Note:

The following is a long version of a manuscript which has been accepted to appear in *Phonology* in a significantly shorter form. The two versions will not be substantially different in terms of content, but the current version goes into considerably more detail.

– Adam Jardine
September 1, 2015

Computationally, tone is different*

Adam Jardine

(long version, short version to appear in *Phonology*)

Abstract

This paper establishes a typological asymmetry between tone and segmental phonology and characterises the asymmetry using a notion of computational complexity taken from Formal Language Theory. UNBOUNDED CIRCUMAMBIENT PROCESSES, phonological processes in which triggers or blockers appear unboundedly far away on both sides of a target, are common in tone but rare in segmental phonology. The evidence for this is based around attestations of Unbounded Tone Plateauing (UTP; Hyman 2011, Odden & Kisseberth 2003), but it is also shown how the 'sour grapes' harmony pathology (Baković 2000) is unbounded circumambient. The paper argues that such processes are not WEAKLY DETERMINISTIC, which contrasts with previous typological work finding segmental phonology to be at most weakly deterministic. Positing that weak determinism bounds segmental phonology but not tonal phonology thus captures the typological asymmetry. It is also discussed why this explanation is superior to any offered by Optimality Theory.

1 Introduction

This paper establishes and then characterises a typological difference between segmental and tonal phonology: UNBOUNDED CIRCUMAMBIENT PROCESSES are well-attested in tone but rare in segmental phonology. Briefly, an unbounded circumambient process is one in which triggers or blockers appear on *both sides* of a target, and there is *no bound*, on *either side*, on the distance between these triggers or blockers and the target. This paper argues that such processes are computationally more complex than those which are commonly attested in segmental phonology.

---

*The author would first like to thank Jeffrey Heinz for going over countless versions of this paper, as well as three anonymous reviewers and associate editor who provided thoughtful and detailed critiques and advice. Thanks also go to Larry Hyman, William Idsardi, Kevin Ryan, the computational linguistics group at the University of Delaware, the members of Jeffrey Heinz's 2012 phonology seminar, and an inquisitive audience at NECPHON 7, all of whom provided valuable feedback over the course of the project. Although many helped with this paper, all errors are my own.

An unbounded circumambient process in tonal phonology that will be central to this paper is UNBOUNDED TONAL PLATEAUING (henceforth UTP; Kisseberth and Odden, 2003; Hyman, 2011), in which any number of tone-bearing units in between two underlying high tones also become high. A simple example from Luganda (Hyman et al., 1987; Hyman and Katamba, 2010; Hyman, 2011) is given below, with high toned vowels (both underlying and surface) marked with an acute accent (ó) and the 'plateau' underlined:

(1) UTP in Luganda (Hyman, 2011, p.231 (52))
/bikópo byaa-walúsiimbi/ → bikópó byáá-wálúsiimbi
cups of-Walusimbi
'the cups of Walusimbi'

UTP is an unbounded circumambient process because the triggering H tones can be any distance away from the affected tone-bearing units. Hyman (2011) observes that UTP is commonly attested in tone, but similar plateauing effects are, with one exception, unattested in segmental phonology. The first contribution of this paper is to document this asymmetry in detail and to show that it is in fact part of a larger generalisation: unbounded circumambient processes are well-attested tone, but extremely rare in segmental phonology.

To highlight this, UTP is compared with with Sour Grapes harmony (Baković, 2000; Wilson, 2003; McCarthy, 2010; Heinz and Lai, 2013), a 'pathological' process predicted by variants of Optimality Theory (henceforth OT; Prince and Smolensky, 1993, 2004) in which target vowels harmonize with a preceding trigger only if there is no following blocking segment. It is shown how Sour Grapes vowel harmony is also an unbounded circumambient process, and that while it is unattested in segmental phonology, a Sour Grapes-like pattern is attested in tone. However, the typological generalization is an *asymmetry*, and not a universal. The two known possible cases of segmental unbounded circumambient processes, Sanskrit *n*-retroflexion and plateauing in KiYaka vowel harmony, are presented and discussed in detail.

The second contribution is to show that UTP and Sour Grapes, by virtue of being unbounded circumambient, are *formally* similar, in terms of computational complexity as measured by Formal Language Theory (FLT). FLT defines complexity classes of input-output MAPPINGS which we can use to categorize phonological processes. This paper shows that UTP is not SUBSEQUENTIAL (Mohri, 1997), as has also been shown for Sour Grapes (Heinz and Lai, 2013). This is illustrated by showing that it is not describable with a DETERMINISTIC FINITE-STATE TRANSDUCER. Intuitively, this is because, as UTP is an unbounded circumambient process, each target must be able to 'look ahead' in either direction to see crucial information in the environment. It is also argued unbounded circumambient processes are outside of the WEAKLY DETERMINISTIC class of mappings, which allow restricted bidirectional lookahead, and thus are FULLY REGULAR, i.e. only describable with NON-DETERMINISTIC finite-state transducers.

The third contribution of this paper is then to understand the unbounded circumambient asymmetry in terms of these classes of mappings. The complexity of unbounded circumambient processes is contrasted with previous work applying FLT to phonology, which has found local and long-distance unidirectional segmental processes to be subsequential (Chandlee, 2014; Chandlee et al., 2012; Heinz and Lai, 2013; Payne, 2014) and unbounded bidirectional processes (in which a feature 'radiates' outwards in two directions, e.g. Arabic emphasis spreading) to be weakly deterministic. The unbounded circumambient asymmetry can thus be captured in terms of a complexity bound on segmental phonology: segmental phonology is restricted to weakly deterministic mappings, but tone is not. This captures the generalisation that segmental phonology exhibits local, unbounded unidirectional, and unbounded bidirectional processes but not—save for two exceptions meriting further study—unbounded circumambient ones.

The logical structure of this paper is as follows. §2 establishes the empirical generalisation that unbounded circumambient processes are more common in tone than in segmental phonology. §3 introduces the FLT notions of complexity that have been applied to phonology and shows how most segmental processes are subsequential. §4 shows that UTP, along with Sour Grapes and other circumambient processes, is not subsequential, and defends the linear representation used to obtain this result. §5 discusses Heinz and Lai (2013)'s conjecture that Sour Grapes is not weakly deterministic, and argues that unbounded circumambient processes in general are not weakly deterministic. This leads to the computational characterisation proposed by this paper: segmental processes are at most weakly deterministic, but tone is not restricted in this way. It is also discussed how Sanskrit *n*-retroflexion and KiYaka vowel harmony fit into this proposal. §6 discusses how Optimality Theory does not offer a unified way of characterising the typological asymmetry. §7 concludes, and mathematical definitions and a proof that UTP is neither left- nor right-subsequential are given in an appendix (§8).

2 The Unbounded Circumambient Asymmetry

This section defines UNBOUNDED CIRCUMAMBIENT PROCESSES and shows how they are common in tonal phonology but extremely rare in segmental phonology. CIRCUMAMBIENT is used here to refer to a process whose application is dependent on the existence of triggers or blockers on both sides of a target; an UNBOUNDED circumambient process is one in which there is *no bound*, on *either side*, on the distance between these triggers/blockers and the target. These terms are discussed in more detail below in §2.1.

The bulk of the evidence for the asymmetry comes from Unbounded Tone Plateauing (Kisseberth and Odden, 2003; Hyman, 2011), which was noted by Hyman (2011) as a tonal process with no common correlate in segmental phonology. §2.2 surveys attestations of variants of UTP in the tonal literature, as a thorough documentation of Hyman (2011)'s claim. §2.3 reviews two known segmental unbounded circumambient processes of mid-vowel harmony in KiYaka (Hyman, 1998, 2011) and *n*-retroflexion in Sanskrit (Whitney, 1889; Macdonell, 1910; Schein and Steriade, 1986; Hansson, 2001; Graf, 2010).

§§2.4 and 2.5 summarize related generalizations in other typological work which support the conclusion that unbounded circumambient processes are rare in segmental phonology, and discuss why Sanskrit and KiYaka are exceptional for these generalizations as well. §2.5 also discusses how the asymmetry is not confined to UTP but really the *class* of unbounded circumambient processes, by way of Sour Grapes vowel harmony (Wilson, 2003; McCarthy, 2010; Heinz and Lai, 2013), which in Copperbelt Bemba H-spreading (Bickmore and Kula, 2013) has an attested correlate in tone.

2.1 Definition of 'unbounded circumambient'

A precise definition of UNBOUNDED CIRCUMAMBIENT PROCESS is given in (2). Crucially, this property is atheoretical and thus agnostic to specific theories of representation and processes, as will be shortly discussed in more detail.

(2) An UNBOUNDED CIRCUMAMBIENT PROCESS is a process for which:

a. its application is dependent on information (i.e., the presence of a *trigger* or *blocker*) on *both* sides of the target, and

b. *on both sides*, there is no bound on how far this information may be from the target

While this definition is atheoretical, to help illustrate the concept, a rewrite rule representation of this type of process is given in (3). In (3), X and Y are nonempty, they surround the target, and there is no bound on the distance between them.

(3) Shape of unbounded circumambient processes in a rule-based framework:
$A \quad \rightarrow \quad B \quad / \quad X(U)\_\_(V)Y$ (X and Y are nonempty, U and V may be of any length)

Save for the 'unbounded' part, this is actually a very standard-looking rule. Take, for example, an intervocalic voicing rule like in (4):

(4) [−sonorant] → [+voice] / V __ V

The rule in (4) crucially refers to the existence of vowel triggers on both sides of the target; intervocalic voicing is thus a circumambient process. It is, however, *bounded*, because only vowels adjacent to the target count as triggers. *Unbounded* circumambient processes, in contrast, seem less natural. An example is the imaginary rule in (5):

(5) [+syl] → [+back] / $\begin{bmatrix} \text{+syl} \\ \text{+back} \end{bmatrix}$ (U) __ (V) $\begin{bmatrix} \text{+syl} \\ \text{+back} \end{bmatrix}$ (U and V are any string of segments)

For the rule in (5), whether a [$-$back] vowel becomes [$+$back] crucially depends on the presence of two [$+$back] vowels on both sides of the target. However, these [$+$back] vowels may be separated from the target segment by the strings X and Y. Because X and Y can be of any length, when the process applies, it must be able to 'look behind' for a [$+$back] segment *over any distance* in the left context (i.e., over X), and it must also be able to 'look ahead' over any distance for a [$+$back] vowel in the right context (i.e., over Y). Exactly what 'any distance' means will be qualified shortly in Section 2.1.1.

The rule representations in (3) through (5) are purely illustrative, and the idea of unbounded circumambient processes is broader than what can be represented in linear rules. For example, the 'crucial information' in the environment is more general than the traditional rewrite rule environment—as we will see, X or Y might contain information about the presence of a blocker. An imaginary such rule is given in (6):

(6) [$-$sonorant] $\rightarrow$ [$+$nasal] / [$+$nasal] (U) __
(except in this situation: [$+$nasal] (U) __ (V) [$+$nasal])

Here, the target will nasalise only in the case in which it is preceded by a [$+$nasal] segment and *not* followed by another [$+$nasal] segment. While for brevity (6) has been written out as a rule, such conditions are more intuitively expressed with Optimality Theory (henceforth OT; Prince and Smolensky, 1993, 2004) constraints than rule-based formalisms.

Furthermore, the definition of unbounded circumambient applies equally well to autosegmental representations. From an autosegmental standpoint, the 'target' referred to in (2) is any unit affected by the changing of association lines and the distance from the 'trigger' is measured on the timing tier. These choices will become clear in §4.4.

2.1.1 What is unbounded?

Because 'unbounded' is critical to the definition in (2), is important to set the criteria by which we decide a phonological process is unbounded. Intuitively, an unbounded process is one which operates over multiple units, like segments or tone-bearing units (TBUs), for which the correct generalisation does not refer to a bound on how many units over which it may operate. As linguists may differ as to what constitutes evidence for a process having 'no bound,' this paper considers the criteria in (7):

(7)
a. The source authors characterise the process as unbounded, and there is no evidence to the contrary
b. Examples exist of the process operating over multiple units
c. Examples exist of the process applying even when productive word or phrase formation processes extend its domain

On its own, criterion (7a) does not constitute strong evidence for a process being unbounded, and thus no processes are included here which only meet (7a).

The criterion in (7b) refers to a process operating over more than one unit. For this criterion, the longer the span to which a process is seen to apply, the more convincing a data point is as an example of an unbounded process. Take the following data from Johore Malay nasal spreading (Walker, 1998, citing Onn 1980), in which any number of sonorant segments following a nasal are nasalised:

(8) Johore Malay (Walker, 1998, (4))

a. baŋõn ‘to rise’
b. mãkan ‘to eat’
c. mãj̃ãŋ ‘stalk (palm)’
d. pəŋãw̃ãsan ‘supervision’

In (8c) and (d), the process operates on targets separated from their triggers by two units; ex., there are two segments [ãw̃] intervening between the the second [ã] and the (original) nasalisation trigger [ŋ]. Because of this, Johore Malay nasal spreading satisfies criterion (7b), and most phonologists would have no trouble considering it ‘unbounded.’ The processes discussed in this paper have examples where the trigger and target are separated by at least two units, and thus satisfy (7b).

Even more convincing evidence for the unboundedness of a process is (7c), in which there are examples of a process applying over a domain whose length can be increased. For example, in Chizigula verbs (Kenstowicz and Kisseberth, 1990), an underlying H tone shifts rightward to the penult, no matter how many morphemes are introduced. In the following, a surface H is marked with an accute accent [á], and the position of the underlying H is marked with an underline:

(9) Chizigula (Kenstowicz and Kisseberth, 1990, p.166)

a. ku-lo̲mbéz-a ‘to ask’
b. ku-lo̲mbez-éz-a ‘to ask for’
c. ku-lo̲mbez-ez-án-a ‘to ask for each other’

Because the shifting rule applies no matter how many morphemes extend its domain, it can be considered ‘unbounded’ under (7c). Furthermore, H shift in Chizigula applies even when a noun is added to the verb phrase:

(10) Chizigula (Kenstowicz and Kisseberth, 1990, p.172,5)

a. ku-fi̲s-íz-a ‘to hide for’
b. ku-fi̲s-a ma-tungúja ‘to hide tomatoes’
(cf. ku-guha ma-tunguja ‘to take tomatoes’)

Thus, the rule also applies regardless of morphosyntactic operations extending its domain. This is the sense of ‘unbounded’ captured by (7c). As we will see, many attestations of

Unbounded Tone Plateauing are unbounded by this criterion, but neither of the circumambient segmental processes discussed are.

For some, (7b) is enough evidence that a process is unbounded, especially if there are examples of it operating over three or more units—as Kenstowicz (1994) put it, "phonological rules do not count past two" (p.372). In contrast, some researchers may only consider (7c) sufficient. However, the evidence presented here shows there is a typological asymmetry regardless of which criterion one considers. By either (7b) or (7c), unbounded circumambient processes are far more common in tone than in segmental phonology.

One final technical point about measuring unboundedness is that distance will be measured relative to the particular kind of unit targeted by a process. Thus, for tone, it will be measured in TBUs, in vowel harmony in vowels, and in consonantal processes in segments. Slightly different choices could have been made here (e.g., counting consonants in vowel harmony), but these would have no effect on the main results.

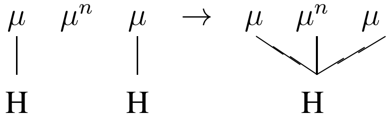
Having established the criteria for classifying a process as unbounded and circumambient, we now turn to a survey of the languages with Unbounded Tone Plateauing, the main unbounded circumambient tonal process to be discussed in this paper.

2.2 Unbounded Tone Plateauing

This section surveys eight languages with some form of Unbounded Tone Plateauing (UTP, Kisseberth and Odden, 2003; Hyman, 2011), a phenomenon in which any number of L–toned or unspecified (∅) TBUs (here, TBU will be assumed to be the mora) surface as H if they are in between two Hs, but surface as L otherwise.

Kisseberth and Odden (2003) motivate UTP as a repair for a long distance constraint against "toneless moras between Hs" (p.67). Hyman and Katamba (2010) formalise UTP in Luganda (the process is referred to in that paper as 'H tone plateauing') as follows:

(11)



Essentially, UTP is an unbounded H-spreading process that is *only* triggered when another H is present further along in the domain. UTP thus fits the definition of an unbounded circumambient process because whether or not the process applies depends on two Hs that a) are on both sides of the affected TBUs and b) can be arbitrarily far away from any one of the affected TBUs.[1]

Using data from a number of sources, the following subsections establish Hyman (2011)'s claim that UTP is a well-attested tonal process, and that it is unbounded by criterion (7b) in all cases and by criterion (7c) in most. Data from Luganda (which is pointed out by Hyman) and Digo are examined in depth first, and then examples in other

---

[1] This can be contrasted with *bounded* plateauing, in which only one ∅ TBU becomes H in between two Hs. This pattern is attested, for example, in Kihunde (Goldsmith, 1990).

languages are briefly reviewed. Note that, with regards to distance between triggers and TBUs, UTP is ‘symmetrical’—that is, the first target in the plateau is the same distance from the right trigger as the last target in the plateau is from the left trigger. Thus, when measuring distance for criterion (7b), the maximum distance seen in any example from a target to the left trigger will also be the maximum distance seen from a target to the right trigger.

2.2.1 Luganda

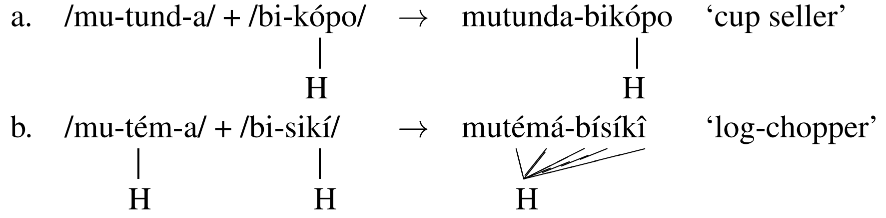
UTP is pervasive in Luganda (Hyman et al., 1987; Hyman and Katamba, 2010), occuring both word-internally and in the phrasal phonology. We begin with noun compounds. Luganda TBUs can be either H or unspecified (∅) underlyingly, and Luganda nouns may have no underlying tone or a H tone. Lexically, a L tone is inserted after a H, causing a falling HL contour when the H is on a final syllable. In some cases, these intermediate L tones contrast with ∅, but not with regards to UTP (technically, they delete in the UTP environment), so otherwise they will not factor into the discussion. Here, tones are marked using standard notation of an acute accent [á] for underlying and surface H tones, with underlying unspecified and surface low tones unmarked. The falling contour is marked with a circumflex on the vowel, [â].[2]

(12) Luganda nouns (from Hyman and Katamba, 2010, (2),(3),(26),(52))

a.	/ki-kópo/	kikópo	‘cup’
b.	/ki-sikí/	kisikî	‘log’
c.	/ki-tabo/	kitabo	‘book’
d.	/mu-tund-a/	mutunda	‘seller’ (from /-tund-/ ‘to sell’)
e.	/mu-tém-a/	mutéma	‘chopper’ (from /-tém-/ ‘to chop’)

If a toneless noun and a noun with an underlying H tone are put together in a compound, both nouns are pronounced as they are in isolation (13a). However, when both nouns have an underlying H, a ‘plateau’ of high tones occurs between them (13b). Autosegmental diagrams accompany the following examples to illustrate the process (ignoring the intermediate Ls). In compounds, the plural /bi-/ is used for nouns from the IV /ki-/ noun class. It carries no tone.

(13) Luganda compounds (Hyman and Katamba, 2010, (26) & (52))



a. /mu-tund-a/ + /bi-kópo/ → mutunda-bikópo ‘cup seller’

b. /mu-tém-a/ + /bi-sikí/ → mutémá-bísíkî ‘log-chopper’

---

[2] For clarity of exposition, the tones transcribed here are those of what Hyman and Katamba call the ‘intermediate’ forms before the superimposition of phrasal boundary tones. The boundary tones are inserted late in the derivation and play no role in the phonology.

In (13b) there is a plateau over three unspecified TBUs. This shows target toneless TBUs three TBUs away from their triggers—the mora (represented by) /a/ in /mu-tém-a/ 'chopper' is two TBUs from its right trigger H /í/ in /bi-sikí/ 'log,' and the second /i/ in /bi-sikí/ is three TBUs from its left trigger /é/ in /mu-tém-a/. Thus Luganda satisfies the 'multiple unit' criterion for unboundedness in (7b), and examples showing much longer distances between trigger and target will be shown momentarily.

UTP operates over syntactic phrases as well, satisfying (7c). Under certain morphosyntactic conditions (explicated in Hyman and Katamba (2010)), noun-verb sequences can also form a domain for UTP. In (14a) below, /walúsimbi/ 'Walusimbi (proper name)' is an adjunct, and thus does not form a phonological phrase with the verb. Save for H spreading onto the second TBU of the long vowel in [tw-áá-mu-lab-a] 'we saw him', the TBUs in between the two H tones surface as L (the two H tones in the verb that do not surface are changed to L by a version of Meussen's Rule). In (14b), however, the two words do form a phonological phrase, and plateauing occurs accross all TBUs in between the surviving Hs.

(14) Luganda verb+noun combinations (Hyman and Katamba, 2010, (14))

a. /tw-áa-mú-láb-a walúsimbi/ → tw-áá-mu-lab-a walúsimbi
| | | | |/ |
H H H H H H
'we saw him, Walusimbi'

b. /tw-áa-láb-w-a walúsimbi/ → tw-áá-láb-wá wálúsimbi
| | |
H H H H
'we were seen by Walusimbi'

A more extreme example can be seen when proclitics are added to the noun. Proclitics generally do not have tone, as the lack of H in the following data shows:

(15) Luganda toneless proclitics (Hyman and Katamba, 2010, (33))

a. byaa=ba=mulondo '(it's) those of the Mulondos (Mulondo and his people)'
b. na=ku=byaa=ba=mulondo 'and on those of the Mulondos'

Luganda UTP 'sees' over these as well, no matter how many are stacked on to the noun:

(16) Luganda verb+noun sequences with proclitics (Hyman and Katamba, 2010, (35))

a. /tw-áa-láb-a byaa=walúsimbi/ → tw-áá-láb-á byáá-wálúsimbi
| | |
H H H H
'we saw those of Walusimbi'

b. /tw-áa-génd-a na=byaa=ba=walúsimbi/ → tw-áá-génd-á ná=byáá=bá=wálúsimbi
| | |
H H H H
'we went with those of Walusimbi'

These last few examples show UTP operating over a six-TBU span of toneless TBUs created by prefixation, and thus triggers are separated from their targets by five TBUs on each side. As such, Luganda UTP also satisfies criterion (7c) for unboundedness, and satisfies (7b) with a value of 5 units.

2.2.2 Digo

Digo verbs (Kisseberth, 1984) show complex interactions between underlying H tones, including tonal plateauing. Underlyingly, Digo is a privative H/∅ system. Both verb roots and affixes may carry a H tone, although this is not obligatory. To illustrate, the verb root in (17a), /tsukur/ 'take,' is 'toneless'; it does not carry an underlying H, and thus is pronounced with all low tones when its affixes are also underlyingly toneless. When the third person plural object prefix /á/, which carries a H tone, is concatenated to the root in (17b), its tone surfaces as a rising/falling pattern on the final two TBUs. Kisseberth (1984) argues that this is the realisation of the underlying H tone associated with /á/, which has shifted to the end of the word. This is illustrated autosegmentally in (18) following the same scheme as in the preceding section.

(17) Digo (Kisseberth, 1984, (29))
a. ni-na+tsukur-a
'I am taking'
b. ni-na+a-tsukǔr-â
'I am taking them'

(18) /ni-na+á-tsukur-a/ →ni-na+a-tsukǔr-â
| \/
H H
'I am taking them (=17b)'

Kisseberth analyses the resulting rising/falling pattern as a late, second shift leftward of the H. This is unimportant to the issue at hand; it is simply important to note the initial shift of the H from the prefix to the end of the word.

The forms in (17) have the toneless first person prefix /ni/; if this is substituted for the third person singular subject prefix /á/, which also has an underlying H associated with it, a H-tone pleateau occurs from the object prefix to the end of the root:

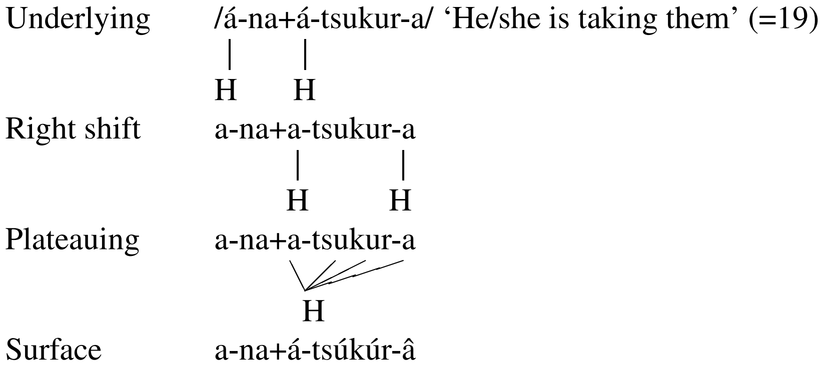
(19) a-na-á-tsúkúr-â 'He/she is taking them' (Kisseberth, 1984, (29))

Again, we see the presence of two H tones creating a long-distance plateau across the length of the root. The Digo example is complicated by the shifting of tones; Kisseberth analyses it as a two-step process in which the first H shifts to the initial vowel of a 'verbal complex' (marked with the '+' boundary), the second H shifts to the end of the word (as in (18)), and then the first H then triggering a plateau across the root between them.[3] This analysis is illustrated by the AP derivation in (20) below.

---

[3] That this is truly an interaction between the two processes of rightward shift and plateauing, and not simply rightward spreading triggered by two Hs, can be seen in forms with voiced obstruents. Voiced obstruents act as depressor consonants and block plateauing but have no effect on rightward shift. If the simple rightward spreading analysis were correct, we would not expect to see the rising/falling intonation on the final syllable (the manifestation of a shifted H) in forms with depressor consonants. Instead, we do,

(20)



The domain for both right shift and plateauing is actually larger than just the verb—they both also apply to verb+noun constructions. For brevity, surface tonal patterns are given here, with underlying H tones marked by underlining. In the following, (21) shows right shift applying when there is only one H in the verb, and (22) shows plateauing over the phrase resulting from two Hs associated with the verb.

(21) a. ku+a̲fŭnâ ‘to chew’
b. nazi ‘coconut’
c. ku+a̲fun-a năzî ‘to chew a coconut’ (Kisseberth, 1984, (63))

(22) a. a-ka̲+tsú̲kú ts-â ‘he has cleaned’
b. chi-ronda ‘wound’
c. a-ka̲+tsú̲kú ts-á chí-róndâ ‘he has cleaned a wound (Kisseberth, 1984, (65))

The phrase in (22c) results in a plateau over six TBUs, which shows target TBUs three TBUs away from their triggers and thus satisfies criterion (7b) for unboundedness. Also shown in (22c) is a plateau created over a span of (nearly) the length of two words, a domain provided by the phrasal syntax. As such, Digo UTP also satisfies criterion (7c).

This section has shown, from Luganda and Digo, two clear examples of UTP which satisfy all three unboundedness criteria in (7). The following section briefly describes UTP processes in other languages, all of which satisfy (7b) and most of which also satisfy (7c).

2.2.3 Other languages

Kisseberth and Odden (2003) cite a plateauing process in Xhosa very similar to Digo. In Xhosa, underlying H tones shift to the antepenult. A phrase with two H tones shows a plateau between the first H and the second, shifted H on the antepenult. Again, here surface forms are listed with UR H TBUs underlined:

---

as the following show (Kisseberth, 1984, p.135–137):
i. /á-na-dúndurik-a/ → a-ná-dundurǐkâ ‘he/she is walking stealthily’ (*a-ná-dúndurika)
ii. /á-na-ni-dúngir-a/ → a-na-ní-dungǐrâ ‘he/she is piercing for me’ (*a-na-ní-dúngir-a)

(23) Xhosa (Kisseberth and Odden, 2003, pp. 67-8)
a. u̲-ku-qonóndis-a ‘to emphasise’
b. ndi-fú̲n-a ‘I want’
c. ndi-fú̲n ú̲-kú-qónónóndis-a ‘I want to emphasise’

The second H shifts to the third /o/ in /ú-ku-qononondis-a/ ‘to emphasise,’ and becomes the right trigger for UTP; this shows triggers on both sides affecting targets four TBUs away. Kisseberth and Odden (2003) note that such a process is ‘common in the Nguni languages’ spoken in southern Africa, as do Cassimjee and Kisseberth (2001) (Digo, on the other hand, is spoken in northeast Africa in Kenya and Tanzania). In Zulu (Laughren, 1984; Cassimjee and Kisseberth, 2001; Downing, 2001), another Nguni language, a single H simply shifts to the antepenult (if it originates on a prefix) or penult (if it originates on a stem). In forms with two Hs, a plateau forms between the Hs. In Zulu, the two H tones do not fuse; instead, the first spreads up to the second, creating a downstep (marked with !):

(24) Zulu (Yip, 2002, p. 158, citing Laughren (1984))
a. i-si-hlaːlo ‘seat’ → i-sí-hlaːlo b. ámàkhòsánà → ámákhós!ánà (no gloss)

Laughren (1984) specifically states that “the rule only applies to a H which is followed by a LH tonal sequence” (p. 221; Yip represents the LH as a single H, which is then downstepped), and while she only gives examples of the plateau operating over two TBUs, analyses the process (which she calls High Tone Spreading) as operating over an arbitrary number of TBUs.

Other examples of UTP can be found throughout Bantu. In KiYaka (also known as Yaka; Kidima, 1990, 1991), “all toneless syllables flanked by Hs become H by rightward spreading of the H to the left domain” (Kidima, 1991, p.44). The following show plateauing alternations. Kidima (1991) gives KiYaka tonal assignment a complex accentual analysis; the underlying Hs marked in the following data result from a tonal assignment rule.[4] In the following examples, *bakhoko* ‘chickens’ is not assigned a tone, prompting the alternations, most notably in the initial and final vowels of *ba ngwaasi* ‘of uncle’ in (25a) and (c).

(25) KiYaka (Kidima, 1991, (68), p. 180)

a. bakhoko ba ngwaá̲si
chickens of uncle
‘Uncle’s chickens’

---

[4] A surface distinction between regular H and raised H—the latter occuring on accented syllables with an associated H—is ignored in these transcriptions.

b. bakhoko ba kabeé̲nga
chickens of red
‘The red chickens’

c. bakhoko ba kabeé̲ngá bá ngwáá̲sí Málóó̲ngi
chickens of red of uncle Maloongi
‘Uncle Maloongi’s red chickens’

Example (25c) shows two plateaus, each with their target TBUs separated from their triggers by two TBUS.

UTP also occurs outside of Bantu. In Saramaccan (Roundtree, 1972; Good, 2004; McWhorter and Good, 2012), a creole language spoken in Suriname, a phrasal version of UTP occurs accross words in certain syntactic configurations. This analysis follows Good (2004), who posits an underlying H/L/∅ distinction for TBUs in Saramaccan. The following nouns in isolation show different tonal configurations.

(26) Saramaccan nouns (Good, 2004)
a. wómì ‘man’ /wómi/
b. mùjɛ́ɛ̀‘woman’ /mujɛ́ɛ/
c. wàjàmákà ‘iguana’ /wajamáka/
d. sɛ̀mbɛ̀‘person’ /sɛ̀mbɛ̀/
e. àmɛ̀ɛ̀ká ‘American’ /amɛɛká/

According to Roundtree (1972), “all changeable low tones [=Good (2004)’s ∅ TBUs] between highs in successive morphs in certain syntactic positions are changed to high...” (p. 314). One such syntactic position is an adjective-noun sequence. In the following examples, the final /o/ in /hánso/ ‘handsome’ is realised as low [ó] before /sɛ̀mbɛ̀/ ‘person’ (27a) but high before /wómi/ ‘man’ (27b) and /mujɛ́ɛ/ ‘woman’ (27c). Relevant vowels are emphasised in bold.

(27) Saramaccan phrases
a. /dí hánso sɛ̀mbɛ̀/ → dí háns**ò** sɛ̀mbɛ̀
the hansome person “the hansome person” (Roundtree, 1972, p.315)
b. /dí hánso wómi/ → dí háns**ó** wómì
” ” man “the handsome man” (Roundtree, 1972, p.324)
c. /dí hánso mujɛ́ɛ/ → dí háns**ó** m**ú**jɛ́ɛ̀
” ” woman “the handsome woman” (Roundtree, 1972, p.316)
d. /dí wajamáka=dé á óbo/ → dí w**á**j**á**máká=dé á óbo
the iguana=there have eggs “the iguana there has eggs” (Good, 2004, p. 28)
e. /dí taánga amɛɛká wómi/ → dí tààng**á** **á**m**ɛ́ɛ́**ká wómì
the strong American man “the strong American man” (McWhorter and Good, 2012, p.48)

Note that in (27b) the initial /u/ of /mujɛ́ɛ/ 'woman' also surfaces as H, illustrating plateauing over two TBUs. In (27e) a plateau occurs over four TBUs, the final /a/ of /taánga/ 'strong' and the first three vowels of /amɛɛká/ 'American,' thus Saramaccan UTP satisfies (7b), with its targets and triggers separated by three TBUs on each side.

Similar to the system in Saramaccan (and noted by Good) is the intonational phonology of the Uto-Aztecan language Papago (Hale and Selkirk, 1987), in which H tones are associated "to each stressed vowel and to all vowels in between" and L tones are associated "to each unstressed vowel preceding the first stress in the tonal phrase" and "to each unstressed vowel following the last stress in the tonal phrase" (Hale and Selkirk, 1987, pp.152–153); they list examples of plateaus created in between stressed vowels three TBUs apart. Do and Kenstowicz (2011) also discuss plateauing in between Hs in certain intonational phrases in South Kyungsang Korean, giving a spectrogram showing a plateau between the two H tones in *seccók khaliphonía* 'Western California' (pp. 3 & 11), which shows unspecified TBUs affected by triggers two TBUs away (on either side).

2.2.4 UTP: summary

This section has surveyed a number of languages with some variation of UTP. Each satisfied the criteria for unboundedness in (7b). The attestations are, with the maximum number of TBUs seen in between target and trigger given in parentheses: Luganda (5); Digo (3); Xhosa (4); Zulu (2); Ki-Yaka (3); Saramaccan (3); Papago (2); South Kyungsang Korean (2). As UTP is 'symmetrical,' these distances apply to both the left and right trigger. Most of these examples were also phrasal in nature, also satisfying requirement (7c).

2.3 Unbounded circumambient processes in segmental phonology

This section discusses two unbounded circumambient processes in segmental phonology. These are the only two potential attestations of which the author is aware. One is Sanskrit *n*-retroflexion (Whitney, 1889; Macdonell, 1910; Schein and Steriade, 1986; Hansson, 2001; Graf, 2010; Ryan, 2015). The other is 'plateauing harmony' in KiYaka (Hyman, 1998, 2011), which also has UTP, as explained in §2.2.3. Both satisfy the unboundedness criterion in (7b), but neither satisfy (7c).

2.3.1 Sanskrit

Sanskrit *n*-retroflexion is a long-distance process in which an underlying alveolar /n/ becomes retroflex [ɳ] after retroflex /r,ʂ/, which can appear far to the left of the target /n/. In the following examples, both trigger and target will be highlighted with underlining:

(28) Sanskrit *n*-retroflexion (Hansson, 2001, p. 225, citing Schein and Steriade (1986))

	UR	SR	Gloss
a.	/iʂ̱-ṉaː-/	iʂ̱-ɳ̱aː-	‘seek (pres. stem)’
b.	/cakʂ̱-aːṉa-/	cakʂ̱-aːɳ̱a-	‘see (middle part.)’
c.	/kṟp-a-maːṉa-/	kṟp-a-maːɳ̱a-	‘lament (middle part.)’

The longest distance between trigger /r/ and target /n/ in (28) is five segments (four counting the long vowel as one segment), so the triggering of *n*-retroflexion satisfies criterion (7b) for an unbounded process. This long-distance process by itself is unidirectional, though, and thus would not fit the definition of an unbounded circumambient process in (2). However, one of the many restrictions on *n*-retroflexion may potentially give the process an unbounded circumambient quality.

Hansson (2001) states that retroflexion fails “when there is also an /ʂ/ or /r/ *later* in the word” (p.230, emphasis original).[5] He cites the following data from (Macdonell, 1910) (second, blocking /ʂ/ or /r/ also underlined; syllabic rhotics are not transcribed for typographic clarity):[6]

(29) Blocking of Sanskrit *n*-retroflexion (Hansson, 2001, p. 225, citing Macdonell (1910))

	Attested	Unattested	Gloss
a.	pṟaː-ṉṟtyat	*pṟaː-ɳ̱ṟtyat	from -nrt- ‘dance’
b.	paṟi-ṉakʂ̱ati	*paṟi-ɳ̱akʂ̱ati	‘encompasses’
c.	-ṉiʂ̱ʈʰaː-	never *-ɳ̱iʂ̱ʈʰaː-	‘eminent’
d.	-ṉiʂ̱ʂidʰ-	never *-ɳ̱iʂ̱ʂidʰ-	‘gift’
e.	-ṉiṟɳiɟa-	never *-ɳ̱iṟɳiɟa-	‘adornment’
f.	-ṉṟmɳa-	never *-ɳ̱ṟmɳaː-	‘manhood’
g.	pṟa-ṉaŋkʂ̱yati	*pṟa-ɳ̱aŋkʂ̱yati	‘causes to dissappear’ (Monier-Williams, 1899)

Given the data in (28), one would expect, for example, that the underlying /n/ in (29b) [pari-nakʂati] ‘encompasses’ would surface as [ɳ], because it follows a trigger /r/ for *n*-retroflexion. However, it instead surfaces as [n], which is attributed to the ‘blocking’ /ʂ/ three segments to the right.

Recall that the definition in (2a) for circumambient includes not just triggers, but blocking segments as well. Thus, Sanskrit *n*-retroflexion fits the definition for circumambient in a slightly different way than UTP, because information about the presence of both the triggers and blockers is crucial. The evidence in the examples here also fit criterion (7b) for unboundedness. As just mentioned, (29g) shows a blocker 3 segments away to

---

[5] Ryan (2015) points out that blocking only occurs when a root boundary lies between the trigger and the target, which is abstracted away from here. The reader is referred to Ryan (2015) to a detailed discussion of Sanskrit retroflexion.

[6] Many thanks to Kevin Ryan for pointing (29g) out to me and for enlightening me about Sanskrit in general.

the right of the target, and (28c) shows a trigger 5 segments to the left. Thus, at least by criterion (7b), Sanskrit *n*-retroflexion is an unbounded circumambient process. However, it should be noted that Ryan (2015)'s study of several Sanskrit corpora finds no examples of retroflexion being blocked when a long vowel or multiple syllables intervene between the target and blocker. In fact, he finds a few explicit counterexamples in which blocking fails over a long vowel, though due to the paucity of forms which test the generalisation he refrains from concluding whether this failure is "principled or accidental" (p. 28). Thus, while the evidence for Sanskrit *n*-retroflexion satisfies criterion (7b) for unboundedness given here, whether or not it is truly unbounded is the subject of some doubt.

Furthermore, while Sanskrit *n*-retroflexion satisfies (7b), the 'multiple unit' criterion for unboundedness, to the author's knowledge there is no documented evidence satisfying (7c), i.e., that the distance this second blocker can be from the target may be extended by a morphological or syntactic process. Both Whitney (1889) and Macdonell (1910) state that *n*-retroflexion occurs in some phrases, e.g. preposition-noun constructions. However, no evidence is given for the *blocking* of *n*-retroflexion occuring in such an environment, and so while there is evidence satisfying (7c) in one direction, no such evidence exists for the other direction.

2.3.2 KiYaka vowel harmony

The other attested unbounded circumambient segmental process is vowel height harmony in KiYaka. On the page following his description of UTP, Hyman (2011) cites KiYaka (Hyman, 1998, where it is referred to as Yaka) as a rare example of vowel 'plateauing.' In KiYaka, the initial vowel of the perfective suffix /ile/ lowers to a mid [e] when the vowel in the stem is also mid (30c and d below). Otherwise, a progressive harmony converts the final /e/ to [i] ([l] and [d] are in complementary distribution, with [d] occurring before [i], [l] elsewhere; thus the [l]/[d] alternation in (30) and (31). Furthermore, this /l/ turns to [n] following a root nasal; thus the allomorph [ene] in (31b)). The lowering of this middle /i/ to [e] does not happen to the applicative suffix [ila], which does not end in a mid vowel. The following examples show this alternation, with the /i/ to [e] change in question underlined:

(30) KiYaka mid-vowel 'plateauing' (Hyman, 2011, p. 501)

	Gloss	Root	+Applicative /ila/	+Perfect /ile/
a.	'obstruct'	/kik/	kik-ila	kik-idi
b.	'bind'	/kas/	kas-ila	kas-idi
c.	'pay attn.'	/keb/	keb-ila	keb-e̲le
d.	'clear brush'	/sol/	sol-ila	sol-e̲le

That this is part of a more general process lowering high vowels to mid if and only if they are between two mid vowels can be seen in (31). This process can take place at least over three vowels:

(31) KiYaka mid-vowel plateauing (Hyman, 1998, p. 19(6a,e,&f))

	Gloss	Stem	+Final Vowel /a/	+Perfect /ile/
a.	'to send'	/hit-ik/	hit-ik-a	hit-ik-idi
b.	'lower'	/bet-ilik/	bet-idik-a	bet-elek-ele
c.	'to do an about-face'	/kel-umuk/	kel-umuk-a	kel-omok-ene

As in UTP, the plateau of mid-vowels shows triggers on both sides of the target vowels. Examples (b) and (c) in (31) thus show triggers two vowels from the right and two vowels from the left from their targets. As such, KiYaka vowel harmony satisfies criterion (7b) for an unbounded process. However, according to Hyman (1998), plateauing alternations can only be seen with verb roots and the perfective /-ile/; it thus does not satisfy (7c).

2.3.3 Segmental processes: summary

This subsection has presented two potential examples of unbounded circumambient segmental processes, Sanskrit *n*-retroflexion and KiYaka vowel harmony. Both are circumambient and both satisfy the criterion for unboundedness given in (7b). However, to the best of the author's knowledge, no evidence satisfying (7c) exists for either.

2.4 Empirical summary

The preceding sections presented eight attestations of UTP, an unbounded circumambient process in tonal phonology, and two examples of separate unbounded circumambient processes in segmental phonology. All satisfied criterion (7b), with the greatest distance between target and trigger being 5 TBUs. Seven examples of UTP were shown to operate over domains extended by morphology or syntax (thus satisfying criterions (7c)). Neither segmental process satisfied (7c), and in KiYaka, the distribution of the process was quite limited. Both satisfied (7b), although in KiYaka the greatest attested distance between trigger and target was three vowels, and in Sanskrit there was no evidence for blocking beyond the syllable following the target.

Most importantly, to the best of the author's knowledge, these are the only examples of such processes in segmental phonology. This is notable, given the wide attestation of long-distance segmental processes in general, as documented in the comprehensive surveys on feature-spreading harmony (Rose and Walker, 2004), vowel harmony (Baković, 2000; Nevins, 2010), consonant harmony (Rose and Walker, 2004; Hansson, 2001, 2010), and consonantal disharmony (Suzuki, 1998; Bennett, 2013).

There thus is a typological asymmetry between tonal and segmental processes: unbounded circumambient patterns are extremely rare in segmental processes, but well-attested in tone, at the very least in the variants of UTP discussed. A comparison between proportions of circumambient unbounded processes found in typological surveys (such as those just mentioned) of segmental and tonal processes would be ideal, but comparable surveys of tonal processes, or even particular kinds of processes, do not exist (to the best of the author's knowledge). Regardless, the evidence reviewed in this paper clearly shows an asymmetry despite the absence of such surveys for tone.

It should be noted that bidirectional spreading processes, in which a feature spreads outwards from a single trigger in two directions, are common in segmental harmony. An example is Arabic emphasis spreading (Al Khatib, 2008), in which an emphatic gesture spreads in an unbounded fashion in both directions from an underlying emphatic segment. In the following example from Southern Palestinian Arabic, emphasis spreads to the left (32a), to the right (32b), and to both the left and right (32c) (the spread is blocked by high front segments, such as /j/).[7]

(32) Palestinian Arabic (Al Khatib, 2008, (1))

a. /balːaːsˤ/ → [bˤaˤlˤːaˤːsˤ] 'theif'
b. /sˤajːaːd/ → [sˤaˤjːaːd] 'hunter'
c. /ʔatˤfaːl/ → [ʔaˤtˤfˤaˤːlˤ] 'children'

Such bidirectional processes are not uncommon in segmental phonology; other examples include nasal spread in Capanahua and Southern Castillian (Safir, 1982), and stem-control analyses of vowel harmony (Baković, 2000) and consonant harmony (Hansson, 2001, 2010). While these processes apply in an unbounded fashion, and they operate in two directions, they only have one trigger, and thus are not circumambient.

2.5 Discussion: A larger generalisation

It is important to emphasize that the definition of unbounded circumambient is atheoretical, which allows us to compare processes irrespective of their analysis given a particular theory. Viewed in this way, the unbounded circumambient asymmetry unifies Hyman (2011)'s observation that unbounded plateauing effects are (nearly) unseen in segmental phonology with other researchers' characterizations of long-distance segmental processes.

For example, Hansson (2001) finds long-distance consonant harmony processes operate in one of two ways. One is left-to-right, which means a process can look ahead, but never *also* look behind. The other way is bidirectionally, in the root-dominance way described above, where one trigger harmonises outward in both directions. Similarly, Wilson (2003, 2006b), reviewing the typologies of nasal, emphasis, and vowel harmonies, characterises segmental spreading processes as 'myopic'. This means that even segmental spreading processes which affect multiple segments proceed in a local fashion, never 'looking ahead' beyond immediately adjacent segments. This is different for unbounded circumambient processes, which depend on information on both sides of the target, unboundedly far away. Thus, any myopic process is not an unbounded circumambient process, and any unbounded circumambient spreading process is necessarily not myopic.

---

[7] Thanks to an anonymous reviewer for providing this example.

Thus, neither spreading nor consonant harmony fits the definition of unbounded circumambient. This is particularly interesting because these processes all have different analyses in Optimality Theory—long-distance, correspondence-based agreement in the case of consonantal harmony (Rose and Walker, 2004; Hansson, 2001), and local agreement or 'feature sharing' in spreading processes (Wilson, 2003; McCarthy, 2010, inter alia)—but a property that remains constant is that the processes are not unbounded circumambient. Thus, Hansson (2001) and Wilson (2006b)'s generalisations, as well as Hyman (2011)'s observation that plateauing is almost unattested in segmental phonology, all are in concord with the cross-linguistic generalisation that unbounded circumambient processes are not well-attested in segmental phonology. It should be noted that this also means KiYaka vowel harmony and Sanskrit *n*-retroflexion are exceptions to the myopic spreading generalisation.[8] Instead of invalidating this generalisation, however, this fact highlights how atypical these two processes are.

One last piece of evidence for the unbounded circumambient asymmetry is related to this discussion. It regards the unattested 'sour grapes' vowel harmony pattern (henceforth 'Sour Grapes'; Wilson, 2003; McCarthy, 2010; Heinz and Lai, 2013).[9] Sour Grapes has received some attention in the literature because it is a non-myopic harmony pattern predicted to exist by ranking permutations of classic OT with AGREE constraints, but it is not attested in segmental harmony. Sour Grapes is also unbounded circumambient, and, as will be discussed momentarily, a Sour Grapes-like process appears in tone.

Sour Grapes works as follows. Given a spreading [+F] feature (which targets underlying [–F] segments) and a blocking feature [!F], [–F] segments become [+F] after another [+F], provided that there is no blocking segment, [!F], following in the word:

(33) a. [–F]^n → [–F]^n (no trigger, no blocker → no harmony)
b. ...[+F][–F]^n → ...[+F]**[+F]**^{n} (trigger, no blocker → **harmony**)
c. ...[+F][–F]^n[!F]... → ...[+F][–F]^n[!F]... (trigger, blocker → no harmony)

In other words, [+F] spreads if and only if it can spread all the way.[10] Note that Sour Grapes is not myopic: the spread of the [+F] has to 'look ahead' to see if there is a blocker before it can apply. As such, it is also an unbounded circumambient pattern, because the presence or absence of both [+F] and [!F] segments, which can be any distance apart, bears on the realisation of [–F] segments in between (this blocking aspect makes it circumambient in a similar way to Sanskrit *n*-retroflexion).

However, a Sour Grapes-like pattern does exist in tone. In Copperbelt Bemba (Bickmore and Kula, 2013, 2014a,b), underlying H tones undergo one of two spreading processes, bounded ternary spreading or unbounded spreading. The latter is blocked by the

---

[8] If, as Gafos (1999) and Hansson (2001) argue, *n*-retroflexion is best analysed as spreading.

[9] The term 'sour grapes', originally due to Padgett (1995), refers to behaviours under certain formulations of OT in which either all features in a particular domain assimilate or none do. See also McCarthy (2010).

[10] This characterisation is thanks to an anonymous reviewer.

presence of another H tone. In phrase-final forms, unbounded spreading applies to the rightmost H. (Copperbelt Bemba has an underlying privative H/∅ contrast.)

(34) Bemba unbounded spreading (Bickmore and Kula, 2013, (1)&(17))

	UR	SR	gloss
a.	/u-ku-tul-a/	ù-kù-tùl-à	'to pierce'
b.	/bá-ka-fik-a/	bá-ká-fíká	'they will arrive'
c.	/bá-ka-mu-londolol-a/	bá-ká-mú-lóóndólól-á	'they will introduce him/her'
d.	/tu-ka-páapaatik-a/	tù-kà-páápáátík-á	'we flatten'

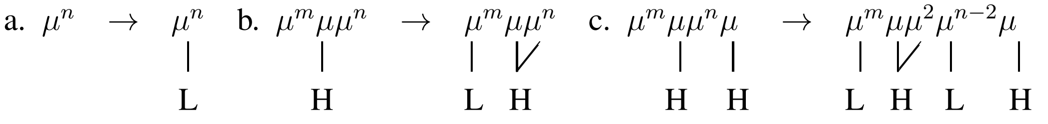
Bounded spreading occurs when another H appears to the right. Bounded spreading obeys the OCP; it will spread up to two additional TBUs, maintaining at least one L TBU before the second H. In the following examples, /kó/ is a post-verbal locative clitic. All other intervening TBUs surface with a L tone.

(35) Bemba bounded spreading (Bickmore and Kula, 2013, (18), 2014b, (9))

	UR	SR	gloss
a.	/bá-ka-pat-a kó/	bá-ká-pát-à kó	'they will hate'
b.	/bá-ka-londolol-a kó/	bá-ká-lóòndòlòl-à kó	'they will introduce them'
c.	/tu-ka-béleeng-el-an-a kó/	tù-kà-béléég-él-àn-à kó	'we will read for e.o'
d.	/tu-ka-lás-a Kapembuá/	tù-kà-lás-á Kápèèmbwá	'we will hit Kapembwa'

The formalisations in (36) summarise the facts. When no Hs are present, as in (36a), all TBUs surface as L (c.f. (34a) ù-kù-tùl-à 'to pierce'). When one H is present, it spreads to all remaining TBUs in the domain (and the rest surface as low, as in (34d)). When two Hs are present (36c), the first only spreads to the next two TBUs (c.f. (35c) tù-kà-bélééng-él-àn-à kó 'we will read for e.o').

(36)



$$\text{a. } \mu^n \rightarrow \begin{array}{c}\mu^n\\ |\\ \text{L}\end{array} \quad \text{b. } \begin{array}{ccc}\mu^m & \mu & \mu^n\\ & | & \\ & \text{H} & \end{array} \rightarrow \begin{array}{ccc}\mu^m & \mu & \mu^n\\ | & | & /\\ \text{L} & \text{H} & \end{array} \quad \text{c. } \begin{array}{cccc}\mu^m & \mu & \mu^n & \mu\\ & | & & |\\ & \text{H} & & \text{H}\end{array} \rightarrow \begin{array}{ccccc}\mu^m & \mu & \mu^2 & \mu^{n-2} & \mu\\ | & | & / & | & |\\ \text{L} & \text{H} & & \text{L} & \text{H}\end{array}$$

(c.f (34a)) (c.f (34d)) (c.f (35c))

Note that, modulo the bounded spreading, the formalisations in (36) are almost identical to the Sour Grapes generalisations in (33). In other words, in Copperbelt Bemba the second H can be seen as a blocker for unbounded spread. This makes it an unbounded circumambient process, because the realisation of unspecified TBUs depends on the presence or absence of Hs on *both* sides which can be arbitrarily far away. Thus, Copperbelt Bemba is a Sour Grapes-like pattern in tone. As a tonal process, it does not seem particularly aberrant, unlike Sanskrit *n*-retroflexion and KiYaka vowel harmony. Thus, it provides more evidence for the unbounded circumambient asymmetry.

2.6 Empirical conclusion

This section has made clear an asymmetry between tonal and segmental processes. Unbounded circumambient processes—i.e., processes in which crucial information about the environment lie arbitrarily far away on both sides of the target—are well-attested in tone, but extremely rare in segmental phonology. This section has also showed how Sour Grapes vowel harmony is like UTP: they are both unbounded circumambient processes.

The remainder of the paper is concerned with the following question: How can we understand this typological asymmetry? The following sections answer this question in the positive by arguing that any unbounded circumambient process has a certain level of computational complexity.

This characterization of the asymmetry is then argued in §6 to be superior to those offered by current theories of phonology, as they do not treat unbounded circumambient processes in a unified way.

3 The Computational Complexity of Phonological Mappings

This section introduces the computational concepts of complexity central to this paper and reviews previous work establishing that most segmental processes are SUBSEQUENTIAL. The section is structured as follows. §3.1 first introduces Formal Language Theory (FLT) and how complexity viewed in terms of FLT is relevant to phonology. §3.2 gives a broad overview of the particular complexity classes of interest to this paper—the REGULAR, LEFT- and RIGHT-SUBSEQUENTIAL, and the WEAKLY DETERMINISTIC mappings—and how they relate to phonology. §§3.3 and 3.4 then go into detail about how the subsequentiality of a phonological process can be determined using FINITE-STATE TRANSDUCERS and illustrate how most common segmental processes are subsequential.

3.1 Formal language complexity and cognitive complexity

The field of FLT started as a way to study the formal properties of natural language patterns by studying the relationships between STRINGSETS, or sets of strings, and the expressive power of grammars that describe them.[11] FLT characterisations of natural language patterns have been argued to reflect domain-specific cognitive biases for and against of patterns of a certain level of complexity. For example, one of the first results was that the REGULAR class of stringsets is insufficient to describe English syntax, and that English syntax is at least CONTEXT-FREE (Chomsky, 1956). Phonology, on the

---

[11] The term FORMAL LANGUAGE, or just simply LANGUAGE, is perhaps more commonly used in the FLT literature than STRINGSET, but this paper uses the latter term in order to avoid confusion with natural languages.

other hand, appears to be at most regular (Johnson, 1972; Kaplan and Kay, 1994).[12] All regular stringsets are context-free, but not all context-free stringsets are regular, so formalisms which can describe all context-free grammars must be more EXPRESSIVE than those that can only describe regular stringsets. Because this expressive power correlates with an increase in computational resources necessary for parsing and generating strings, the context-free stringsets are said to be more COMPLEX than the regular stringsets. FLT notions of complexity have been explicitly linked to cognitive complexity (Rogers and Pullum, 2011; Rogers and Hauser, 2010; Folia et al., 2010), and indeed, results from artificial language learning experiments provide evidence in support of the psychological reality of the context-free/regular division between syntax and phonology (Lai, 2012, to appear).

The goal of this paper is to use this notion of complexity to characterise the unbounded circumambient asymmetry discussed in §2. Instead of stringsets, however, the following sections study the relationships between regular MAPPINGS, which are string-to-string relations in which each input string is paired with at most one output string (although a single output may have multiple inputs).[13] Analagous to the regular/context-free division for stringsets, classes of mappings can be classified according to their complexity. The distinctions important for this paper center around the property of SUBSEQUENTIALITY (Mohri, 1997), which can be defined in terms of FINITE-STATE TRANSDUCERS (FSTs). FSTs are idealised machines that match pairs of strings; they are described in more detail below. While FSTs in general can describe any regular mapping, subsequential FSTs describe a more restricted set of mappings.

3.2 Overview: Formal language complexity and phonology

The classes of interest to this paper are the LEFT- and RIGHT-SUBSEQUENTIAL mappings, the WEAKLY DETERMINISTIC mappings, and the REGULAR MAPPINGS. Figure 1 below depicts the relationships between these classes as a nested hierarchy in which more complex classes properly include lesser ones. The left- and right-subsequential mappings

---

[12] Johnson and Kaplan and Kay show that SPE-style rewrite rules generate REGULAR RELATIONS, provided that they are not allowed to apply to their own output (for discussion of these results and why this restriction is empirically adequate see Heinz (2011)). A relation is not a stringset but a set of pairs of strings (and so similar to, but more general than, the mappings discussed in this paper), but because the set of outputs of a regular relation is a regular stringset (Rabin and Scott, 1959; Beesley and Karttunen, 2003), any generalisation about surface forms resulting from phonological processes is describable with regular stringsets. For more on the regular/non-regular split between phonology and syntax, see Heinz and Idsardi (2011, 2013).

[13] In cases of phonological free variation, one underlying representation may have more than one potential surface representation. Such a relation is not a mapping, as one input string is paired with multiple output strings. While the study of such relations in the framework presented here is important, the focus of the present study is on subclasses of *mappings*, and so free variation shall not be considered. For formal machines which are similar to the subsequential transducers introduced here but can handle (finite) free variation, see the P-SUBSEQUENTIAL TRANSDUCERS of Mohri (1997) or the SEMI-DETERMINISTIC TRANSDUCERS of Beros and de la Higuera (2014).

are provably less complex than the weakly deterministic mappings (Heinz and Lai, 2013) and the regular mappings (Mohri, 1997), and the weakly deterministic mappings have been conjectured to be less complex than the regular mappings (Heinz and Lai, 2013). Assuming this conjecture to be true (more on this below), this means that all weakly deterministic mappings are also regular mappings, but not all regular mappings are weakly deterministic. Mappings conjectured to be regular but not weakly deterministic will be sometimes referred to as FULLY REGULAR.

These classes and the relationships between them will be discussed in more detail below; for formal definitions and proofs of the relationships, the reader is referred to Heinz and Lai (2013). Importantly, while these relationships will be defined here in terms of FSTs, the complexity hierarchy they define is *independent* of their representation with FSTs, and they hold regardless of the formalism describing them (for example, a FST-independent characterisation of subsequential mappings is used in the Appendix).

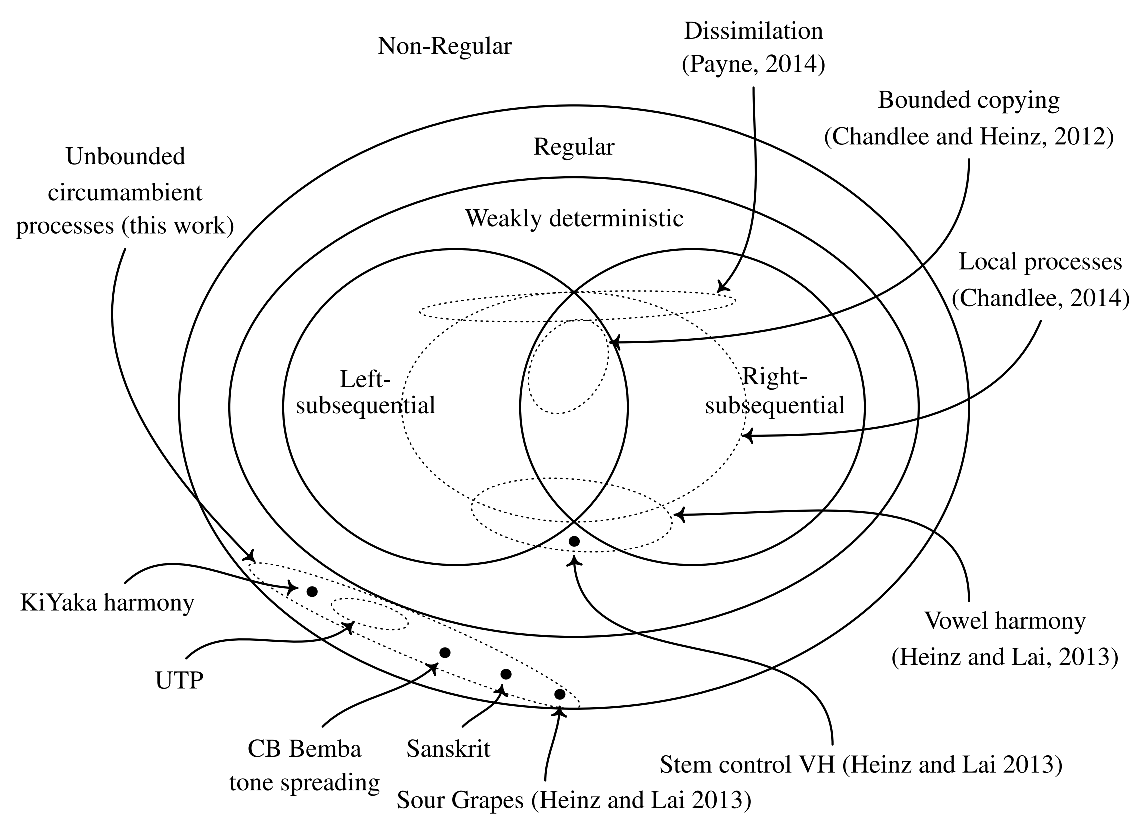


Figure 1: Phonology and the subregular classes

Figure 1 also depicts where typological research examining segmental processes as mappings place these processes in the complexity hierarchy. This work has found all of them to be within the weakly deterministic class, with most in the less complex left-and right-subsequential classes. This has led to the hypothesis, in the same vein as the discussion above, that that the weakly deterministic class forms a bound on the complexity of phonology. As Heinz and Lai (2013) discuss, this hypothesis is supported by the absence of Sour Grapes, which they prove to be neither left- nor right-subsequential and conjecture to not be weakly deterministic.

Figure 1 also shows the other unbounded circumambient processes discussed in this paper belonging to the fully regular region. As will be explained in detail in §4, this is because, like Sour Grapes, UTP and other unbounded circumambient processes are neither left- nor right-subsequential, and there is no known weakly deterministic characterisation of them. By revising the above hypothesis, then, we have a characterisation for the typological asymmetry established earlier: *segmental* phonology, but not *tonal* phonology, is at most weakly deterministic. The existence of Sanskrit *n*-retroflexion and KiYaka are admitted exceptions, but as already discussed in §2, they are not only rare but also exceptions to Wilson's myopia generalisation. These exceptions shall be discussed in §5.3.

This section and §§4 and 5 are devoted to a detailed discussion of these formal and empirical claims. The remainder of current section introduces FSTs and shows how most segmental processes are either left- or right-subsequential. This provides background for the main proof in §4 that UTP is not subsequential, a generalisation which is extended to other unbounded circumambient processes. §5 then argues that unbounded circumambient processes are also not weakly deterministic, as opposed to bidirectional spreading processes.

3.3 Finite-state transducers, subsequentiality, and determinism

This subsection introduces FSTs informally and discusses SUBSEQUENTIAL TRANSDUCERS, a type of FST which are strictly less expressive than the fully regular NONDETERMINISTIC TRANSDUCERS (Mohri, 1997). As subsequential transducers can describe left- and right-subsequential mappings, this provides the formal background necessary to understand the results of computational studies of segmental typologies discussed in §3.4 and the analysis of UTP in §§4 and 5.

First, let us look at how mappings work using an example mapping to model a very simple phonological process. The example below shows a highly simplified version of regressive nasal place assimilation, in which any underlying /n/ preceding a /p/ becomes an [m] in the output. Abstracting away from other (i.e., non-labial) consonants and vowels by simply representing them as 'C's and 'V's, (37) below highlights the difference in behavior we want: /n/ becomes [m] before a /p/, but not when /n/ precedes a /V/ or a /C/. To highlight that this is a mapping, and not a rule, the 'maps to' arrow ($\mapsto$) is used in this and following examples, instead of the rewrite rule arrow ($\rightarrow$).

(37) UR SR
...nV... ↦ ...nV...
...nC... ↦ ...nC...
...np... ↦ ...mp...

In this mapping, any 'n' preceding a 'p' in an input string will appear as a 'm' in the output string; this could just as easily be the result of rules as it is the result of an interaction between constraints. As the ellipses in (37) suggest, this is an *infinite* mapping, pairing any sequence of input 'C','V','n','p' to some sequence of 'C','V','n','m','p' where the nasal place assimilation generalisation holds. An expansion of the mapping represented in (37) is given in (38) below:

UR		SR	UR		SR
CVCV	↦	CVCV	CVnp	↦	CVmp
CVnC	↦	CVnC	VnpC	↦	VmpC
CnVC	↦	CnVC	VnpV	↦	VmpV
CCCC	↦	CCCC	npCC	↦	mpCC
...			...		

(38) A subset of the simple nasal place assimilation mapping

Note that input 'C's,'V's, 'p's, and and any 'n' that does not precede a 'p' are mapped to themselves in the output. As the last row shows, the mapping also includes strings entirely made of consonants. This is unthinkable in a real-world linguistic situtation, in which phonotactic constraints and other phonological processes would no doubt apply. However, this is no different from isolating a phonological generalisation with, for example, a SPE-style rewrite rule that is 'blind' to other rules in the phonology, or a partial OT constraint ranking capturing one specific phenomenon. With that said, let us see how this mapping can be exactly represented with a FST.

3.3.1 Defining the mapping with a FST

To see how FSTs pair inputs with outputs, let us use the nasal place assimilation mapping in (38) as an example. Figure 2 below gives a machine which describes this mapping.

A FST comprises a set of STATES (pictorially, the circles labeled with numbered qs) and TRANSITIONS (the labeled arrows) between the states. Transitions are triggered by particular symbols in the input string (shown on the left of the colon in the label), specify what symbol or symbols should be output when the transition is taken (shown on the right of the colon, so the labels are of the format '⟨input⟩:⟨output⟩'). Unlabelled arrows mark the START STATE from which the machine is allowed to begin reading a string; in Figure 2, the machine starts on state q_0. The transitions and states define what input/output string pairs are accepted by a machine. If, read sequentially, the input string takes the machine from the start state through a path of transitions to an ACCEPTING STATE (accepting states are marked by double circles), and all of the transition outputs match the output string, the

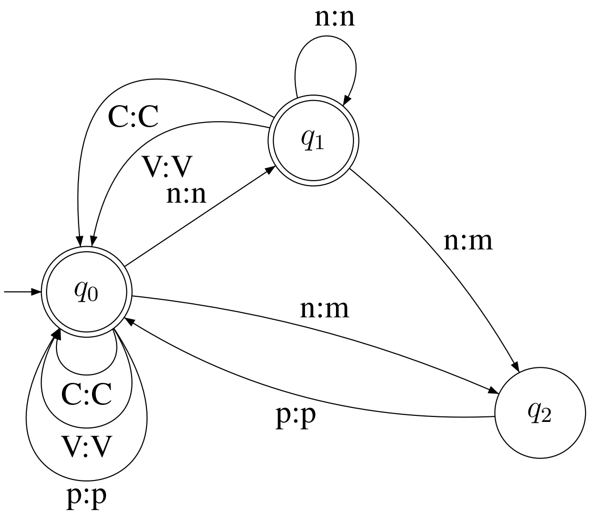


Figure 2: A nasal place assimilation FST

input/output pair is in the mapping described by the FST. If this fails, the input/output pair is not in the mapping. There are different varieties of FSTs, each with different properties which affect the class of mappings which they are able to describe. For example, in the FST in Figure 2, there are two transitions out of state q_0 with 'n' on the input side of their label. This means the machine is NONDETERMINISTIC. As will be discussed in more detail below, nondeterministic FSTs are more formally expressive than those that are deterministic, and thus they are able to describe more complex mappings.

To see how this machine works, let us walk through some simple examples from (38). According to the mapping in (38), an input 'CVnC' should return an output 'CVnC.' As the output string looks the same as the input string, this is a rather boring example, but it will be useful to see how the input 'CVnV' moves the machine from state to state, outputting 'CVnV' as it goes. A derivation is given below in (39), highlighting each state, input symbol, and output as the machine reads 'CVnC':

Input:		C		V		n		C	
State:	q_0	$\rightarrow$	q_0	$\rightarrow$	q_0	$\rightarrow$	q_1	$\rightarrow$	q_0
Output:		C		V		n		C	

(39) A derivation for CVnC $\mapsto$ CVnC in the FST in Figure 2

In the above derivation for CVnC $\mapsto$ CVnC, the machine first sees a 'C' and so it takes the 'C:C' transition, represented by the looped arrow on the machine in Figure 2, from state q_0 to itself. Recall that a label like 'C:C' means "take a C on the input and output a C." This first transition can be seen in the first step of the derivation in (39), where the machine starts at q_0, sees a 'C' in the input, outputs a 'C' and remains at q_0. The next 'V' in the input triggers the 'V:V' transition, and the machine again stays on q_0. Next, there is an ‘n’ in the input, which gives the machine two options. It can take the ‘n:n’ transition and move to q_1, or take the ‘n:m’ transition to q_2. Note, however, that from q_2, unless the next symbol in the input is ‘p,’ the machine cannot make a transition. As the symbol following ‘n’ is a ‘C,’ the only possible path is to first take ‘n:n’ to q_1 and then ‘C:C’ back to q_0. ‘CVnC’ has now been (trivially) transformed to ‘CVnC.’

It is clear, then, that if the input is instead ‘CVnp,’ the presence of the ‘p’ in the input forces the machine to take ‘n:m’ to $q2$ after the V. This maps the ‘n’ to an ‘m’ and thus transforms ‘CVnp’ to ‘CVmp.’ This derivation is given below:

Input:		C		V		n		p	
State:	q0	→	q0	→	q0	→	q2	→	q0
Output:		C		V		m		p	

(40) A derivation for CVnp $\mapsto$ CVmp in the FST in Figure 2

The reader can verify that this machine will take any permutation of ‘C’s, ‘V’s, ‘n’s and ‘p’s as an input, outputting ‘n’s as ‘m’s only before a ‘p.’ The reader can also verify that it would be impossible for the machine to accept an input that does *not* obey the generalisation; for example, CVnp $\mapsto$ CVnp, which does not show the assimilation pattern, will not be accepted by the machine, as there is no set of transitions that take ‘CVnP’ and map it to ‘CVnp’. Thus, the FST in Figure (2) describes exactly the mapping represented in (37) and (38).

3.3.2 Determinism and nondeterminism

Recall that in state q_0 of the FST in Figure 2, there are two options given an input ‘n’: transition ‘n:n’ to state q_1 or transition ‘n:m’ to state q_2. When a FST has, in any of its states, more than one possible transition given a certain input, it is NONDETERMINISTIC. In contrast, DETERMINISTIC FSTs are those which have at most one transition per input symbol at each state. Determinism is a defining property of the classes of left- and right-subsequential mappings, which as discussed above are a less complex subclass of the regular class of mappings (see, ex., Mohri, 1997).[14] In other words, deterministic FSTs are *less powerful* than nondeterministic FSTs: not every mapping describable with a nondeterminstic FST can be described with a deterministic FST, although the converse is true.

The nasal assimilation mapping above happens to be able to be modeled with a deterministic FST. This is accomplished by creating a machine that ‘waits’ for one transition—i.e., doesn’t output anything—any time it sees an ‘n,’ in order to see whether or not the subsequent symbol is a ‘p.’ Figure 3 gives such a machine. Because this machine employs a ‘waiting’ strategy, we need additional notation for what happens when the input string ends while the machine is waiting. State q_1 is labeled ‘q_1:n,’ which means that an

---

[14] For simplicity, I will abstract away from the other properties here.

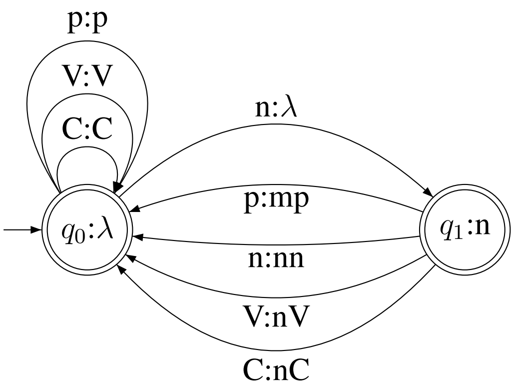


Figure 3: A deterministic nasal place assimilation FST

additional 'n' will be added to the output string machine when it reaches the end of the input on that state. This extension is merely notational and does not change any of the formal properties of the machine (Mohri, 1997). The symbol λ (as in the output for state q_0 Figure 3), like the $\varnothing$ symbol in phonological rewrite-rules, means 'don't write anything.' Thus, the label 'q_0:λ' means nothing extra is written when the machine ends on that state.

The 'waiting' strategy employed by these machines is as follows. Notice the 'n:λ' transition from q_0 to q_1. This means that the machine doesn't output when it sees an 'n' from state q_0, it simply moves to state q_1. Let us thus call state q_1 a 'wait' state, as at q_1 the machine 'waits' to see what the subsequent symbol is (if any) before outputting a symbol corresponding to the input 'n.' If the following input symbol is 'p,' it writes out 'mp'; for any other symbol, it writes both the 'n' and that symbol. If the string ends there, it simply outputs 'n.' This obtains the same input/output mapping as in the machine in Figure 2. The following derivations for the inputs 'CVnC' and 'CVnp' verify how this new machine in Figure 3 works:

Input:		C		V		n		C	
State:	q_0	$\rightarrow$	q_0	$\rightarrow$	q_0	$\rightarrow$	q_1	$\rightarrow$	q_0
Output:		C		V				nC	

a. CVnC $\mapsto$ CVnC

Input:		C		V		n		p	
State:	q_0	$\rightarrow$	q_0	$\rightarrow$	q_0	$\rightarrow$	q_1	$\rightarrow$	q_0
Output:		C		V				mp	

b. CVnp $\mapsto$ CVmp

(41) Derivations for inputs ‘CVnC’ and ‘CVnp’ for the new FST

The reader can confirm that the machine in Figure 3 accepts the same mapping as the one in Figure 2; however, in contrast to Figure 2, in each machine there is only one transition per input symbol at each state. They accomplish this by looking ahead a finitely bounded (one symbol) interval to see if there is a trigger for the change from ‘n’ to ‘m.’ Notice that because the number of states is finite (and we are not allowed to create any more throughout the derivation), *there can only ever be a finite number of wait states*. This idea of bounded lookahead is a crucial difference between mappings which can be modeled with and deterministic machines and those which require nondeterminism. There are regular mappings which require a machine that has unbounded lookahead. This can still be done with a finite state machine if it is nondeterministic because the nondeterminism allows it to ‘postpone’ a decision about a particular input symbol indefinitely. A concrete example will be shown with UTP in §4, but first, it is necessary to review the literature studying phonological processes from a finite-state perspective.

3.4 Subsequentiality and segmental phonology

Computational analyses of typologies of segmental processes have shown that they are describable with deterministic FSTs, with some variation regarding directionality. Mohri (1997) described two kinds of subsequential mappings: RIGHT-SUBSEQUENTIAL mappings and LEFT-SUBSEQUENTIAL mappings. Left-subsequential mappings can be described by a deterministic FST reading an input string left-to-right, as in the example in the previous section, whereas right-subsequential mappings can be described by a deterministic FST reading the input string right-to-left. Both classes are subclasses of the regular mappings; collectively, they can be referred to as the SUBSEQUENTIAL mappings.

The discussion from the previous section gives an intuition for how local processes are subsequential. Chandlee (2014) finds that 94% of the processes in the P-Base database of phonological patterns (Mielke, 2004) fall into a subset of the union of the left- and right-subsequential mappings she terms the Input Strictly Local functions. This includes epenthesis, deletion, metathesis, substitution, and partical reduplication (Chandlee et al., 2012; Chandlee and Heinz, 2012; Chandlee, 2014).

Work studying long-distance segmental processes such as vowel harmony (Gainor et al., 2012; Heinz and Lai, 2013) and dissimilation (Payne, 2014) has also found them to be largely left- or right-subsequential. The one exception is stem-control analyses of vowel harmony, in which a feature ‘radiates’ outwards to vowels in both prefixes and affixes (Archangeli and Pulleyblank, 1994; Baković, 2000). Bidirectional spreading patterns like this (and the Arabic emphasis example discussed in §2.4) are weakly deterministic because, as discussed in §5, they use the same subsequential FST reading *both* left-to-right and right-to-left.

It is perhaps less intuitive how long-distance processes can be subsequential, so this shall briefly be reviewed here. Any segmental process that is left- or right-subsequential has bounded lookahead in at least one direction—recall that the nasal place assimilation FST in §3.3 only required one 'wait state' to see what it should do with an input 'n'. Progressive long-distance processes are also left-subsequential, as they largely only depend on a trigger in the left context. Imagine a long-distance progressive consonantal harmony process in which a feature [–F] becomes [+F] after some other consonant specified [+F], no matter how early in the word this consonant appeared (an example is /l/ in KiYaka becoming [n] in suffixes attaching to a root containing a nasal; see Odden (1994)). Such a mapping, given in (42) (with the changed feature highlighted in bold), is describable with the deterministic FST in Figure 4.

(42) ...[+F]...[–F]... $\mapsto$ ...[+F]...**[+F]**...

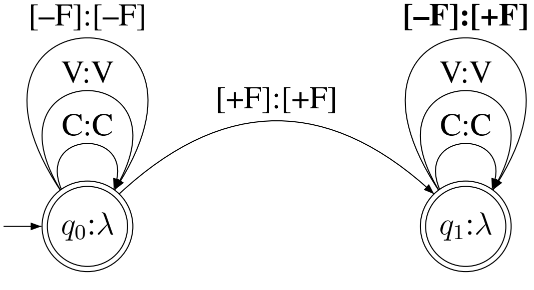


Figure 4: A deterministic FST for progressive consonant harmony (C=consonant, V=vowel)

This is impossible for a *regressive* harmony process, in which [–F] becomes [+F] *before* some [+F] segment an unspecified distance *later* in the word:

(43) ...[–F]...[+F]... $\mapsto$ ...**[+F]**...[+F]...

This requires unbounded lookahead in the left-to-right direction; in order to determine the output for a target [–F] in the input, a FST reading left-to-right would have to wait indefinitely to see if a trigger [+F] appears later in the string. However, this mapping is *right*-subsequential because reading the input *right-to-left* requires no lookahead. Reading right-to-left can be thought of as reversing the input, feeding it into the FST, and then reversing the output. If we reverse the strings in the mapping in (43), we get exactly the same mapping as in (42). We can thus describe the reverse of (43) with a deterministic FST, and so it is right-subsequential.

Thus, local processes and unidirectional long-distance processes are either left- or right-subsequential. Again, the intuition is that any process with a bounded lookahead in at least one direction is subsequential. As the literature cited above establishes, this is true for the vast majority of segmental phonology. However, there are two classes of phonological processes which are neither left- nor right-subsequential: bidirectional spreading processes and unbounded circumambient processes. The next section shows this for UTP. §5 will then argue that unbounded circumambient processes are not weakly deterministic and thus more complex than bidirectional spreading processes. This then provides a characterisation for the unbounded circumambient asymmetry: there is a weakly deterministic bound on segmental phonology, but not for tone.

4 Unbounded circumambient processes are not subsequential

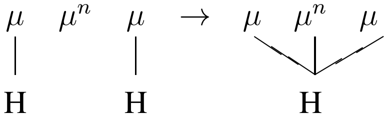
This section proves that UTP is neither left- nor right-subsequential. This is compared to a similar result for Sour Grapes by Heinz and Lai (2013), and is then generalised to the class of unbounded circumambient processes. Additionally, §4.4 defends the use of a linear representation of UTP.

4.1 UTP as a mapping

To analyze UTP using the computational framework for studying string-to-string mappings outlined in the previous section, we must use a string representation. The representation used here will mark associations to H tones on each TBU, which, as detailed in §4.4, represents featural information in a way parallel to the string representations used in the literature cited above. One may object to the string-based representation as not being faithful to analyses of the process which invoke autosegmental representations. This objection shall be addressed in §4.4. For now, the reader can understand the result as follows: *if* UTP is viewed with this kind of string representation, *then* it is not subsequential.

The Unbounded Tone Plateauing (UTP) generalisation, originally formalised in (11) in §2, is repeated below in (44). Its string-based counterpart is given in (45). In (45), H represents a TBU associated to a H tone, and ∅ represents an unspecified TBU. The superscripts m, n, and p, representing any natural number, make more explicit that this change happens for words of any length.

(44) a.



(45) a. $\varnothing^n \mapsto \varnothing^n$

b. $\varnothing^m\text{H}\varnothing^n \mapsto \varnothing^m\text{H}\varnothing^n$

c. $\varnothing^m\text{H}\{\varnothing,\text{H}\}^n\text{H}\varnothing^p \mapsto \varnothing^m\text{HH}^n\text{H}\varnothing^p$

(where m, n and p are natural numbers greater than or equal to zero)

The linear mapping in (45) makes explicit every possible situation in UTP. In (45a) and (45b), in which there are fewer than two H-toned TBUs in the underlying form, no plateauing occurs, and the input surfaces faithfully. Plateauing occurs instead when there are two or more Hs in the input (see, e.g., (14b) from Luganda). This is seen in (45c): all TBUs in between the first and last H-toned TBUs surface as H. The notation $\{\varnothing,\mathrm{H}\}^n$ denotes a string of n TBUs, either H or ∅.

4.2 UTP is not subsequential

The UTP mapping in (45) is not left- nor right-subsequential. As shall be shown, this is because it requires unbounded lookahead in both directions. As §5.3 will discuss in detail, this holds for any unbounded circumambient process. This section presents an informal illustration of the proof; for the full proof, see the Appendix (§8).

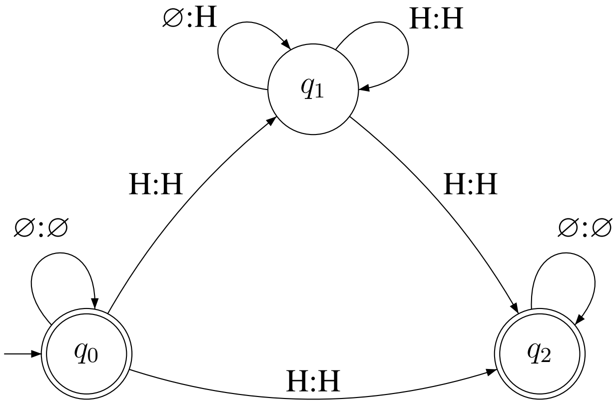


Figure 5: A non-deterministic FST for UTP

The UTP mapping in (45) is *regular*, as it can be modeled with a nondeterministic FST. This FST is given in in Figure 5. In this FST, underlying ∅ TBUs will only be output as H in state q_1, which is only reachable if there is another input H following; one can think of q_1 as the 'plateau' state. To illustrate, (46) contrasts derivations for the inputs H∅∅∅ and H∅∅H:

(46) a. H∅∅∅ ↦ H∅∅∅

Input:		H		∅		∅		∅	
State:	q_0	→	q_2	→	q_2	→	q_2	→	q_2
Output:		H		∅		∅		∅	

b. H∅∅H ↦ HHHH

Input:		H		∅		∅		H	
State:	q_0	→	q_1	→	q_1	→	q_1	→	q_2
Output:		H		H		H		H	

The input H∅∅∅ is mapped to H∅∅∅ because, as can be seen in (46a), the machine goes to state q_2 after seeing the first H, as there is no other H in the input. From state q_2, all input ∅s are output as ∅. Thus there is no change in the output. In contrast, given the input H∅∅H, the machine moves to state q_1 after the first H, because another H follows in the string. From q_1, all input ∅s are changed to H. In this way, the FST models the plateauing pattern, as it changes ∅s following a H to H if and only if another H follows. Note, however, that this takes advantage of nondeterminism; at state q_0 there are two transitions to take on an input H.

This nondeterminism is *necessary*, as there is no way to capture this mapping with a deterministic FST. Recall that a deterministic FST must have at most one transition per input symbol at every state. We cannot determinise the FST in Figure 5. To see why not, let us attempt the 'waiting' strategy employed in the Figure 3 FST in §3.3.2. While there are many ways to try this, this discussion follows one. A proof in the Appendix ensures that all will fail, but the discussion here is intended to give an intuition as to why.

In the deterministic FST in Figure 6 below, state q_2 is a waiting state representing the knowledge that a sequence H∅—which may be a plateauing environment—has been seen in the input. For the following diagrams, the state labels are again augmented with additional output symbols, as in the deterministic nasal place assimilation FST.

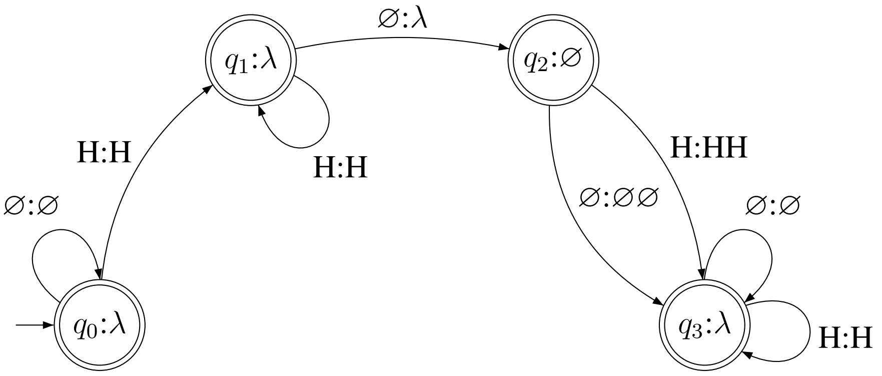


Figure 6: First attempt at a deterministic FST for UTP

The FST in Figure 6 outputs ∅ as ∅ until it sees a H, which sends it to state q_1. If it sees a ∅ in q_1, it needs to make a decision as to whether to output it as H (if it is in the plateauing environment; i.e. that another H is coming down the line) or a ∅ (if no H follows). The FST in Figure 6 'waits' one symbol to see if there will be a H or a ∅ in the input. It will thus correctly transform ∅H∅H, with only one intervening ∅ TBU, to ∅HHH. However it incorrectly maps inputs for which the second H is farther away; for example, the input ∅H∅∅H would be mapped to ∅H∅∅H (itself). Thus, because there is only one wait state, the machine in Figure 6 can only describe plateaus of at most three Hs. The following machine in Figure 7 thus adds an additional wait state to try and remedy this situation.

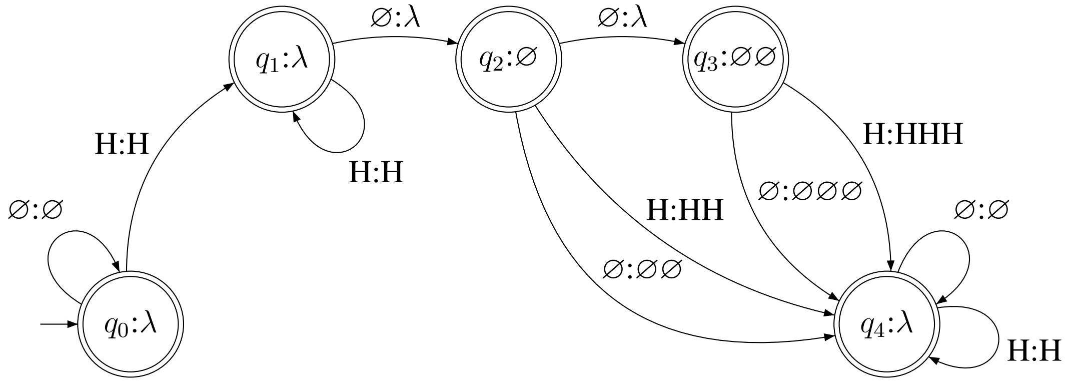


Figure 7: Second attempt at a deterministic FST for UTP

The FST in Figure 7 is much like the one in Figure 6, except that it has two wait states. It again outputs ∅ as ∅ until it sees a H, which sends it to state q_1. If it sees a ∅ in q_1, it 'waits,' outputting nothing. If it sees a H, then it outputs two H's and returns to q_0. This captures, as the FST in Figure 6 did, a plateau one TBU long. If it does *not* see a H, but instead a ∅, it 'waits' again, outputting nothing and moving to state q_3. If it then sees a H, it can make its decision regarding the previous two ∅ TBUs, and output them as two Hs.

Thus, Figure 7 correctly maps ∅H∅H ↦ ∅HHH and ∅H∅∅H ↦ ∅HHHH. However, it incorrectly maps inputs like ∅H∅∅∅H, where two Hs are separated by three or more ∅s, to themselves. By now it is perhaps obvious that we are on a wild goose chase; any 'wait n symbol' strategy will fail for any mapping in the UTP relation whose input string includes the sequence H$\varnothing^{n+1}$H. However, given the restriction of determinism, 'wait n symbols' is the best we can do (and, because we can't add states, we can't increase the value of n). Simply reversing the string, in an attempt to create a right-subsequential transducer, will not help us; the position of the first triggering H is just as arbitrarily far to the right as the second is to the left. Thus, a deterministic FST representation of the UTP mapping is impossible, and so it lies outside both the left- and right-subsequential classes.

4.3 Interim conclusion: Unbounded circumambient processes are not subsequential

That UTP is neither left- nor right-subsequential follows from its unbounded circumambient nature: as triggers may lie any distance away on either side of a given target, a FST describing the mapping requires unbounded lookahead in both directions.

Heinz and Lai (2013) prove this is also true for Sour Grapes, for the same reasons. For Sour Grapes, the fate of an input segment that can potentially assimilate rests on whether or not a trigger appears to one side *and* whether or not a blocker appears to the other. They show that this means it cannot be described by a deterministic FST reading in either direction. We can then see why no unbounded circumambient process can be subsequential. For unbounded circumambient processes, it is *definitional* that crucial information may appear on either side of the target, unboundedly far away. This means any unbounded circumambient process will require unbounded lookahead in both directions, and will not be subsequential. This result is then key to characterising the unbounded circumambient asymmetry in terms of computational complexity, as discussed further in §5.

As stated at the outset, this result depends on a particular kind of string-based representation. Thus, before moving on to compare unbounded circumambient processes with bidirectional spreading, as will be taken up in §5, it is necessary to address the potential objection that these results are invalid because they hinge on this string representation for UTP.

4.4 Autosegmental representations and linear representations

Tone is widely analyzed with autosegmental representations (though not universally so; for recent representational alternatives see Cassimjee and Kisseberth (2001) and Shih and Inkelas (2014)), while the notion of subsequentiality is defined in terms of strings. Thus, the preceding result showing that UTP, a tonal process, is not subsequential was based on a string representation. This section explains and defends two interrelated facts about such string representations:

(47) a. Each symbol in a string represents associations to a timing tier unit in any corresponding autosegmental representation.
  b. Therefore, 'unbounded lookahead' is determined by what translates to autosegmental terms as distance on the *timing* tier, not the *melody* tier.

As discussed below, (47a) is a common assumption about the relationship between string representations and autosegmental representations, and is one implicit in the string representations used in the computational literature cited in §3.4. As for (47b), measuring distance as on the timing tier as opposed to the melody tier is a representational *choice*. Indeed, Kornai (1995) compares different ways of encoding string representations of autosegmental representations. The following sections explain string representations based on (47) in detail and justify them by arguing the following three points. One, as seen in the preceding and following sections, given these string representations, neither UTP nor Sour Grapes is subsequential, which is a step towards understanding to why neither is well-attested in the segmental typology. Furthermore, this choice of representation studies the properties of processes without further representational assumptions relevant to locality on the melody tier. Finally, any further studies of complexity using different representations would have to duplicate the results using the particular representation here, and would not refute them.

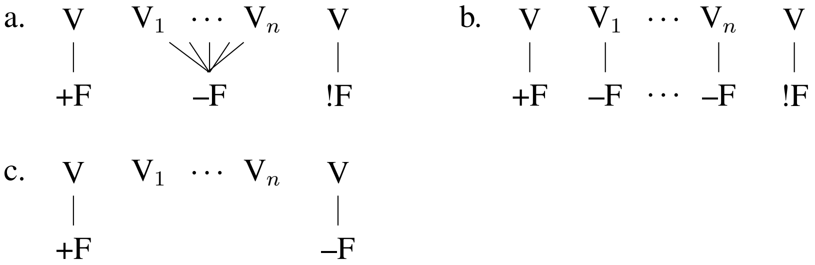
4.4.1 Translating between string and autosegmental representations

Let us compare string and autosegmental representations of the Sour Grapes mapping, repeated below in (48) from (33).[15]

(48) a. $[–F]^n \mapsto [–F]^n$ (no trigger, no blocker → no harmony)
b. $...[+F][–F]^n \mapsto ...[+F]\mathbf{[+F]}^n$ (trigger, no blocker → **harmony**)
c. $...[+F][–F]^n[!F]... \mapsto ...[+F][–F]^n[!F]...$ (trigger, blocker → no harmony)

We have a number of options for representing the featural contrasts made by the symbols in (48) autosegmentally. Some possible autosegmental representations for the underlying form $[+F][–F]^n[!F]$ from (48c) are given below in (49).

(49)



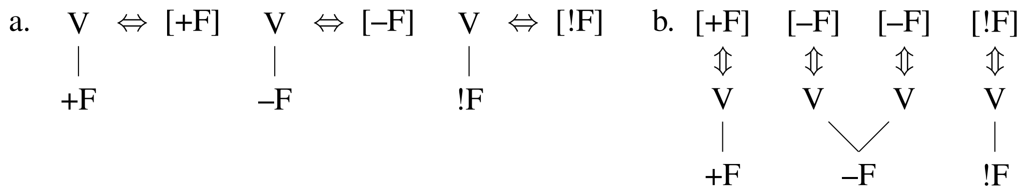
In (49a) and (b), the target [–F] vowels in the string are represented as underlyingly associated to [–F] features in the autosegmental diagram. In (49a), a single [–F] feature is associated to multiple vowels, whereas in (49b), each [–F] vowel is associated to its own [–F] feature (in violation of the Obligatory Contour Principle (Leben, 1973; McCarthy, 1986)). Here, a [!F] autosegment is used as shorthand for some vowel which is also associated to some other feature [+G] which prevents [+F] from spreading to it (e.g., as in [+low] vowels in Akan (Clements, 1976), which block the spreading of a [+ATR] feature). In (49c), the targets are analyzed as underspecified on the [±F] tier, and the blocker is analyzed as underlyingly specified as [–F] (as in Clements (1976)'s analysis of [+low] vowels in Akan as underlyingly [–ATR]).

One property that all possible autosegmental analyses share is that each symbol in a string in (48a) corresponds to the featural associations of a particular timing tier unit in (49). To give an explicit example of this, the following is a translation between symbols in UR strings in (48) and the autosegmental information at each timing tier unit in (49a). The translation for each string symbol [+F], [–F], and [+F] is given in (50a), and an example correspondence between the string [+F][–F][–F][!F] and an autosegmental representation is given in (50b).

---

[15] The [+F], [–F], and [!F] used here correspond to Heinz and Lai's $+$, $-$, and $\boxminus$, respectively.

(50)



While simplified in the sense that it focuses on one feature, translations like in (50) are what phonologists commonly, if implicitly, use when moving back and forth between linear strings of feature bundles and autosegmental representations. Again, what is important is that each symbol in the string corresponds to a timing tier unit in the autosegmental representation—the property originally highlighted in (47a). This is no less true of the string representations used by the computational work cited in §3.4, as this example from Sour Grapes shows.

As (50) makes clear, information about the *melody* (or featural) tier units in the autosegmental representation is obscured in the string representation. In (50b), the string does not encode the multiple associations of the [–F] autosegment, and it also does not show that [+F] and [!F] are only separated by a single [–F] autosegment. This ambiguity is apparent in (49), in which different autosegmental interpretations of the same string representation have different information on their featural tiers.

4.4.2 Representational assumptions and lookahead

As a result, when a FST reads in a symbol in a string representation like in (48), it can thus be thought of in autosegmental terms as reading the featural information of a particular timing tier unit. Thus, because subsequentiality depends on lookahead in terms of a FST reading such a string representation, the results regarding subsequentiality in the work of Heinz and Lai (2013) and others are based on what translates in autosegmental terms to lookahead on the *timing* tier, and *not* the melody tier. Thus, when using a string representation which follows (47a), subsequentiality is determined independently of whether or not there is an autosegmental analysis in which triggers and targets are local on the melody tier (as in (49a) and (c)).

There are a few arguments for this assumption. One is that locality on the melody tier depends on certain representational assumptions such as underspecification and the Obligatory Contour Principle, both of which have been argued against as phonological universals (see Inkelas (1995) for the former, Odden (1986) for the latter). The string representation is agnostic to an assumption about underspecification ([–F] could be written [0F], but this difference does not matter in terms of unbounded lookahead over the string).

More importantly, as detailed in this paper, it is by measuring unboundedness on the timing tier which distinguishes UTP and Sour Grapes from common segmental processes via its non-subsequentiality. Thus, while this choice is an assumption on which the results in this and the literature cited in §3.4 is based, it appears to be correct in that it helps us to characterise the unbounded circumambient asymmetry. How this relates to different representational assumptions will be discussed momentarily.

Returning to UTP, it is then important to establish that the linearisation used in the previous section is indeed comparable to that of previous studies of subsequentiality, such as for Sour Grapes in (48), in that it also measures lookahead in terms of the timing tier. The linearisation in (45) can be made explicit with a mapping along the lines of the one in (50) as given below in (51).

(51) a. $\begin{array}[t]{c}\mu\\ | \\ \mathrm{H}\end{array} \Leftrightarrow \mathrm{H} \qquad \mu \Leftrightarrow \varnothing$ b. $\mu \quad \begin{array}[t]{c}\mu\\ | \\ \mathrm{H}\end{array} \quad \mu \Leftrightarrow \varnothing\mathrm{H}\varnothing$

As can be seen in (51b), the string symbols H and ∅ can encode both contrasts in association in the underlying representation and the changes in these associations in the surface, as originally claimed in (47a). Additionally, the string representation in (51b) preserves the linear order of the TBU tier of the autosegmental representation, just as in (50). Thus, lookahead is measured in terms of the timing tier, as originally claimed in (47b). Note again that, as they simply encode the associations to each timing tier unit, linearisations like in (50) and (51) are very general, and can be applied to any set of autosegmental representations for a particular process.

4.4.3 Representation and subsequentiality

This concludes the arguments for why this particular kind of representation is used in this paper. However, an important question remains: what if we *don't* use this particular representation? The answer is that this is a valid direction for research, but any results along these lines will not change the result argued for in this paper.

It is known that, formally, representation is related to expressive power (Medvedev, 1964; Rogers et al., 2013). A potentially fruitful direction for future research is to study the relationship between representation and computational complexity as it relates to phonology. There has been much important work studying the computational properties of autosegmental representations, notably the automata-theoretic work of Kay (1987), Wiebe (1992), Bird and Ellison (1994), and Yli-Jyrä (2013). Also, various linear encodings of autosegmental representations are discussed in detail in Kornai (1991, 1995). However, this work does not yet offer a hierarchy of complexity like the one presented in §3.4. Future research may use this work in computational autosegmental phonology as a starting point to study the relationship between representation and computational complexity of phonological processes.

It may even be possible to derive non-subsequentiality of unbounded circumambient processes from some aspect of representation (such as lookahead on the melody tier). This requires, however, that if such representations were translated into the string representations according to (47), non-subsequentiality in the string representations would somehow emerge. Thus, any such explanation based on representation would *duplicate* the computational results outlined here—it would not *refute* them. That is, it would have to uphold the fact that, when viewed as string mappings for for which (47) is true, much of segmental phonology is subsequential while UTP and Sour Grapes are not.

5 Unbounded circumambient processes and weak determinism

Having shown the result that UTP and Sour Grapes, and by extension any unbounded circumambient process, are not left- or right-subsequential, and having defended the linear representations upon which these results are based, there is one final distinction to be made. This section argues that unbounded circumambient processes are computationally distinct from bidirectional spreading processes, first introduced in §2.5, because the latter are weakly deterministic but the former are not. This leads to the characterisation of the unbounded circumambient asymmetry in terms of a weakly deterministic complexity bound on segmental phonology which is not present in tone.

5.1 Bidirectional spreading and the weakly deterministic class

Unbounded circumambient processes are not the only class of process which is not left- or right-subsequential. In patterns of stem-control harmony, or cases of bidirectional spreading, as in the Arabic emphasis spreading discussed in §2.5, a feature spreads outward both to the right and the left:

(52) ...[–F]...[+F]...[–F]... $\mapsto$ ...**[+F]**...[+F]...**[+F]**...

Such a mapping requires unbounded lookahead in either direction: a target may follow or precede the trigger, in any direction. Thus, as Heinz and Lai (2013) also show, such cases (which shall henceforth be referred to under the umbrella term BIDIRECTIONAL SPREADING) are also not subsequential. However, there is a crucial difference between bidirectional spreading and unbounded circumambient processes. Bidirectional spreading hinges on a single trigger whose influence spreads outward. In contrast, unbounded circumambient processes hinge on *two* triggers/blockers whose targets lie between.

Heinz and Lai (2013) observe that bidirectional spreading processes are essentially the same unidirectional mapping applied left-to-right and then right-to-left. They propose a superclass of the subsequential mappings, called the WEAKLY DETERMINISTIC mappings, which includes this kind of process. Weakly deterministic mappings are those which can be *decomposed* into a left- and right-subsequential mapping such that the left-subsequential mapping is not allowed to change the alphabet or length of the string.

The process in (52) can be decomposed into two left- and right-subsequential mappings describable by the consonant harmony FST in Figure 4 in the following way. First the input string is read by the FST left-to-right (applying the left-subsequential mapping), then the resulting output is fed back into the FST right-to-left (applying the right-subsequential mapping to the output). This decomposition is schematised below in (53):

(53) a. (left-subsequential) ...[–F]...[+F]...[–F]... $\mapsto$ ...[–F]...[+F]...**[+F]**...
b. (right-subsequential) ...[–F]...[+F]...[+F]... $\mapsto$ ...**[+F]**...[+F]...[+F]...

This composition of the two mappings is special because it does not change the alphabet or the length of the string. Intuitively, this is because these bidirectional processes can be thought of as one unidirectional process operating in two directions. This is highlighted by the fact that both sub-mappings use the *same* FST. Thus, bidirectional spreading is weakly deterministic, as first seen in Figure 1.

The weakly deterministic class is defined as such by Heinz and Lai (2013) as a restriction on Elgot and Mezei (1965)'s result that any regular mapping can be decomposed into a left-subsequential and right-subsequential mapping, as long as the left-subsequential mapping is allowed to enlarge the alphabet. This result holds because the left-subsequential mapping can 'mark-up' the string with extra symbols in the first mapping, and then erase them in the second.[16] However, no attested segmental mappings studied in the literature cited above require such a markup.

5.2 Unbounded circumambient processes and the weakly deterministic class

In contrast, Heinz and Lai (2013) argue that Sour Grapes requires such a markup, and thus is not weakly deterministic.[17] This is because Sour Grapes is not simply the application of the same process in two directions. Because Sour Grapes is unbounded circumambient, for any decomposition into two sub-mappings, the left-subsequential process must somehow encode whether or not a [+F] has been seen to the right of the [–F] targets. It is difficult to see how this can be done without intermediate markup, and thus Heinz and Lai (2013) conjecture that Sour Grapes is not weakly deterministic. The exact same arguments apply to UTP. Let us look at why more concretely.

UTP can be decomposed into two subsequential mappings with an *augumented* alphabet, as in (54). First, the left-subsequential process marks all ∅ following a H as ?,

---

[16] This is not unlike the use of abstract intermediate forms in early derivational phonology, e.g. (Clements, 1977).

[17] *Proving* that a pattern is not in a complexity class requires an abstract characterization of that class which allows for analytical tools such as the pumping lemmas for regular and context-free languages (Hopcroft et al., 2006). No such tools exist yet for the weakly deterministic class. Another complexity class for which no such proof exists is P, or the class of problems that can be solved in polynomial time. Many problems in the NP class conjectured to be outside of P are considered computationally intractable (see Idsardi, 2006; Heinz et al., 2009, for such discussion on OT), although no proof exists that $P \neq NP$ (Fortnow, 2009).

and then the right-subsequential process changes all ? preceding a H to H (and all ? not preceding a H to ∅).

(54)	Input	∅∅∅H	H∅∅∅	H∅∅H
		↧	↧	↧
	a. (left-subsequential)	∅∅∅H	H???	H??H
		↧	↧	↧
	b. (right-subsequential)	∅∅∅H	H∅∅∅	HHHH
	Output	∅∅∅H	H∅∅∅	HHHH

Crucially, this decomposition relies on the intermediate ? symbols to 'carry forward' the information that a H appears to the left in the string. This allows the right-subsequential mapping to correctly apply without any unbounded lookahead. However, like Sour Grapes, it is hard to see how there could be a similar decomposition which uses only H and ∅, maintains string length in the left-subsequential sub-mapping, and obtains the exact same mapping. For example, instead of using ? to mark a string of ∅ following a H, the left-subsequential function could change it to an alternating string of ∅ and Hs. Thus, H∅∅∅∅∅∅ would be mapped to H∅H∅H∅H. However, there would be no way for the right-subsequential function to distinguish this from an input of H∅∅∅∅∅H, which the left-subsequential function would also map to H∅H∅H∅H. In general, using the same alphabet to create an encoding will always lead to distinctions between input strings being lost, and so such an encoding is bound to fail.

Again, this is based on the unbounded circumambient nature of the process: because there is crucial information on either side of the targets, it is necessary to mark targets to the right of the left trigger in order for the right-subsequential function to correctly process them without unbounded lookahead. Thus, Heinz and Lai (2013)'s conjecture also applies not just to UTP, but to any unbounded circumambient process.

An important implication of Heinz and Lai (2013)'s conjecture is that there exist FULLY REGULAR mappings outside of the weakly deterministic class, and thus that the weakly deterministic class of mappings is, as depicted in Figure 1, a proper subclass of the regular class of mappings.[18] This section has argued that unbounded circumambient processes fall into this fully regular class (also depicted in Figure 1), and that this distinguishes them in complexity from bidirectional spreading.

---

[18] However, not all fully regular mappings would correspond to some unbounded circumambient mapping. One example from Mohri (1997) of a mapping which is neither left- nor right-subsequential is the mapping $a^n \mapsto b^n$ if n is even and c^n if n is odd. This mapping is also almost certainly not weakly deterministic. However, this mapping depends not on two targets but on whether or not the string is evenly divisibly by 2. This fully regular mapping would look bizarre as a phonological process, even as a tonal one, so the attested unbounded circumambient tonal processes must belong to some as-yet undiscovered subclass of the fully regular mappings, as to be discussed in §7.

5.3 The weakly deterministic hypothesis

Thus, this conjecture provides a way to capture the unbounded circumambient asymmetry. Heinz and Lai (2013) posit a WEAKLY DETERMINISTIC HYPOTHESIS for phonology: phonology is at most weakly deterministic. This hypothesis correctly predicts the absence of unbounded circumambient processes like Sour Grapes in the typology of vowel harmony. However, this paper has shown that unbounded circumambient processes are well-attested in tone. This paper thus proposes that the weakly deterministic bound *only* applies to segmental phonology, and not *tonal* phonology. This accurately predicts fully regular mappings such as UTP and Copperbelt Bemba H spread to exist in tone. Furthermore, to propose that tone is more computationally complex than segmental phonology, as has been done here, is in line with Hyman (2011) and others' assertions that tone can 'do more' than phonology.

An explanation for how the weakly deterministic bound manifests in the phonological system shall be left for future work, although it is possible to speculate on a few points. One, as mentioned in §3.1, formal language complexity correlates with an increase in computational resources necessary for parsing and generation. It could be that tone has access to such resources because prosodic information more commonly interacts with syntax (see, e.g., Hyman and Katamba, 2010) and thus requires more powerful computation. This issue of computational power can also be directly related to learning, as empirical work suggests that formal complexity constrains phonological learning (Heinz, 2010; Lai, 2012, to appear; McMullin and Hansson, 2015; Moreton and Pater, 2012; Rogers et al., 2013).

5.4 Explaining the exceptions to the weakly deterministic hypothesis

The weak determinism hypothesis in its strongest form cannot account for the rare cases of unbounded circumambient segmental processes discussed in this paper, Sanskrit *n*-retroflexion and KiYaka vowel harmony. There are a few ways to reconcile these exceptions with the weakly deterministic hypothesis.

One possibility is that the accounts of these processes in the literature were incorrect in classifying them as unbounded. As brought up in §2.3.1, Ryan (2015) has observed that blocking of Sanskrit *n*-retroflexion may be bound to the adjacent syllable. In KiYaka, the greatest attested distance was only three vowels, and was restricted to a particular morphological context. This is in stark contrast to UTP, for which §2.2 documented a number of attestations, all of which were clearly unbounded.

Another explanation comes from potentially interfering factors. For example, the presence of an unbounded circumambient process in the tonal phonology may license an unbounded circumambient process in segmental phonology. This may be the case for KiYaka, which as noted in §2.2.3, also has UTP. It could be that KiYaka speakers, having first internalized plateauing process in tonal phonology, may then be able to generalise it to their segmental phonology.

Finally, it may simply be that the constraint against fully regular mappings in segmental phonology is not categorical but somehow gradient, and thus admits exceptions. One way this may manifest is through a learning bias in which individuals are more receptive to some patterns than others (Moreton, 2008; Wilson, 2006a). For example, Moreton (2008) shows how learners are less receptive to certain vowel-consonant dependencies than vowel-vowel dependencies. That the former kind of pattern is attested, although significantly less than the latter, is explained by learners overcoming this bias when exposed to enough data. A related explanation for the rare unbounded circumambient segmental processes would be that human children show a strong preference for weakly deterministic mappings when learning segmental phonology, but may change to a more general learner in the face of sufficient data.

As Staubs (2014) shows, gradient typological generalisations may also result from the transmission of patterns among multiple learning agents. For example, he shows that stress patterns with larger stress windows are less likely to be passed on in successive generations as learners need to see longer words in order to obtain the correct generalisation. At this point, it is only possible to speculate how the complexity of unbounded circumambient processes may be related to such an explanation. However, it is likely that non-weakly deterministic patterns require more kinds of data to learn. As tone is known to operate over much longer domains, this may provide such data in a way that is rarer in segmental phonology. Thus tonal unbounded circumambient processes may be more likely to be passed on from generation to generation.

Hence, there are a number of reasons why the cases of Sanskrit *n*-retroflexion and KiYaka vowel harmony do not immediately invalidate the proposal offered here. Even if they cannot ultimately be explained away, they are exceptional for other theories besides this one. As mentioned, for example, both processes violate Wilson (2006b)'s myopia generalisation. Furthermore, as to be discussed in §6, Hyman (1998)'s OT analysis of KiYaka vowel harmony utilises the same Sour Grapes-type behaviour that Wilson (2003, 2006b) and McCarthy (2010) attempt to remove from OT. Finally, Sanskrit *n*-retroflexion and KiYaka vowel harmony appear to be the only such cases—Hyman (2011) states, for example, that KiYaka is the "only one example" of such a process of which he is aware (p. 218). As such, the development of the potential explanations above shall be left to future research.

In sum, this paper has characterised the unbounded circumambient asymmetry in terms of computational complexity and identified the weakly deterministic boundary as the relevant difference between tone and segmental phonology. The following section contrasts this account with ones couched in OT, which are shown to be unable to provide a unified characterisation of the typological assymetry.

6 The Unbounded Circumambient Asymmetry in Optimality Theory

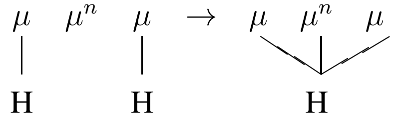
This section argues that the cross-linguistic generalisation from §2 is difficult to capture with Optimality-Theoretic constraint interaction (Prince and Smolensky, 1993, 2004). This is because current theories of OT do not provide a unified characterisation of the unbounded circumambient processes. As shall be argued, this means that making OT empirically adequate with regards to the asymmetry runs into a ‘duplication of effort’ problem. Banning the particular non-local effects that generate segmental unbounded circumambient processes requires changes both to how OT manipulates segments autosegmentally but also to how OT compares candidates. Thus, OT does not provide a unified characterisation of the asymmetry comparable to the one based on computational complexity put forth in the preceding sections.

The structure of this section is as follows. §6.1 and §6.2 present OT analyses of UTP and KiYaka vowel harmony, respectively, to illustrate how OT generates unbounded circumambient processes.[19] §6.3 then discusses what restrictions might be added to OT to capture the unbounded circumambient asymmetry, reaching the conclusion that a number of unrelated changes are necessary.

6.1 UTP in Optimality Theory

This section provides an OT account of UTP, and shows that it is possible using constraints which have also been posited for segmental phonology. Hyman and Katamba (2010)’s formalisation of UTP is repeated below:

(55)



This process can be derived from through the interaction of standard principles of Autosegmental Phonology (AP). The first is the Obligatory Contour Principle (OCP; Leben, 1973; McCarthy, 1986), which bans adjacent, identical autosegments.

(56) OCP: Autosegments adjacent on a tier must be distinct. Assign one violation mark for every sequence of adjacent, identical autosegments.

The underlying structure in (55) violates this OCP constraint, and so this will be one of the constraints that motivates the surface plateau.[20]

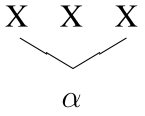
---

[19] Sanskrit will not be discussed here, as, to the author’s knowledge, the only analysis of the blocking condition is by Ryan (2015), who uses not classic, parallel OT but serial Harmonic Grammar.

[20] For expositional clarity, the analyses here only consider URs with up to two underlying Hs. Accomodating three or more Hs will have no significant effect on the analyses discussed.

Another principle is a general constraint on gapped structures, in which associations skip over potential timing tier units. An example is given in (57), in which the autosegment α is associated to the first and third target X but not the second. Constraints against gapped structures have been discussed in the literature in both vowel harmony (Levergood, 1984; Archangeli and Pulleyblank, 1994; Ito et al., 1995; Ringen and Vago, 1998; Walker, 1998) and tone (Archangeli and Pulleyblank, 1994; Yip, 2002).

(57) A gapped structure (from Archangeli and Pulleyblank, 1994, p. 33, (42b))



A constraint banning this structure is given in (58) This markedness constraint shall also be ranked highly, ensuring that any changes to avoid an OCP violation will not result in a gapped structure.

(58) NoGap: Multiply linked features cannot skip TBUs. Assign one violation mark for every gapped structure in the output.

Yip (2002) posits basic Faithfulness constraints for tones, such as Max-T, banning the deletion of tones. While tone-specific, Max-T follows the schema of the original Max-IO constraint family (McCarthy and Prince, 1995). In fact, Ringen and Vago (1998) use an almost identical Max$_{\text{subseg/rt}}$ constraint against the deletion of featural autosegments for vowel harmony. Another constraint of Yip's to be used here is *Assoc, which bars the addition of association lines.[21] Similar constraints have been proposed for autosegmental analyses of vowel harmony, such as in (Pulleyblank, 1996).

(59) Max-T: Do not remove underlying tones. Add one violation mark for every tone in the UR that is not in the SR.

(60) *Assoc: Do not insert association lines. Add one violation mark for every association line in the SR that is not in the UR.

One more Faithfulness constraint is necessary to pick out fusion as the optimal repair for OCP. Faithfulness constraints against fusion have been posited for both segmental phonology (e.g., Pater, 2004) and tone (Meyers, 1997). The version here is taken from Meyers (1997):

(61) Uniformity-IO(X): If a and b are distinct elements of type X in the input, then their output correspondents a' and b' are also distinct elements of type X (Meyers, 1997, p. 852).

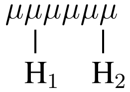
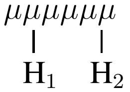
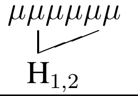
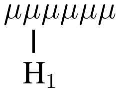
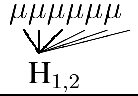
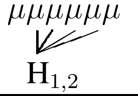
---

[21] Although its name contains an asterisk, which is usually used for Markedness constraints, Yip explicitly discusses *Assoc as a Faithfulness constraint. Her naming convention is followed here.

The constraint Uniformity-IO(T), henceforth abbreviated as Unif, militating against fusion of tones, must be ranked below the Markedness constraints listed above. Fusion is a natural repair for the OCP—this captures the generalisation is that a single, multiply associated autosegment is preferred over a sequence of adjacent, similar autosegments.

The correct output obtains if the Markedness constraints OCP and NoGap are ranked above Unif and *Assoc but no other Faithfulness constraints. This is illustrated in (62) below with an example input in which two Hs appear separated by two morae. Because there are multiple morae in between them, NoGap bars candidate (b), in which the Hs are simply fused. The low ranking of *Assoc instead favours candidate (e), in which the Hs are fused and just enough association lines are added to satisfy NoGap.

(62)

 μμμμμμ H_1 H_2	OCP	NoGap	Max-T	Unif	*Assoc
a.  μμμμμμ H_1 H_2	*!				
b.  μμμμμμ $H_{1,2}$		*!		*	
c.  μμμμμμ H_1			*!		
d.  μμμμμμ $H_{1,2}$				*	***!*
☞ e.  μμμμμμ $H_{1,2}$				*	**

Thus, the correct outputs for UTP obtain from the following ranking: {OCP,NoGap, Max-T}≫{Unif,*Assoc}≫Spec-T. The constraints listed here all have segmental counterparts which have been posited by previous researchers to exist in Con; a theory of OT containing them thus predicts plateauing processes in segmental phonology as well. This can easily be seen by replacing the Max-T in the tableau in (62) with a generic max constraint for autosegments, replacing the H tone in the input and candidates with a vowel feature and changing the morae to vowels.

However, to generate an unbounded circumambient segmental plateauing pattern, it is not necessary to appeal to autosegmental representations. The following review of Hyman (1998)'s analysis of KiYaka shows that segmental plateauing can even be captured with local linear constraints.

6.2 KiYaka Vowel Harmony

This subsection reviews Hyman (1998)'s analysis of KiYaka vowel harmony, which, as discussed in §2.3.2, is a rare example of a circumambient, bidirectional vowel harmony process. With the addition of two FAITHFULNESS constraints marking prosodically strong and weak positions in the root, Hyman's analysis generates the process with one unidirectional, local agreement MARKEDNESS constraint. First, let us review the data. The following forms are repeated from (30) and (31) in §2.3.2. When harmony applies, high vowels assimilate to surrounding mid vowels, as when the perfective /-ile/ attaches to (63c) and (d) (harmonising vowels are underlined; again, there are /l/→[d] and /l/→[n] changes in this suffix which will not be discussed here).

(63) KiYaka mid-vowel plateauing

	Gloss	Stem	Low FV	Mid FV (+ perf. /ile/)
a.	'obstruct'	/kik/	kik-ila (+appl. /ila/)	kik-idi
b.	'send'	/hit-ik/	hit-ik-a (+FV /a/)	hit-ik-idi
c.	'pay attn.'	/keb/	keb-ila (+appl. /ila/)	keb-ele
d.	'to do an about-face'	/kel-umuk/	kel-umuk-a (+FV /a/)	kel-omok-ene

This harmony does not apply when only a single mid vowel either precedes the target high vowel, as when the roots in (63c) and (d) are followed by a suffix with a low final vowel ([keb-ila] 'pay attn.+appl' and [kel-umuk-a] 'to do an about-face', respectively), or follows the high vowel, as in (63a) and (b) ([kik-idi] from /kik-ile/ 'obstruct+perf.' and [hit-ik-idi] from /hit-ik-ele/ 'send+perf.', respectively). In the latter case, progressive high vowel harmony changes the suffix /ile/ to [idi].

Hyman (1998)'s OT analysis of these facts are as follows. First, following other Bantuists, he posits two positionally-based IDENT constraints to govern the behaviour of the initial and final vowels (he notes this approach is based Beckman (1997)). The first is the IDENTV$_1$, which militates against change in the first mora of a stem. This constraint is ranked highly, reflecting that the initial stem mora is often prominent in Bantu languages. The second is the lower-ranked IDENTFV, which militates against change in the final vowel of a word.

(64) FAITHFULNESS constraints for KiYaka (Hyman, 1998, (16); violation rules added)

a. IDENTV$_1$: preserve the mora and features of the first stem vowel. Assign one violation mark for every change to the first stem vowel.

b. IDENTFV: preserve the mora and features of the final vowel. Assign one violation mark for every change to the first stem vowel.

The relevant markedness constraint is an agreement constraint Hyman calls PLATEAU; this constraint simply assigns violation marks when a non-mid vowel precedes a mid vowel:

(65) Plateau: *HM, *LM. Assign one violation for every high or low vowel preceding a mid vowel in the input.

As the following tableaux will illustrate, it is enough to rank Plateau in between IdentV$_1$ and IdentFV to generate the correct pattern (for brevity, the faithfulness ranking that determine the specific changes in the vowels will not be discussed here). Let us first show that when single mid vowel precedes all of the high vowels in the input, as in (63d) /kel-umuk-a/ 'to do an about-face+FV', the most faithful candidate wins, as it does not violate Plateau:

(66)

	/kel-umuk-a/	IdentV$_1$	Plateau	IdentFV
☞	kel-umuk-a			
	kil-umuk-a	*!		
	kil-umuk-e			*!

Thus, correctly, the ranking produces no change for a UR like /kel-umuk-a/. The next tableau illustrates the output when a single mid vowel follows high vowels in the UR for (63a) /kik-ile/ 'obstruct+perf.'. In this situation, Plateau is violated, so a non-faithful candidate will surface. Because IdentFV is the lowest ranked faithfulness constraint, it is the final vowel which changes, resulting in progressive harmony. The tableau in (67) is identical to Hyman (1998)'s (16).

(67)

	/kik-ile/	IdentV$_1$	Plateau	IdentFV
	a. kik-ile		*!	
	b. kik-ele		*!	
	c. kek-ele	*!		
☞	d. kik-idi			*

In (67), the vowels in the suffix of the faithful candidate (67a) *kik-ile violate Plateau. However, simply changing the middle vowel to [e], as in (67b) *kik-ele, is not enough, because the first and second vowel still violate Plateau. However, the initial /i/ must remain as such, less IdentV$_1$ be violated. Instead, candidate (67d), in which the final vowel /e/ is edited to [i], violating IdentFV but satisfying both Plateau and IdentV$_1$, wins.

Finally, we turn to the case of two /e/s on either side of high vowels. The UR (63d) /kel-umuk-ile/ 'to do an about-face+pref.' is used to illustrate. Again, the vowels in the /ile/ suffix violate Plateau, precipitating the change in the surface form. However, because the initial vowel is itself mid, the optimal change is instead to convert all of the intervening high vowels into mid:

(68)

/kel-umuk-ile/	IDENTV$_1$	PLATEAU	IDENTFV
a. kel-umuk-ile		*!	
b. kel-umuk-idi			*!
c. kel-umuk-ele		*!	
d. kel-umok-ele		*!	
☞ e. kel-omok-ele			

Candidate (68e) wins because, unlike (68b), it does not violate IDENTFV. It does so by changing *all* of the intervening vowels—note that candidates (68c) and (d), in which the vowel harmony ‘stops’ midway, still violate PLATEAU. This will hold true no matter how long the input word is. Crucially, this non-local effect is created using local constraints by a ‘global comparison’ made possible by the infinite candidate set in GEN: under standard OT, candidates like (68e), in which *all* intervening vowels are changed, can be compared with candidates in which only *some* of the intervening vowels have been changed. This is very similar to how optimisation produces Sour Grapes, a point that will be returned to below.

We have now seen how the constraint ranking IDENTV$_1$$\gg$PLATEAU$\gg$IDENTFV produces the correct output given each type of input. Thus, Hyman (1998)’s analysis of KiYaka vowel harmony generates, with a simple set of linear constraints, a pattern in which an unbounded number of high vowels will change to mid if and only if there are two /e/s on either side of them. Again, this is an unbounded circumambient process; it is circumambient because the trigger /e/s of the mid-harmony must be on both sides of the targets, and it is unbounded because it applies to target high vowels no matter how far to the left or right from the trigger mid vowels they are.

6.3 Characterising the asymmetry in OT

This section has reviewed OT analyses of UTP, the tonal unbounded circumambient process at the center of this paper, and Hyman (1998)’s analysis of KiYaka vowel harmony. For UTP, the motivating concepts were the OCP, a constraint against gapped structures, and fusion. For KiYaka, the analysis rested on positional identity and a local agreement constraint. These analyses used standard theoretical machinery proposed elsewhere in the literature for segmental phonology, but were able to generate unbounded circumambient processes. How, then, can OT account for the asymmetry in unbounded circumambient processes between tone and segmental phonology? It is argued here that in order to do so, there is a ‘duplication of effort’ problem in trying to characterise the empirical generalisation in OT: any changes would not only have to address autosegmental manipulation of segments, but also situations in which OT’s GEN allows comparison of non-local effects with local ones. Thus, it is possible to account for the typological asymmetry in OT, but not in a way that captures the insight behind the generalisation in any unified way.

To start, there is a choice between categorically banning segmental unbounded circumambient processes from OT, or introducing some set of biases that accept rare cases like KiYaka. The former is much simpler for ‘classical’ OT, which, through permutations of rankings of the universal constraint set CON, categorically states whether or not a process should exist at all. The latter can be done by introducing weighted constraints or ranking volumes (Wilson, 2006a; Bane and Riggle, 2009, inter alia). For the sake of concision, this discussion will only concern categorical OT, however gradient OT grammars suffers very similar duplication of effort problems in trying to characterise the unbounded circumambient asymmetry.

For a categorical banning of segmental unbounded circumambient processes, one solution is to pick and choose which constraints should be in CON. To prevent the segmental version of the unbounded plateauing as explained in §6.1, one could remove from CON the segmental versions of the constraints responsible for the fusion of autosegments and the filling in of the intervening material. For example, as NOGAP partially motivates plateauing, perhaps there should be no such MARKEDNESS constraint for segmental phonology. This would require revisiting previous vowel harmony analyses that used NOGAP (e.g., Ringen and Vago, 1998), but it is possible.

Another answer is to modify the autosegmental representations for segmental information that appear in GEN. As plateauing is partially motivated by fusion of autosegments, it may be possible to eliminate the process by banning candidates featuring the fusion of non-tonal featural autosegments. A further option is to abolish autosegmental representations for segmental features from GEN altogether and use agreement-based analyses for segmental harmony (as in the analyses of Baković, 2000; Rose and Walker, 2004; Hansson, 2001).

However, the case of Sour Grapes shows that OT does not need recourse to autosegmental representations to generate unbounded circumambient processes. Sour Grapes is unattested in segmental phonology but predicted by parallel OT and local AGREE constraints checking the agreement in a feature of adjacent segments. Its unbounded behaviour is achieved through exactly through the same ‘global comparison’ used in Hyman (1998)’s analysis of KiYaka vowel harmony (see McCarthy (2010) for a thorough explanation of how Sour Grapes is generated in OT). This global comparison is due to the nature of optimisation, which allows comparison of candidates with non-local changes to those with local ones. One proposal for dealing with this problem is McCarthy (2010)’s analysis of spreading in Harmonic Serialism (HS), which restricts GEN to only produce candidates with a single change (with winning candidates fed back into the grammar until no changes are more optimal). As only candidates with single changes are compared, global comparison cannot take place.

However, it is unlikely that a change to HS alone would also remove autosegmental plateauing from the predicted segmental typology, as there is a derivational account for the process in which fusion occurs first and then the gapped structure is gradually filled in. We then have a duplication of effort problem: in order to categorically remove segmental unbounded circumambient processes from the typology predicted by OT theories of phonology, it is necessary to at least adopt one of the changes proposed for Sour Grapes *and* one of the changes adopted for segmental plateauing.

In sum, to account for the unbounded circumambient asymmetry in OT requires adopting a number of unrelated solutions which all conspire to keep unbounded circumambient processes out of segmental phonology. Thus, while it appears technically possible to have a theory of OT that is empirically adequate with regards to the unbounded circumambient asymmetry, it does not provide a unified explanation. In contrast, the FLT-based account presented here draws a clear line: a weakly deterministic bias present for segmental phonology but not for tone predicts the rarity of unbounded circumambient processes in the former but not in the latter. It is possible that OT can provide an explanation for the unbounded circumambient asymmetry by 'coopting' this proposal—that is, by stipulating a constraint that the mappings they produce for segmental phonology remain within the weakly deterministic level of complexity. This would be somewhat similar to Tesar (2013)'s work relating learnability in OT with output-driven maps, although it remains to be seen how CON or GEN must be designed to ensure so much mappings are present in the resulting factorial typologies.

7 Conclusion

This paper has made three contributions. One, it has documented the asymmetry in the attestation of unbounded circumambient processes in tonal phonology and segmental phonology. Two, it has shown that UTP is similar to the Sour Grapes, in that they are both unbounded circumambient processes, and that this similarity has formal consequences. The third contribution then accounts for the asymmetry based on the hypothesis that unbounded circumambient processes are fully regular, which posits that they are more computationally complex than processes which do not require unbounded lookahead in two directions. This has been shown to be a superior characterisation than that offered in Optimality Theory, which predicts such processes to be equally attested in segmental and tonal phonology.

The conclusions in this paper raise a number of interesting questions for future research. For one, what computational constraints are there on tone? As all of the processes described here are *at least* describable by non-deterministic finite-state transducers, a reasonable first hypothesis is that tonal processes correspond to the regular mappings. However, this hypothesis is not restrictive enough—as Footnote 18 mentions, there are regular mappings that no phonologist would recognise as even a tonal process. Is there, then, a sub-regular class of mappings which includes (or corresponds to) unbounded circumambient processes?

Furthermore, the discussion in this paper regarding the conjectured difference between weakly deterministic mappings and fully regular mappings raises interesting questions about intermediate representations. As the reader will recall, Elgot and Mezei (1965)'s result shows that, by using intermediate markup, any regular mapping can be characterised as the composition of a left-subsequential and right-subsequential mapping. As almost all segmental processes are weakly deterministic, they require no such markup, at least in the computational sense defined here. As Heinz and Lai (2013) point out, this idea may be brought to bear on the question of how abstract intermediate representations are in phonology. Thus, it is worth studying in more detail the nature of the composition of sub-regular classes of mappings, and how intermediate markup in the Elgot & Mezei sense affects the generative capacity of such compositions. The questions of representation raised at the end of §4.4 can be approached in a similar way—how do changes in representation correlate with changes in generative capacity?

Finally, how can the FLT insight presented here be incorporated into traditional phonological theory? As mentioned in §6.3, the weakly deterministic bound on segmental phonology could simply be stipulated as a restriction on the segmental mappings that autosegmental derivations or OT grammars generate. Another approach would be to appeal to learning, as raised in §5.3. For example, this could involve attempting to integrate the results in learning subsequential mappings into learning OT grammars in a manner similar to Tesar (2013)'s integration of output-driven maps into learning lexicons and OT grammars.

Unfortunately, going into these concerns in detail is beyond the scope of this paper. Instead, this paper's goal is similar to that of Kisseberth (1970)'s prophetic work on conspiracies in Yawelmani. Kisseberth writes that he is not "principally interested in proposing detailed formalism; instead I would like to encourage phonologists to look at the phonological component of a grammar in a particular way" (p.293 Kisseberth, 1970). Regardless of how it is incorporated into our previous understanding of phonology, the unbounded circumambient asymmetry between tonal and segmental phonology is robust, and the best available characterisation of this generalisation is one of computational complexity.

8 Appendix: Mathematical Definitions and Proof

8.1 Notation

Basic knowledge of set theory is assumed. An alphabet is a finite set of symbols; if Σ is an alphabet, let Σ^* denote the set of all finite strings, including the empty string λ, over Σ. Let $|w|$ denote the length of string w. If w and u are strings let wu denote their concatenation. If w is a string and X is a set of strings then let wX denote the set of strings resulting from concatenating w to each string in X. The PREFIXES of a string $w \in \Sigma^*$ are $Pr(w) = \{u \in \Sigma^* | \exists v \in \Sigma^* \text{ such that } w = uv\}$. The prefixes of a set of strings $L \subseteq \Sigma^*$ are $Pr_{set}(L) = w \in \cap_{x \in L} Pr(x)$.

The LONGEST COMMON PREFIX (lcp) of a set of strings is the longest prefix shared by all strings in the set: $lcp(L) = w$ such that $w \in Pr_{set}(L)$ and $\forall w' \in Pr_{set}(L), |w'| \leq |w|$. For example, $lcp(\{aaa, aab\}) = aa$, because aa is the longest prefix shared by both aaa

and aab.

If Σ and Δ are alphabets a RELATION is some subset of $\Sigma^* \times \Delta^*$. A relation R is a MAPPING (or FUNCTION) iff for all $w \in \Sigma^*$, $(w, v), (w, v') \in R$ implies $v = v'$.

The TAILS of x in given a relation R, denoted $T_R(x)$ are $T_R(x) = \{(y, v) | t(xy) = uv, u = lcp(t(x\Sigma^*))\}$. If R is a mapping, it is a SUBSEQUENTIAL mapping iff its sets of tails are finite; that is, the set $\bigcup_{w \in \Sigma^*} \{T_R(w)\}$ is of finite cardinality.

8.2 Subsequential Finite State Transducers

A finite-state transducer (FST) is a six-tuple $(q_i, F, Q, \Sigma, \Delta, \delta)$ where Q is the finite set of states, $q_i \in Q$ is the initial state, $F \subseteq Q$ is the set of final states, and $\delta \subseteq Q \times \Sigma^* \times \Delta^* \times Q$ is the transition function. The recursive extension of the transition function δ^* is defined as:

- $\delta \subseteq \delta^*$
- $(q, \lambda, \lambda, q) \in \delta^*$ for all $q \in Q$
- $(q, x, y, r) \in \delta^*$ and $(r, a, b, s) \in \delta$ implies $(q, xa, yb, s) \in \delta^*$

The relation that a FST describes is defined as $R(t) = \{(x, y) \in \Sigma^* \times \Delta^* \big| \exists q_f \in F$ such that $(q_i, x, y, q_f) \in \delta^*\}$.

A FST is DETERMINSITIC iff $\forall q \in Q$ and for all $q \in Q$ and $\sigma \in \Sigma$, $(q, \sigma, v, r), (q, \sigma, v', r') \in \delta$ implies $v = v'$ and $r = r'$. SUBSEQUENTIAL FSTs (SFSTs) are deterministic FSTs with an added output function $\omega : Q \to \Delta^*$ which specifies for each state an output string to be written when the machine ends on that state. Thus, a SFST is a 7-tuple $(q_i, F, Q, \Sigma, \Delta, \delta, \omega)$. The relation that a SFST describes is defined as $R(t) = \{(x, yz) \in \Sigma^* \times \Delta^* : \exists q_f \in F$ such that $(q_i, x, y, q_f) \in \delta^*$ and $\omega(q_f) = z\}$. Any such R is a subsequential mapping (Mohri, 1997).[22] For any subsequential mapping R there is a canonical SFST for which each state in the machine corresponds to a set of tails in R (Heinz and Lai, 2013).

8.3 Proof UTP is not subsequential

The proof is exactly Heinz and Lai (2013)'s proof for the non-subsequentiality of Sour Grapes.

Proof. The proof is by contradiction. Let UTP be the mapping discussed in the main text. The following shows that for all distinct $n, m \in \mathbb{N}$, $T_{UTP}(\text{H}\varnothing^m) \neq T_{UTP}(\text{H}\varnothing^n)$.

---

[22] Note that, given the definition of determinism, at any state q, there is not necessarily a transition on σ going out of state q. This means that while a SFST always describes a mapping, this could be a PARTIAL mapping—i.e., the domain is not equal to Σ^*. The distinction between partial and total mappings (where the domain *is* Σ^*) does not make a difference for questions of subsequentiality, so it will not be discussed here. Interested readers are encouraged to consult Beesley and Karttunen (2003, ch. 2).

As $\mathbb{N}$ is infinite, this means there must be infinitely many states in the canonical SFST for UTP, which contradicts the definition of FSTs.

If x =H∅, $lcp(UTP(x\Sigma^*))$ =H, because $UTP(x\Sigma^*)$ includes both HLL (which $= UTP($H∅∅$))$ and HHH (which $= UTP($H∅H$))$ and thus there no shared prefix of $UTP(x\Sigma^*)$ longer than H. Thus for all $n \neq 2,$ (∅,L^n) $\notin T_{UTP}$(H∅); i.e., (∅,LL) is the only possible tail with ∅ as the first member of the tuple.

If x =H∅∅, $lcp(UTP(x\Sigma^*))$ =H, because $UTP(x\Sigma^*)$ includes both HLLL and HHHH. Thus for all $n \neq 3$, (∅,L^n) $\notin T_{UTP}$(H∅); i.e., only (∅,LLL) is possible is the only possible tail with ∅ as the first member of the tuple.

We can see then that for any distinct $n \in \mathbb{N}$, (∅,L^k) $\in T_{UTP}$(H∅^{n}) only if $k = n + 1$. For $m \in \mathbb{N}, m \neq n$, (∅,L^j) $\in T_{UTP}$(H∅^{m}) only if $j = m + 1$. Thus for all distinct n and m, $k \neq j$, and so T_{UTP}(H∅^{n}) $\neq T_{UTP}$(H∅^{m}).

References

Al Khatib, Sam (2008). On the directionality of emphasis spread. In *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*.

Archangeli, Diana and Pulleyblank, Douglas (1994). *Grounded Phonology*. Cambridge: MIT Press.

Baković, Eric (2000). *Harmony, dominance and control*. Ph.D. thesis, Rutgers University.

Bane, Maximilian and Riggle, Jason (2009). The typological consequences of weighted constraints. In *CLS 45*.

Beckman, Jill (1997). Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology*, 14.

Beesley, Kenneth R. and Karttunen, Lauri (2003). *Finite State Morphology*. CSLI Publications.

Bennett, William (2013). *Dissimilation, Consonant Harmony, and Surface Correspondence*. Ph.D. thesis, Rutgers, the State University of New Jersey.

Beros, Achilles and de la Higuera, Colin (2014). A canonical semi-deterministic transducer. In *Proceedings of the 12th International Conference on Grammatical Inference (ICGI 2014)*, JMLR Workshop Proceedings, 33–48.

Bickmore, Lee and Kula, Nancy C. (2014a). Mutually-feeding iterative rules in Copperbelt Bemba Phrasal Tonology. Paper presented at the Eighth North American Phonology Conference.

Bickmore, Lee S. and Kula, Nancy C. (2013). Ternary spreading and the OCP in Copperbelt Bemba. *Studies in African Linguistics*, 42.

Bickmore, Lee S. and Kula, Nancy C. (2014b). Prosodic phrasing in Copperbelt Bemba. Ms., under review.

Bird, Steven and Ellison, T. Mark (1994). One-level phonology: Autosegmental representations and rules as finite automata. *Computational Linguistics*, 20.

Cassimjee, F. and Kisseberth, C.W. (2001). Zulu tonology and its relationship to other Nguni languages. In Shigeki Kaji (ed.), *Cross-linguistic studies of tonal phenomena: tonogenesis, Japanese accentology, and other topics.*, 327–359. Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Chandlee, Jane (2014). *Strictly Local Phonological Processes*. Ph.D. thesis, University of Delaware.

Chandlee, Jane, Athanasopoulou, Angeliki, and Heinz, Jeffrey (2012). Evidence for classifying metathesis patterns as subsequential. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, and Alex Trueman (eds.), *WCCFL 29*. Somerville, MA: Cascadilla Press.

Chandlee, Jane and Heinz, Jeffrey (2012). Bounded copying is subsequential: Implications for metathesis and reduplication. In *Proceedings of the 12th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology*, 42–51. Association for Computational Linguistics, Montreal, Canada.

Chomsky, Noam (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 113–124.

Clements, G. N. (1976). *Vowel Harmony in Nonlinear Generative Phonology: An Autosegmental Model*. Bloomington: Indiana University Linguistics Club Publications.

Clements, G. N. (1977). Neutral vowels in Hungarian vowel harmony: an autosegmental interpretation. In *NELS 7*, 49–64.

Do, Young Ah and Kenstowicz, Michael (2011). A note on phonological phrasing in South Kyungsang. Ms., MIT.

Downing, Laura (2001). How ambiguity of analysis motivates stem change in Durban Zulu. *UBC Working Papers in Linguistics*.

Elgot, C. C. and Mezei, J. E. (1965). On relations defined by generalized finite automata. *IBM Journal of Research and Development*, 9. 47—68.

Folia, Vasiliki, Uddn, Julia, De Vries, Meinou, Forkstam, Christian, and Petersson, Karl Magnus (2010). Artificial language learning in adults and children. *Language Learning*, 60. 188–220. ISSN 1467-9922. doi:10.1111/j.1467-9922.2010.00606.x.

Fortnow, Lance (2009). The status of the p versus np problem. *Communications of the ACM*, 52:9. ISSN 0001-0782.

Gafos, Adamantios (1999). *The articulatory basis of locality in phonology.* Garland.

Gainor, Brian, Lai, Regine, and Heinz, Jeffrey (2012). Computational characterizations of vowel harmony patterns and pathologies. In *WCCFL 29*, 63–71.

Goldsmith, John A. (1990). *Autosegmental & Metrical Theory*. Basil Blackwell, Inc.

Good, Jeff (2004). Tone and accent in Saramaccan: charting a deep split in the phonology of a language. *Lingua*, 114. 575–619.

Graf, Thomas (2010). *Logics of Phonological Reasoning*. Master's thesis, University of California, Los Angeles.

Hale, Ken and Selkirk, Elisabeth (1987). Government and tonal phrasing in Papago. *Phonology Yearbook*, 4. 151–183.

Hansson, Gunnar Ólafur (2001). *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. thesis, University of California, Berkeley.

Hansson, Gunnar Ólafur (2010). *Consonant harmony: Long-distance interaction in phonology*. Berkeley: University of California Press.

Heinz, Jeffrey (2010). Learning long-distance phonotactics. *LI*, 41. 623–661.

Heinz, Jeffrey (2011). Computational phonology part I: Foundations. *Language and Linguistics Compass*, 5:4. 140–152.

Heinz, Jeffrey and Idsardi, William (2011). Sentence and word complexity. *Science*, 333:6040. 295–297.

Heinz, Jeffrey and Idsardi, William (2013). What complexity differences reveal about domains in language. *Topics in Cognitive Science*, 5:1. 111–131.

Heinz, Jeffrey, Kobele, Gregory, and Riggle, Jason (2009). Evaluating the complexity of Optimality Theory. *LI*, 40:2. 277–288.

Heinz, Jeffrey and Lai, Regine (2013). Vowel harmony and subsequentiality. In Andras Kornai and Marco Kuhlmann (eds.), *Proceedings of the 13th Meeting on Mathematics of Language*. Sofia, Bulgaria.

Hopcroft, John, Motwani, Rajeev, and Ullman, Jeffrey (2006). *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, third edition.

Hyman, Larry (1998). Positional prominence and the 'prosodic trough' in Yaka. *Phonology*, 15. 14–75.

Hyman, Larry (2011). Tone: Is it different? In John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.), *The Blackwell Handbook of Phonological Theory*, 197–238. Wiley-Blackwell.

Hyman, Larry, Katamba, Francis, and Walusimbi, Livingstone (1987). Luganda and the strict layer hypothesis. *Phonology Yearbook*, 4. 87–108.

Hyman, Larry and Katamba, Francis X. (2010). Tone, syntax and prosodic domains in Luganda. In Laura Downing, Annie Rialland, Jean-Marc Beltzung, Sophie Manus, Cdric Patin, and Kristina Riedel (eds.), *Papers from the Workshop on Bantu Relative Clauses*, volume 53 of *ZAS Papers in Linguistics*, 69–98. ZAS Berlin.

Idsardi, William J. (2006). A simple proof that Optimality Theory is computationally intractable. *LI*, 37. 271–275.

Inkelas, Sharon (1995). The consequences of Optimization for underspecification. In Eugene Buckley and S. Iatridou (eds.), *NELS 25*, 287–302.

Ito, Junko, Mester, Armin, and Padgett, Jaye (1995). Licensing and underspecification in Optimality Theory. *LI*, 571–613.

Johnson, C. Douglas (1972). *Formal aspects of phonological description*. Mouton.

Kaplan, Ronald and Kay, Martin (1994). Regular models of phonological rule systems. *Computational Linguistics*, 20. 331–78.

Kay, Martin (1987). Nonconcatenative finite-state morphology. In *Proceedings, Third Meeting of the European Chapter of the Association for Computational Linguistics*, 2–10.

Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Blackwell Publishing.

Kenstowicz, Michael and Kisseberth, Charles (1990). Chizigula tonology: the word and beyond. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, 163–194. Chicago: the University of Chicago Press.

Kidima, Lukowa (1990). Tone and syntax in Kiyaka. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, 195–216. Chicago: the University of Chicago Press.

Kidima, Lukowa (1991). *Tone and Accent in KiYaka*. Ph.D. thesis, University of California, Los Angeles.

Kisseberth, C. W. (1984). Digo tonology. In G.N. Clements and John A. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, 105–182. Foris Publications.

Kisseberth, Charles and Odden, David (2003). Tone. In Derek Nurse and Gérard Philippson (eds.), *The Bantu Languages*. New York: Routledge.

Kisseberth, Charles W. (1970). On the functional unity of phonological rules. *LI*, 1:3. 291–306. ISSN 00243892.

Kornai, András (1991). *Formal Phonology*. Ph.D. thesis, Stanford University.

Kornai, András (1995). *Formal Phonology*. Garland Publication.

Lai, Regine (2012). *Domain Specificity in Learning Phonology*. Ph.D. thesis, University of Delaware.

Lai, Regine (to appear). Learnable versus unlearnable harmony patterns. *LI*.

Laughren, Mary (1984). Tone in Zulu nouns. In G.N. Clements and John A. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, 183–235. Foris Publications.

Leben, W. R. (1973). *Suprasegmental phonology*. Ph.D. thesis, Massachussets Institute of Technology.

Levergood, Barbara (1984). Rule-governed vowel harmony and the strict cycle. In *NELS 14, Amherst, MA*, 275–293.

Macdonell, Arthur (1910). *Vedic grammar*. Trūbner.

McCarthy, John (2010). Autosegmental spreading in Optimality Theory. In John A. Goldsmith, Elizabeth Hume, and W. Leo Wezels (eds.), *Tones and Features: Phonetic and Phonological Perspectives*, 195–222. De Gruyer Mouton.

McCarthy, John and Prince, Alan (1995). Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, number 18 in University of Massuchusetts Occasional Papers in Linguistics, 249—384. University of Massachusetts.

McCarthy, John J. (1986). OCP effects: gemination and antigemination. *LI*, 17. 207–263.

McMullin, Kevin and Hansson, Gunnar Ólafur (2015). Locality in long-distance phonotactics: evidence for modular learning. In *NELS 44, Amherst, MA*. In press.

McWhorter, John and Good, Jeff (2012). *A Grammar of Saramaccan Creole*. De Gruyter Mouton.

Medvedev, Y. T. (1964). On the class of events representable in a finite automaton. In E. F. Moore (ed.), *Sequential machines – Selected papers*, 215–227. New York: Addison-Wesley.

Meyers, Scott (1997). OCP effects in Optimality Theory. *NLLT*, 15:4. 847–892.

Mielke, Jeff (2004). P-Base 1.95. http://137.122.133.199/jeff/pbase.

Mohri, Mehryar (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, 23:2. 269–311.

Monier-Williams, Monier (1899). *A Sanskrit-English Dictionary*. Great Britain: Oxford University Press. Maintained online at http://www.sanskrit-lexicon.uni-koeln.de/.

Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology*, 25. 83–127.

Moreton, Elliott and Pater, Joe (2012). Structure and substance in artificial-phonology learning. part i: Structure. *Language and Linguistics Compass*, 6. 686—701.

Nevins, Andrew (2010). *Locality in Vowel Harmony*. MIT Press.

Odden, David (1986). On the role of the Obligatory Contour Principle in phonological theory. *Lg*, 62:2. 353–383.

Odden, David (1994). Adjacency parameters in phonology. *Lg*, 70:2. 289–330.

Onn, Farid M. (1980). *Aspects of Malay Phonology and Morphology: A Generative Approach*. Kuala Lampur: Universiti Kebangsaan Malaysia.

Padgett, Jaye (1995). Partial class behavior and nasal place assimilation. In K. Suzuki and D. Elzinga (eds.), *Proceedings of the 1995 Southwestern Workshop on Optimality Theory*.

Pater, Joe (2004). Austronesian nasal substitution and other NC effects. In John McCarthy (ed.), *Optimality Theory in Phonology: A Reader*, 271–289. Oxford and Malden, MA: Blackwell.

Payne, Amanda (2014). Dissimilation as a subsequential process. In *NELS 44*.

Prince, Alan and Smolensky, Paul (1993). Optimality Theory: Constraint interaction in generative grammar. *Rutgers University Center for Cognitive Science Technical Report*, 2.

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

Pulleyblank, Douglas (1996). Neutral vowels in Optimality Theory: A comparison of Yoruba and Wolof. *Canadian Journal of Linguistics*, 41. 295–347.

Rabin, M. O. and Scott, D. (1959). Finite automata and their decision problems. *IBM Journal of Research and Development*, 3:2. 114–125. ISSN 0018-8646. doi: 10.1147/rd.32.0114.

Ringen, Catherine O. and Vago, Robert M. (1998). Hungarian vowel harmony in Optimality Theory. *Phonology*, 393–416.

Rogers, James and Hauser, Marc D. (2010). The use of formal language theory in studies of artificial language learning: a proposal for distinguishing the differences between human and nonhuman animal lerners. In Harry van der Hulst (ed.), *Recursion and Human Language (Studies in Generative Grammar [SGG]) 104*. de Gruyter.

Rogers, James, Heinz, Jeffrey, Fero, Margaret, Hurst, Jeremy, Lambert, Dakotah, and Wibel, Sean (2013). Cognitive and sub-regular complexity. In *Formal Grammar*, volume 8036 of *Lecture Notes in Computer Science*, 90–108. Springer.

Rogers, James and Pullum, Geoffrey (2011). Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information*, 20. 329–342.

Rose, Sharon and Walker, Rachel (2004). A typology of consonant agreement as correspondence. *Lg*, 80. 475–531.

Roundtree, S. Catherine (1972). Saramaccan tone in relation to intonation and grammar. *Lingua*, 29. 308–325.

Ryan, Kevin (2015). Attenuated spreading in sanskrit retroflex harmony. Ms., Harvard University.

Safir, Ken (1982). Nasal spreading in Capanahua. *LI*, 13. 689–694.

Schein, Barry and Steriade, Donca (1986). On geminates. *LI*, 17. 691–744.

Shih, Stephanie and Inkelas, Sharon (2014). A subsegmental correspondence approach to contour tone (dis)harmony patterns. In John Kingston, Claire Moore-Cantwell, Joe Pater, and Robert Staubs (eds.), *Proceedings of the 2013 Meeting on Phonology (UMass Amherst)*, Proceedings of the Annual Meetings on Phonology. LSA.

Staubs, Robert (2014). *Computational modeling of learning biases in stress typology*. Ph.D. thesis, University of Massachusetts, Amherst.

Suzuki, Keiichiro (1998). *A typological investigation of dissimilation*. Ph.D. thesis, University of Arizona, Tucson.

Tesar, Bruce (2013). *Output-Driven Phonology: Theory and Learning*. Cambridge Studies in Linguistics.

Walker, Rachel (1998). *Nasalization, neutral segments, and opacity effects*. Ph.D. thesis, University of California, Santa Cruz.

Whitney, William D. (1889). *Sanskrit grammar*. Oxford University Press.

Wiebe, Bruce (1992). *Modelling Autosegmental Phonology with Multi-Tape Finite State Transducers*. Master's thesis, Simon Fraser University.

Wilson, Colin (2003). Analyzing unbounded spreading with constraints: marks, targets, and derivations. Unpublished ms.

Wilson, Colin (2006a). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30. 945–982.

Wilson, Colin (2006b). Unbounded spreading is myopic. Paper presented at the workshop Current Perspectives on Phonology.

Yip, Moira (2002). *Tone*. Cambridge University Press.

Yli-Jyrä, Anssi (2013). On finite-state tonology with autosegmental representations. In *Proceedings of the 11th International Conference on Finite State Methods and Natural Language Processing*, 90–98. Association for Computational Linguistics.