This allows seemingly non-local retroflexivization to apply locally. Note that the order in which the affixes are merged, as well as the movement of the first prefix, are more or less exactly as in our approach, and that there is a point in the derivation before movement when there is no intervening material between the prefix *abhi* and the root. It is the stage at which a first phonological cycle is triggered in our analysis. Myler (2017), however, assumes that phonological interaction between prefixes is postponed to a later stage, where it is derived by manipulating the order of vocabulary insertion operations.

(62) [_{AugP} [_{PP} [_P *abhi*]] [_{Aug'} [_{Aug} *a*] ... [_{VP} [_{PP} [_P ~~*abhi*~~]] [_V *siñ c*]]]]

This (important) difference notwithstanding, a crucial difference between the two analyses relates to the motivation for movement of *abhi* in (62). There is general agreement that syn-

tactic movement does not apply without restrictions.⁴⁴ Myler suggests that *abhi* is a clitic and should be licensed in the specifier of AugP, but the relation between a past tense augment and syntactic licensing of a clitic is arguably somewhat stipulative. Similarly, an attempt to extend Myler's (2017) analysis to the other data discussed in section 2 is faced with the challenge of finding a simple syntactic trigger. In effect, then, the approach based on syntactic movement is subject to the same criticism as the post-syntactic approaches discussed in the previous subsection. And (abstracting away from the problem with base-merged specifiers noted above) it may well be that the same kind of way out might be available, viz., postulating general linearization constraints as triggers for movement as we have suggested them in sections 3 & 4, which in turn would render the approach more similar to our proposal.

To sum up the results of this subsection, given that alternative approaches to the phenomena we have been concerned with in the present study all seem to encounter major problems (of different provenance), we take it that there is every reason to look for a new approach that avoids these problems on the one hand, and is conceptually simple and theoretically principled on the other. Morphological movement that is triggered by independently motivated alignment constraints in harmonic serialism, and that is systematically interleaved with phonological cycles, strikes us as a promising candidate that meets these requirements.

References

- Abels, Klaus (2012): *Phases. An Essay on Cyclicity in Syntax*. Vol. 543 of *Linguistische Arbeiten*, De Gruyter, Berlin.
- Ackema, Peter & Ad Neeleman (2003): Context-Sensitive Spell-Out, *Natural Language and Linguistic Theory* 21, 681–735.
- Ackema, Peter & Ad Neeleman (2004): *Beyond Morphology*. Oxford University Press, Oxford.
- Adelaar, Willem & Pieter C. Muysken (2004): *The Languages of the Andes*. Cambridge University Press, Cambridge.
- Adger, David (2003): *Core Syntax*. Oxford University Press, Oxford, New York.
- Alexiadou, Artemis & Gereon Müller (2008): Class Features as Probes. In: A. Bachrach & A. Nevins, eds., *Inflectional Identity*. Oxford University Press, Oxford, pp. 101–155.
- Aronoff, Mark (1994): *Morphology by Itself*. MIT Press, Cambridge, Mass.
- Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller & Philipp Weisser (2015): Ergatives Move Too Early. On an Instance of Opacity in Syntax, *Syntax* 18, 343–387.
- Baker, Mark (1985): The Mirror Principle and Morphosyntactic Explanation, *Linguistic Inquiry* 16, 373–415.
- Bennett, William (2015a): Assimilation, Dissimilation, and Surface Correspondence in Sundanese, *Natural Language and Linguistic Theory* 33(2), 371–415.
- Bennett, William (2015b): *The Phonology of Consonants: Harmony, Dissimilation and Correspondence*. Vol. 147, Cambridge University Press, Cambridge.
- Bermúdez-Otero, Ricardo (2008): *Stratal Optimality Theory*. Book Ms., University of Manchester. To appear: Oxford University Press.
- Bills, Garland D., Bernardo Vallejo & Rudolph C. Troike (1969): *An Introduction to Spoken Bolivian Quechua*. University of Texas Press, Austin.
- Bobaljik, Jonathan (2012): *Universals in Comparative Morphology. Suppletion, Superlatives, and the Structure of Words*. MIT Press, Cambridge, Mass.

⁴⁴ Even in approaches where movement is not directly triggered by features or constraints but may apply freely (see, e.g., Chomsky (1981; 2008)), massive overgeneration must be avoided by imposing restrictions, e.g., in the form of filters (which may then be internal or external to syntax).

- Božič, Jurij (2019): Constraining Long-Distance Allomorphy, *The Linguistic Review* 36, 485–505.
- Bowman, Samuel & Benjamin Lokshin (2014): Idiosyncratically Transparent Vowels in Kazakh. In: *Proceedings of the 2013 Annual Meeting on Phonology*. Linguistic Society of America, Washington, DC.
- Bruening, Benjamin (2017): Consolidated Morphology. Book Ms., University of Delaware.
- Caballero, Gabriela & Alice Harris (2012): A Working Typology of Multiple Exponence. In: F. Kiefer et al., eds., *Current Issues in Morphological Theory: (Ir)Regularity, Analogy, and Frequency*. Benjamins, Amsterdam, pp. 163–188.
- Caballero, Gabriela & Sharon Inkelas (2013): Word Construction: Tracing an Optimal Path Through the Lexicon, *Morphology* 23, 103–143.
- Caha, Pavel (2013): Explaining the Structure of Case Paradigms by the Mechanisms of Nanosyntax, *Natural Language and Linguistic Theory* 31, 1015–1066.
- Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.
- Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.
- Chomsky, Noam (2000): Minimalist Inquiries: The Framework. In: R. Martin, D. Michaels & J. Uriagereka, eds., *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.
- Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.
- Chomsky, Noam (2008): On Phases. In: R. Freidin, C. Otero & M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.
- Chomsky, Noam (2019): Lectures. UCLA & MIT.
- Chomsky, Noam & Morris Halle (1968): *The Sound Pattern of English*. MIT Press, Cambridge, Mass.
- Collins, Chris & Edward Stabler (2016): A Formalization of Minimalist Syntax, *Syntax* 19(1), 43–78.
- Elfner, Emily (2016): Stress-Epenthesis Interactions in Harmonic Serialism. In: *Harmonic Grammar and Harmonic Serialism*. Equinox, Sheffield.
- Embick, David (2007): Linearization and Local Dislocation: Derivational Mechanics and Interactions, *Linguistic Analysis* 33, 303–336.
- Embick, David (2010): *Localism vs. Globalism in Morphology and Phonology*. MIT Press, Cambridge, Mass.
- Embick, David & Rolf Noyer (2001): Movement Operations after Syntax, *Linguistic Inquiry* 32, 555–595.
- Embick, David & Rolf Noyer (2007): Distributed Morphology and the Syntax/Morphology Interface. In: *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press, Oxford, pp. 289–324.
- Ermolaeva, Marina & Greg Kobele (2019): Agreement and Word Formation in Syntax. Ms., University of Chicago & Universität Leipzig.
- Frampton, John (2002): Syncretism, Impoverishment, and the Structure of Person Features. In: M. Andronis, E. Debenport, A. Pycha & K. Yoshimura, eds., *Papers from the Chicago Linguistics Society Meeting*. Vol. 38, Chicago, pp. 207–222.
- Freidin, Robert (1992): *Foundations of Generative Syntax*. MIT Press, Cambridge, Mass.
- Georgi, Doreen (2014): Opaque Interactions of Merge and Agree. PhD thesis, Universität Leipzig.
- Grimshaw, Jane (2001): Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT

- Press, Cambridge, Mass., pp. 205–240.
- Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.
- Halle, Morris & Jean-Roger Vergnaud (1987): *An Essay on Stress*. MIT Press, Cambridge, Mass.
- Hansson, Gunnar (2001): Theoretical and Typological Issues in Consonant Harmony. PhD thesis, University of California, Berkeley.
- Harris, Zelig (1945): Discontinuous Morphemes, *Language* 21, 121–127.
- Heck, Fabian & Gereon Müller (2007): Extremely Local Optimization. In: E. Brainbridge & B. Agbayani, eds., *Proceedings of the 26th WECOL*. California State University, Fresno, pp. 170–183.
- Heck, Fabian & Gereon Müller (2013): Extremely Local Optimization. In: H. Broekhuis & R. Vogel, eds., *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 135–166.
- Heck, Fabian & Gereon Müller (2016): On Accelerating and Decelerating Movement: From Minimalist Preference Principles to Harmonic Serialism. In: G. Legendre, M. Putnam, H. de Swart & E. Zaroukian, eds., *Optimality-Theoretic Syntax, Semantics, and Pragmatics*. Oxford University Press, Universität Leipzig, pp. 78–110.
- Hyman, Larry (1994): Cyclic Phonology and Morphology in Cibemba. In: J. Cole & C. Kisseberth, eds., *Perspectives in Phonology*. CSLI, Stanford, pp. 81–112.
- Hyman, Larry (2002): Cyclicity and Base Non-Identity. In: D. Restle & D. Zaefferer, eds., *Sounds and Systems. Studies in Structure and Change. A Festschrift for Theo Vennemann*. De Gruyter, Berlin, pp. 223–239.
- Hyman, Larry (2003): Suffix Ordering in Bantu: A Morphocentric Approach, *Yearbook of Morphology* 2002, 245–281.
- Julien, Marit (2002): *Syntactic Heads and Word Formation: A Study of Verbal Inflection*. Oxford University Press, Oxford.
- Khan, Geoffrey (2008): *The Neo-Aramaic Dialect of Barwar. Volume One: Grammar*. Brill, Leiden, Boston.
- Kiparsky, Paul (1982a): From Cyclic Phonology to Lexical Phonology. In: H. van der Hulst & N. Smith, eds., *The Structure of Phonological Representations*. Vol. 1, Foris, Dordrecht, pp. 131–175.
- Kiparsky, Paul (1982b): Lexical Morphology and Phonology. In: I.-S. Yang, ed., *Linguistics in the Morning Calm*. Hanshin Publishing Company, Seoul, pp. 3–91.
- Kiparsky, Paul (2017): Lectures on Strata in Phonology and Morphology. Ms., Stanford University. (Brugmann Lectures, Universität Leipzig).
- Kobele, Greg (2006): Generating Copies. PhD thesis, UCLA, Los Angeles.
- Krippes, Karl (1996): *Kazakh Grammar with Affix List*. Dunwoody Press, Kensington.
- Kushnir, Yuriy (2018): Prosodic Patterns in Lithuanian Morphology. PhD thesis, Universität Leipzig.
- Mahanta, Shakuntala (2012): Locality in Exceptions and Derived Environments in Vowel Harmony, *Natural Language and Linguistic Theory* 30(4), 1109–1146.
- Marvin, Tatjana (2002): Topics in the Syntax and Stress of Words. PhD thesis, MIT, Cambridge, Mass.
- Matthews, Peter (1972): *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge University Press, Cambridge.
- McCarthy, John (2008): The Serial Interaction of Stress and Syncope, *Natural Language and Linguistic Theory* 26, 499–546.
- McCarthy, John (2010): An Introduction to Harmonic Serialism, *Language and Linguistics*

- Compass* 4, 1001–1018.
- McCarthy, John (2016): The Theory and Practice of Harmonic Serialism. In: J. McCarthy & J. Pater, eds., *Harmonic Grammar and Harmonic Serialism*. Equinox, Sheffield, pp. 47–87.
- McCollum, Adam (2018): The Empirical Consequences of Data Collection Methods: A Case Study from Kazakh Vowel Harmony, *Linguistic Discovery* 16(2), 72–110.
- Merchant, Jason (2015): How Much Context is Enough? Two Cases of Span-Conditioned Allomorphy, *Linguistic Inquiry* 46, 273–303.
- Moskal, Beata & Peter Smith (2016): Towards a Theory Without Adjacency: Hyper-Contextual VI-Rules, *Morphology* 26, 295–312.
- Muhamedova, Raihan (2015): *Kazakh: A Comprehensive Grammar*. Routledge, London.
- Müller, Gereon (2011): Syncretism without Underspecification in Optimality Theory: The Role of Leading Forms, *Word Structure* 4(1), 53–103.
- Müller, Gereon (2018): *Inflectional Morphology in Harmonic Serialism*. Book ms., Universität Leipzig. To appear in the series *Advances in Optimality Theory*, London: Equinox. (<http://home.uni-leipzig.de/muellerg/harm/imhs.pdf>).
- Myler, Neil (2013): Linearization and Post-Syntactic Operations in the Quechua DP. In: T. Biebrauer & I. Roberts, eds., *Challenges to Linearization*. Mouton de Gruyter, Berlin, pp. 171–210.
- Myler, Neil (2017): Exceptions to the Mirror Principle and Morphophonological ‘Action at a Distance’. In: H. Newell, M. Noonan, G. Piggott & L. Travis, eds., *The Structure of Words at the Interfaces*. Oxford University Press, Oxford.
- Noyer, Rolf (1992): Features, Positions, and Affixes in Autonomous Morphological Structure. PhD thesis, MIT, Cambridge, Mass.
- Oltra Massuet, Isabel & Karlos Arregi (2005): Stress-by-structure in Spanish, *Linguistic Inquiry* 36, 43–84.
- Ouhalla, Jamal (1991): *Functional Categories and Parametric Variation*. Routledge, London.
- Pater, Joe (2007): The Locus of Exceptionality: Morpheme-Specific Phonology as Constraint Indexation, *University of Massachusetts Occasional Papers in Linguistics* 32.
- Pater, Joe (2009): Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In: S. Parker, ed., *Phonological Argumentation: Essays on Evidence and Motivation*. Equinox, London, pp. 123–154.
- Perlmutter, David & Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.
- Pesetsky, David (1979): Russian Morphology and Lexical Theory. Ms., MIT. (Generals Paper).
- Pesetsky, David & Esther Torrego (2006): Probes, Goals and Syntactic Categories. Ms., MIT. Lingbuzz/000321.
- Prince, Alan & Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell, Oxford.
- Pullum, Geoffrey (1992): The Origins of the Cyclic Principle. In: *Proceedings of the Regional Meeting of the Chicago Linguistic Society*. Vol. 28.2, pp. 209–236.
- Rhodes, Russell (2012): Vowel Harmony as Agreement by Correspondence, *UC Berkeley Phonology Lab Annual Report* 8.
- Rolle, Nicholas (2019): In Support of an OT-DM Model: Evidence from Clitic Distribution in Degema Serial Verb Constructions, *Natural Language and Linguistic Theory* Published online, 1–59.
- Rose, Sharon & Rachel Walker (2004): A Typology of Consonant Agreement as Correspon-

- dence, *Language* 80, 475–531.
- Shih, Stephanie & Sharon Inkelas (2018): Autosegmental Aims in Surface-Optimizing Phonology, *Linguistic Inquiry* 50, 137–196.
- Starke, Michal (2001): Move Dissolves Into Merge: A Theory of Locality. PhD thesis, University of Geneva.
- Stiebels, Barbara (2006): Agent Focus in Mayan Languages, *Natural Language and Linguistic Theory* 24, 501–570.
- Stiebels, Barbara (2015): Multiple Exponence: Typology, Triggering Factors, and Theoretical Modelling. Ms., Universität Leipzig.
- Stump, Gregory (2001): *Inflectional Morphology*. Cambridge University Press, Cambridge.
- Svenonius, Peter (2016): Spans and Words. In: H. Harley & D. Siddiqi, eds., *Morphological Metatheory*. Benjamins, Amsterdam, pp. 199–220.
- Torres-Tamarit, Francesc (2016): Compensatory and Opaque Vowel Lengthening in Harmonic Serialism. In: *Harmonic Grammar and Harmonic Serialism*. Equinox, Sheffield.
- Trommer, Jochen (2001): Distributed Optimality. PhD thesis, Universität Potsdam.
- Trommer, Jochen (2008): Coherence in Affix Order, *Zeitschrift für Sprachwissenschaft* 27(1).
- Trommer, Jochen (2011): Phonological Aspects of Western Nilotic Mutation Morphology. Habilitation thesis, Universität Leipzig.
- van der Hulst, Harry (2016): Vowel Harmony. In: M. Aronoff, ed., *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.
- Weber, David (1976): *Suffix-as-Operator Analysis and the Grammar of Successive Encoding in Llacon (Huanuco) Quechua*. Instituto Lingüístico de Verano, Yarinacocha.
- Weber, David (1989): *A Grammar of Huallaga (Huanuco) Quechua*. University of California Press, Berkeley.
- Wolf, Matthew (2008): Optimal Interleaving. PhD thesis, University of Massachusetts, Amherst.
- Wunderlich, Dieter (1997): A Minimalist Model of Inflectional Morphology. In: C. Wilder, H.-M. Gärtner & M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 267–298.