Vowel Harmony and Subsequentiality

Jeffrey Heinz and **Regine Lai**
University of Delaware
{heinz,rlai}@udel.edu

Abstract

Attested and 'pathological' vowel harmony patterns are studied in the context of subclasses of regular functions. The analysis suggests that the computational complexity of phonology can be reduced from regular to weakly deterministic.

1 Introduction

The expressivity of ordered rewrite-rule grammars for phonology (Chomsky and Halle, 1968, henceforth SPE) and two-level phonology (Koskenniemi, 1983) are exactly the class of regular relations (Johnson, 1972; Kaplan and Kay, 1994; Beesley and Kartunnen, 2003). Since SPE-style grammars can express virtually any phonological generalization, it follows that the generalizations themselves are regular, even if they are represented with other formalisms (such as OT grammars).

This result can be interpreted cognitively as establishing a universal property of phonological patterns: humanly possible phonological patterns are regular. If correct, this would mean, for example, that humanly possible syntactic patterns which are nonregular are not humanly possible *phonological* patterns (Heinz and Idsardi, 2011; Lai, 2012; Heinz and Idsardi, 2013).

Recent research suggests that stronger universals than "being regular" can be established for phonology. It has been shown that segmental phonotactic patterns are star-free (Heinz et al., 2011), as are virtually all stress patterns (Rogers et al., to appear), and the semantics of two-level rules appear to ensure that these mappings have star-free-like properties, provided the contexts to the rules are star-free (Yli-Jyrä and Koskenniemi, 2006).[1]

This paper examines the hypothesis that subsequentiality is a necessary property of phonological patterns by studying theories of iterative vowel harmony. Informally, a function is left (right) subsequential if there is a finite-state transducer describing the function which processes strings from left to right (right to left) deterministically on the input. We use the term 'subsequentiality' to mean *either* left or right subsequential.

Previous work has found that synchronically attested metathesis patterns and partial reduplication patterns are either left or right subsequential (Chandlee et al., 2012; Chandlee and Heinz, 2012). Also Gainor et al. (2012) show that the vowel harmony generalizations in Nevins (2010) are also left or right subsequential mappings. Gainor et al.'s analysis, while insightful, is incomplete since Nevin's theory of vowel harmony invokes underspecification and other theories of vowel harmony do not. Phonological underspecification is explained in section 3. The linguistic generalizations examined in this paper come from two types of theories of vowel harmony patterns in linguistics which do not use underspecification: traditional directional theories and dominant/recessive/stem-control theories.

We prove that subsequentiality separates directional theories from logically possible but 'pathological' vowel harmony patterns (Wilson 2003). (This claim was also made by Gainor et al. without proof.) It is also shown that dominant/recessive/stem-control theories posit generalizations which are neither left nor right subsequential, but which are *weakly deterministic*. Informally, this means that these generalizations can be decomposed into a left subsequential and right subsequential function without the left-subsequential function marking up its output in any special way. We conjecture this is

---

[1] There are multiple ways to generalize the class of star-free regular sets to regular relations (Benedikt et al., 2001) so it would be interesting to determine more exactly the nature of such two-level rules.

not the case for the pathological patterns. Since subsequential and weakly deterministic functions are proper subclasses of regular relations, these results suggest concretely how the computational complexity of phonology established by earlier researchers can be improved.

Mathematical and phonological preliminaries are given in sections 2 and 3, respectively. Sections 4, 5, and 6 consider the vowel harmony patterns with respect to the regular and subsequential boundaries. and weakly deterministic boundaries, respectively. Section 7 concludes.

2 Preliminaries

2.1 Regular relations and functions

If X denotes a finite alphabet then X^* and X^n denotes the sets of all finite strings and the set of all strings of length n over X, respectively. The length of a string w is $|w|$. The unique string of length zero is denoted λ. A string w of length k can be written $w_1 w_2 \cdots w_k$, where w_i is the ith letter of w. The reverse of a string $w = w_1 \cdots w_k$ is $w^r = w_k \cdots w_1$. For finite alphabets X and Y, a *relation* is a subset of $X^* \times Y^*$. If R is a relation, the *reverse relation* $R^r = \{(x^r, y^r) \mid (x, y) \in R\}$. Note the reverse relation is *not* the inverse relation. A relation R is *length-preserving* iff for all $(x, y) \in R$ it is the case that $|x| = |y|$. It is *length-increasing* iff there exists $(x, y) \in R$ such that $|x| < |y|$.

Any relation $R \subseteq X^* \times Y^*$ is a *function* iff for all $x \in X^*$, there is at most $y \in Y^*$ such that $(x, y) \in R$. In this case, we often write $R : X^* \to Y^*$. For two functions $f : X^* \to Y^*$ and $g : Y^* \to Z^*$, the *composition* of f and g is a function $h : X^* \to Z^*$ such that $h(x) = g(f(x))$. We write $h = g \circ f$.

For all $x \in X^*$, the prefixes of x are $Pr(x) = \{u \in X^* \mid \exists v \in X^* \text{ such that } x = uv\}$. For any set $L \subseteq X^*$, the longest common prefix of L is

$$lcp(L) = w \Leftrightarrow w \in \bigcap_{x \in L} Pr(x) \quad \wedge$$
$$\left(\forall w' \in \bigcap_{x \in L} Pr(x)\right)[|w'| \leq |w|] \quad (1)$$

Regular relations are those describable by *finite-state transducers* (FSTs).[2] A finite-state transducer T is a tuple (Q, X, Y, I, F, δ) where

[2]In the algebraic theory of automata, these are called *rational* relations (Berstel, 1979; Sakarovitch, 2009).

Q is a finite set of states, X and Y are finite alphabets, $I, F \subseteq Q$ are the initial and final states, respectively, and $\delta \subseteq Q \times X^* \times Y^* \times Q$ is the transition function. For all FSTs, the transition function δ is recursively extended to δ^* in the usual way. The relation that a finite state transducer $T = (Q, X, Y, I, F, \delta)$ recognizes/accepts/generates is

$$R(T) = \Big\{ (x, y) \in X^* \times Y^* \mid (\exists q_i \in I)$$
$$(q_f \in F) \big[(q_i, x, y, q_f) \in \delta^* \big] \Big\} . \quad (2)$$

Let RR denote the class of regular relations. This class is closed under concatenation, Kleene closure, union, composition, and inversion, but not intersection or complement.

A *subsequential* transducer is a tuple $(Q, q_0, X, Y, \sigma, \delta)$, where Q is a finite set of states, X and Y are finite alphabets, $q_0 \in Q$ is the initial state, and $\sigma \subseteq Q \times Y^*$ is the output function. Informally, subsequential transducers are weighted acceptors that are deterministic on the input, and where the weights are strings and multiplication is concatenation.

Formally, the transition function $\delta \subseteq Q \times X \times Y^* \times Q$ is deterministic:

$$(q, a, u, r), (q, a, v, s) \in \delta \Rightarrow u = v \wedge r = s .$$

The transition function δ is also recursively extended to δ^*. The relation that a subsequential transducer $T = (Q, q_0, X, Y, \sigma, \delta)$ recognizes/accepts/generates is

$$R(t) = \Big\{ (x, yz) \in X^* \times Y^* \mid (\exists q \in F)$$
$$\big[(q_0, x, y, q) \in \delta^* \wedge z = \sigma(q) \big] \Big\} . \quad (3)$$

Since subsequential transducers are deterministic, the relations they recognize are functions.

Functions recognized by subsequential transducers are called *left subsequential*. A function f is *right subsequential* iff its reverse f^r is left subsequential. Observe that for all $w \in X^*$, the image of a right subsequential function f of w can be calculated by reversing w, processing the result with the subsequential transducer T recognizing f^r, and then reversing the result. Formally,

$$(\forall w \in X^*)[f(w) = T(w^r)^r]. \quad (4)$$

However, there is another way to state the above without the reversing function $(\cdot^r)$. This is to recognize that $f(w)$ can be computed by applying T

to w from right to left, instead of from left to right. When T processes w right to left, we write $\overleftarrow{T}(w)$ and when T processes w left to right, we write $\overrightarrow{T}(w)$. Then we can restate Equation 4 as

$$(\forall w \in X^*)[f(w) = \overleftarrow{T}(w)]. \qquad (5)$$

Let LSF and RSF denote the class of left and right subsequential functions, respectively.

Theorem 1 (Mohri 1997) *The following hold:*

1. $LSF, RSF \subsetneq RR$.

2. $RSF^r = LSF$.

3. LSF and RSF are incomparable.

There are some relevant subclasses and generalizations of subsequentiality. A subsequential transduction T is *sequential* iff for all $q \in Q$, it is the case that $\sigma(q) = \lambda$.[3] Mohri (1997) generalizes subsequentiality to p-subsequentiality (allowing up to p outputs for each input), preserving many important properties. Mohri's generalizations are important here because there are likely to be a bounded number of exceptions, or optional forms, in actual vowel harmony systems that fall outside the purview of the 1-subsequential analysis presented here, but which would presumably not fall outside a p-subsequential analysis (not presented here).

Elgot and Mezei proved the following.

Theorem 2 (Elgot and Mezei 1965) *Let* $T : X^* \rightarrow Y^*$ *be a function. Then* $T \in RR$ *iff there exists* $L : X^* \rightarrow Z^* \in LSF$, *and* $R : Z^* \rightarrow Y^* \in RSF$ *with* $X \subseteq Z$ *such that* $T = R \circ L$.

What this decomposition means is that the computation of $T(x) = y$ can be accomplished by (1) reading x sequentially from left to right with a subsequential transducer, which transforms it into a word z possibly marking it up with additional symbols; (2) reading the resulting word z from right to left with another subsequential transducer and writing from right to left the final output y. As their proof makes clear, this decomposition of T is possible because the alphabet Z may be strictly

larger than X, and so z can be marked-up with additional symbols which carry additional information.[4]

Finally, we review one important property of subsequential transducers and regular sets. For any function $f : X^* \rightarrow Y^*$ and $x \in X^*$, let the *tails* of x in f be defined as

$$TL_f(x) = \{(y,v) \mid f(xy) = uv \wedge \\ u = lcp(f(xX^*))\}. \qquad (6)$$

Every subsequential transducer T computing a function f admits a *canonical* form, where the states of T are in one-to-one correspondence with $TL_f(x)$ for all $x \in X^*$.

Theorem 3 (Oncina et al. 1993) $f \in LSF \Leftrightarrow \{TL_f(x) \mid x \in X^*\}$ *has finite cardinality.*

This theorem is the functional counterpart to the Myhill/Nerode relation. Recall that for any set of strings L, the tails of a word w with respect to L is defined as $TL_L(w) = \{u \mid wu \in L\}$. This relation partitions the set of all logically possible strings into a finite set of equivalence classes iff the set L is regular. These equivalence classes are the basis for constructing the smallest deterministic acceptor for a regular language.

Similarly, in the construction of the canonical subsequential transducer for a left subsequential function, the states correspond to the sets of tails defined in (6) above. There is a rich literature on subsequential functions (Elgot and Mezei, 1965; Berstel, 1979; Oncina et al., 1993; Mohri, 1997; Roche and Schabes, ; Sakarovitch, 2009).

2.2 Weak Determinism

Here we introduce the notion of *weak determinism*. Informally, these are regular functions which decompose into left and right subsequential functions as in Elgot and Mezei's theorem but *without* the mark-up given by the intermediate, larger alphabet. Thus, they are not necessarily deterministic, but they are "more" deterministic than regular functions where Elgot and Mezei decomposition *requires* the intermediate mark-up.

While the mark-up can be accomplished by introducing new symbols (as done in Elgot and

---

[3]Sakarovitch (2009) prefers the term 'sequential' for subsequential functions and the term 'pure sequential' for sequential functions. While his arguments are reasonable (pp. 651-2), we adopt the more widely adopted terminology.

[4]Berstel (1979) provides an updated proof. This book is out of print but the first four chapters are available for download at `http://www-igm.univ-mlv.fr/˜berstel/LivreTransductions/LivreTransductions.html`. The theorem and proof begin on page 117 in the online version and on page 126 in the printed version.
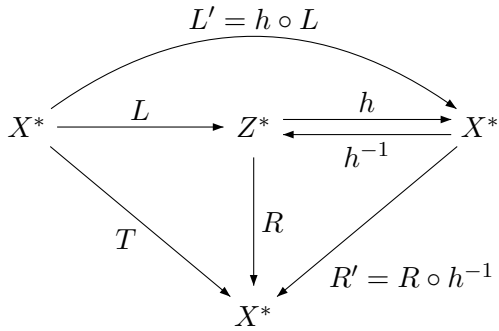
Figure 1: Decompositions of regular function T with $X \subseteq Z$.

Mezei's proof), for any alphabet with at least two symbols, the mark-up can also be accomplished by *coding* these new symbols as strings formed over the original alphabet.[5] Figure 1 illustrates. Let X be a finite alphabet containing the symbols $\{a, b\}$. Consider any *alphabet-preserving* regular function $T : X^* \to X^*$. By Elgot and Mezei's theorem, there exists left subsequential $L : X^* \to Z^*$ and right subsequential $R : Z^* \to X^*$ with $X \subseteq Z$ such that $T = R \circ L$. Let $h : Z^* \to X^*$ be a function which encodes each word w in Z^* by coding each symbol in w as follows. Assume some enumeration of the symbols in Z and the rewrite the nth symbol of Z as $ab^n a$. For example if $Z = \{a, b, c\}$ and $w = cab$ then $h(w) = abbaaaaba$. It is not difficult to verify that h is length-increasing and that both h and h^{-1} are subsequential functions. Letting $R' = R \circ h^{-1}$ and $L' = h \circ L$, it follows that $T = R' \circ L'$ and that both R' and L' have domain and co-domain X^*.

For this reason, it is not sufficient to require that the decomposition be *alphabet-preserving* (i.e. $Z = X$) to avoid any mark-up. It is also necessary that the first factor L not be *length-increasing*. This is because the only way to unambiguously encode a larger alphabet into a smaller one is with length-increasing functions (like h in the above example).

Definition 1 *A regular function T is* weakly deterministic *iff there exists $L : X^* \to X^* \in LSF$, and $R : X^* \to X^* \in RSF$ such that L is not length-increasing and $T = R \circ L$. The class of weakly deterministic functions is denoted* WD.

The corollary below is immediate from this definition and Elgot and Mezei's theorem.

Corollary 1 *$LSF, RSF \subseteq WD \subseteq RR$.*

---

	noun	genitive	gloss
a.	ip	ip-in	*rope*
b.	el	el-in	*and*
c.	son	son-un	*end*
d.	pul	pul-un	*stamp*

Table 1: Examples illustrating a fragment of the Vowel harmony from Turkish (Nevins 2010:32).

(a) w	$f(w)$	(b) w	$f(w)$
/ip-un/	[ip-in]	/−C+C/	[−C−C]
/el-un/	[el-in]		
/son-un/	[son-un]	/C+C+C/	[C+C+C]
/pul-un/	[pul-un]	...	
...			

Table 2: Examples showing fragments of the phonological function describing Turkish back harmony assuming the underlying genitive morpheme is /-un/.

The vowel harmony analysis below is sufficient to go a step further and demonstrate a separation between LSF, RSF on one side and WD on the other. We conjecture that one unattested 'pathological' vowel harmony patterns separates WD from RR.

3 Vowel Harmony

Vowel harmony is a pattern wherein vowels assimilate with respect to some feature. Table 1 shows two allomorphs of the genitive suffix, [-in] and [-un]. The allomorph is predictable based on the front/back dimension of the preceding vowel: if the preceding vowel is front then [-in] occurs, but if it is back then [-un] occurs. (Turkish also has rounding harmony, which is not shown here.)

Phonological analysis conceives of the phonological grammar as a function which maps an abstract lexical representation (the 'underlying form') to a more concrete—but still abstract—phonological representation (the 'surface form') (Hyman, 1975; Kenstowicz, 1994; Hayes, 2009). Phonological transcriptions, like the ones in Table 1, represent surface forms.

To illustrate, consider a simple phonological analysis of the Turkish forms above, which posits the underlying form of the genitive suffix to be /-un/ and a mapping f which derives the surface forms as shown in Table 2(a). Table 2(b) describes the mapping only in terms of the relevant details where [+] indicates [+ back] vowels, [−] indi-

cates [−back] vowels, and C consonants.

Nevins' (2010) analysis of vowel harmony utilizes underspecification. We illustrate this concept ostensively with the Turkish example above. Instead of positing the underlying form of the suffix to be either /-in/ or /-un/, underspecification theory would posit it to be /-Vn/ where V is high, unrounded vowel unspecified for backness. In Nevins' theory, underlying vowels which have feature specifications can spread those features only to vowels unspecified for those features. Underspecification is not congruent with research in Optimality Theory (Prince and Smolensky, 2004), which, by the principle of the rich base, requires every underlying form to be considered (including those where every vowel is fully specified). Gainor et al. (2012) show that the iterative mappings Nevins describes are subsequential, but this does not address those theories (like OT) which may not consider underlying forms to permit underspecification. All the vowel harmony patterns considered in this paper do not admit any underspecification whatsoever.

A traditional view of vowel harmony analyzes vowel harmony patterns as either instances of *progressive harmony* (PH) or *regressive harmony* (RH). Informally, progressive harmony means the value of a feature can be thought to spread from left to the right (as in the Turkish example above). Conversely, regressive harmony can be thought of as spreading from right to left. This is illustrated with examples (a-d) in Table 3.

Other theories of vowel harmony reject that directionality is a primitive of the theory and argue that vowel harmony is either *dominant/recessive* (DR) or *stem-controlled* (SC) (Baković, 2000; Krämer, 2003) (see also (Archangeli and Pulleyblank, 1994)). Dominant/recessive theories analyze vowel harmony patterns by postulating that a particular feature value of a harmonizing feature is the dominant one. The DR function in Table 3 identifies the [+] value as the dominant one; so any underlying representation containing the harmonizing feature with the value [+] will surface so that the harmonizing feature in all vowels will also be [+]. Stem-controlled analyses are similar to dominant-recessive theories, however the feature that spreads is determined not by its value but instead by the morphological position of the vowel to which the feature belongs (for instance, is the vowel in a stem or affix?).

An additional complication is that variations of the above functions are introduced by *neutral* vowels, which never undergo harmony. They come in two kinds: *transparent* vowels which permit features to spread through them, and *opaque* vowels, which block the spread of harmony, but trigger their own harmony domain. Some effects of neutral vowels are shown in rows (e-f) in Table 3. (Symbols [⊟] and [⊖] are [−F] vowels that are opaque and transparent, respectively. Likewise we use [⊞] and [⊕] to denote opaque and transparent vowels which are [+F].)

Additionally, the phonological literature includes discussion of logically possible, unattested and unnatural vowel harmony patterns that are predicted by classical approaches to vowel harmony in OT. These patterns include *sour grapes* (SG) (Padgett, 1995; Wilson, 2003) and *majority rules* (MR) (Lombardi, 1999; Baković, 2000). Informally, SG is like progressive harmony except that vowels only harmonize if every vowel is guaranteed to harmonize. For example, if an opaque vowel occurs after the trigger, no non-neutral vowel harmonizes with the trigger. Majority Rules instantiates the following rule: If the number of segments with αF is greater than the number of segments with $-\alpha F$, then segments with αF are the triggers of harmony and segments with $-\alpha F$ are the targets and undergo change. Because phonologists consider SG and MR to be bizarre, they are referred to as 'pathologies' (Wilson, 2003; Wilson, 2004; Finley, 2008) and it is a strike against a theory if it predicts the existence of either SG or MR.

Henceforth, let $X = Y = \{+, -, C, \boxminus, \boxplus\}$. These symbols represent equivalence classes of a partition of the phonemic inventory of any language which exhibits progressive harmony for the feature F. The symbols $+$ and $-$ represent the classes of all harmonizing vowels which are $+F$ and $-F$, respectively. Phonemes invisible to harmony are in the C class; this includes consonants and transparent vowels. The symbol $\boxplus$ ($\boxminus$) refers to opaque vowels which are $+F$ ($-F$), which block the spread of $-F$ ($+F$), and which begins a new harmonic domain spreading $+F$ ($-F$).

The vowel harmony mappings in this paper are all, in fact, same-length relations. Furthermore, they are sequential. These additional properties do not appear to be shared by other phonological processes. Epenthesis and deletion are common

	w	PH(w)	RH(w)	DR(w)	SG(w)	MR(w)
a.	/+ − −/	[+ + +]	[− − −]	[+ + +]	[+ + +]	[− − −]
b.	/− + +/	[− − −]	[+ + +]	[+ + +]	[− − −]	[+ + +]
c.	/− − −/	[− − −]	[− − −]	[− − −]	[− − −]	[− − −]
d.	/− + −/	[− − −]	[− − −]	[+ + +]	[− − −]	[− − −]
e.	/+ − ⊟/	[+ + ⊟]	[− − ⊟]	[+ + ⊟]	[+ − ⊟]	[− − ⊟]
f.	/+ ⊖ −/	[+ ⊖ +]	[− ⊖ −]	[+ ⊖ +]	[+ ⊖ +]	[− ⊖ −]

Table 3: Example mappings of underlying forms (w) given by progressive harmony (PH), regressive harmony (RH), dominant/recessive harmony (DR), sour grapes harmony (SG), and majority rules harmony (MR). Symbols [+] indicates a [+F] vowel and [−] indicates a [−F] vowel where "F" is the feature harmonizing. Symbols [⊟] and [⊖] are [−F] vowels that are opaque and transparent, respectively.

cross-linguistically, and the metathesis patterns analyzed by Chandlee and Heinz (2012) are not sequential (though they are subsequential). For this reason, we keep the analysis focused at the level of subsequentiality.

4 The regular boundary

In this section we show that the regular boundary is sufficient to distinguish the pathological Majority Rules pattern from the attested progressive and regressive harmony patterns.

Formally, MR functions can be defined as follows. Let $|w|_{+F}$ and $|w|_{-F}$ denote the number of participating vowels (i.e. non transparent vowels) which are $+F$ and $-F$, respectively, in $w \in X^*$. Then we define a Majority Rules Harmony pattern as any same-length function which at a minimum obeys the following:

$$MR(w) = \begin{cases} +^{|w|} \text{ if } |w|_{+F} > |w|_{-F} \\ -^{|w|} \text{ if } |w|_{-F} > |w|_{+F} \end{cases} \quad (7)$$

The result below seems to be widely known (see Riggle (2004, chapter 7, section 5)) though we have not been able to find a proof in print.

Theorem 4 *Majority Rules is not regular.*

Proof By way of contradiction, suppose that MR is a regular relation. Since regular relations are closed under inverse, so is MR^{-1}. The image of a regular set under a regular relation is also a regular set (see Roche and Schabes (, pp. 41-43)). Therefore, $MR^{-1}(+^*)$ is a regular set. Since regular sets are closed under intersection, it follows that $MR^{-1}(+^*) \cap (+^* -^*)$ is regular as well. Call this set S.

However, S is in fact *not* a regular set. Since MR is length preserving, for all odd $k \in \mathbb{N}$, it
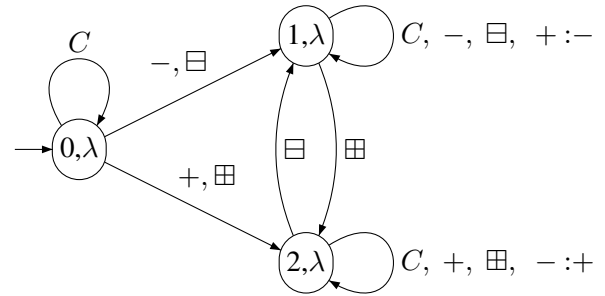

Figure 2: A subsequential transducer which recognizes iterative, progressive harmony.

is the case that $MR(+^k) \cap (+^* -^*) = +^m -^n$ where $0 \leq m, n \leq k$ and $m + n = k$ and $m > n$. Furthermore, for all for all odd k, it is the case that $TL_S(+^k)$, includes $-^n$ for all $n < k - 1$ but excludes $-^n$ for all $n > k + 1$. Thus there is a distinct Nerode-equivalence class for each odd k, and hence S is not a regular set, and therefore MR is not a regular relation. □

On the other hand, progressive and regressive harmony are regular; in fact, subsequential. For concreteness we analyze a canonical progressive harmony pattern which includes neutral vowels.

The subsequential transducer T_{PH} in Figure 2 faithfully captures the PH function. Labels on the transitions are interpreted as in Beesley and Kartunnen (2003): commas delimit multiple transitions; the label $x{:}y$ means x is the input and y the output; absence of a colon means the input and output are identical. The rightmost symbol interior to a state is the output of the σ function.

While T_{PH} presupposes languages have opaque vowels, it can be modified as needed to remove this assumption without losing subsequentiality. If a language has no opaque vowels, those transitions can be removed. Since any subgraph of the transducer shown in Figure 2 is also subsequen-

tial, this establishes the left subsequentiality of iterative, progressive harmony patterns without underspecification.[6]

As for iterative regressive harmony patterns, they are simply the reverse of iterative progressive patterns. In other words, for all $w \in \Sigma^*$, $RH(w) = \overleftarrow{T_{PH}}$.

5 The subsequential boundary

In this section we show that while the regular boundary is not sufficient to separate the pathological Sour Grapes pattern from attested harmony patterns, the subsequential boundary *is* sufficient.

Padgett (1995) defines *Sour Grapes Harmony* as "Either all features must spread, or none will..." For concreteness, consider a progressive Sour Grapes pattern. The form $+ - - -$ would be mapped to $+ + + +$ as in progressive harmony, but the form $+ - - \boxminus$ is mapped to $+ - - \boxminus$ because the opaque vowel will not become $+F$, and so the spreading process grumpily chooses not to spread at all. Therefore, a progressive Sour Grapes Harmony pattern is defined as any length-preserving function which at a minimum includes the following mappings for all $n \in \mathbb{N}$:

$$SG(+-^n) = ++^n \wedge SG(+-^n\boxminus) = +-^n\boxminus \quad (8)$$

There is a finite state transducer which describes this fragment of the SG function, shown in Figure 3. As a total function, for SG to be regular, it is important that the image of the complement of $\cup_{n \in \mathbb{N}}\{+-^n\} \cup_{n \in \mathbb{N}} \{+ -^n \boxminus\}$ under SG also be regular. Pictorially, this would mean that the fragment shown in Figure 3 is a *subgraph* of the full SG pattern. Crucially, however, there can also be no transition from state 2 bearing the input symbol $\boxminus$ that can lead (even eventually) to a final state.

We now prove the main theorem of this paper.

Theorem 5 *SG is neither left nor right subsequential.*

Proof We show that, for all distinct $n, m \in \mathbb{N}$, the tails of $+-^n$ is not the same as the tails of $+-^m$. This immediately implies that the canonical left subsequential transducer would have infinitely many states, and hence that any SG pattern meeting Equation 8 is not left subsequential.

---

[6]It is true that T_{PH} does not model progressive harmony patterns where transparent vowels can trigger harmony. It is not difficult to modify T_{PH} to accommodate this without sacrificing subsequentiality.
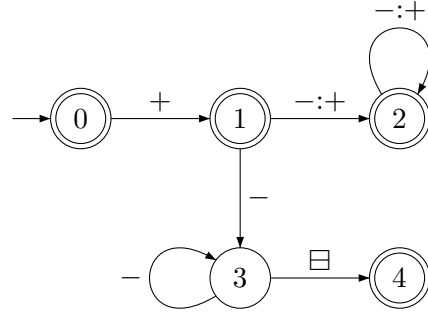

Figure 3: A non-deterministic transducer which recognizes a fragment of SG harmony.

To illustrate, consider $x = +-$. Since $SG(+ - X^*)$ includes elements $+++$ and $+-\boxminus$ (mapped from e.g. $+ - -$ and $+ - \boxminus$, resp.), it follows that $lcp(SG(+ - X^*)) = +$. Therefore, $(-, ++) \in TL_{SG}(+-)$. Observe that $(-, +^n) \notin TL_{SG}(+-)$ for all $n \neq 2$ since SG is length-preserving.

More generally it is the case that $(-, +^{n+1}) \in TL_{SG}(+-^n)$ and $(-, +^m) \notin TL_{SG}(+-^n)$ for all $m \neq n + 1$. Therefore there are infinitely many distinct sets of tails for functions conforming to (8), and thus no SG pattern is subsequential.

A similar argument (omitted) establishes that any SG pattern is not right subsequential. $\square$

Consequently, the subsequential boundary separates SG and MR from PH and RH.

6 The weakly deterministic boundary

As mentioned earlier, the dominant/recessive and stem-control theories of vowel harmony reject the directionality generalizations inherent in the PH and RH mappings. If these theories are correct, then it is important to see what boundary (if any) separates MR and SG from vowel harmony patterns described with dominant/recessive and stem-control based generalizations. We show that the mappings these theories posit are, like SG, not subsequential. However, we believe there is an interesting complexity difference between SG on the one hand and DR and SC on the other. In particular, we show that DR and SC are weakly deterministic. We conjecture that SG is not weakly deterministic and provide the intuition behind this conjecture, though its proof currently escapes us.

A dominant/recessive analysis of vowel harmony says if the word contains the dominant value of the harmonizing feature, then other vowels in the word take on the dominant value for this feature. For example, if the $[+]$ value for harmonizing feature F is the dominant one and an under-

lying representation contains a vowel specified as +F, then all other non-neutral vowels in the word will be realized as +F as well.

Therefore, we can define a function as dominant/recessive as any length-preserving function which includes the following mappings:

$$\forall w \in \{+,-\}^*,$$
$$DR(w) = \begin{cases} +^{|w|} & \text{if } (\exists 0 \leq i \leq |w|)[w_i = +] \\ -^{|w|} & \text{otherwise} \end{cases}$$
$$(9)$$

The next two theorems establish that DR harmony is properly weakly deterministic.

Theorem 6 *DR is neither left nor right subsequential.*

Proof The proof is similar to the one for SG. We show that, for all distinct $n, m \in \mathbb{N}$, the tails of $-^n$ is not the same as the tails of $-^m$.

Consider first $x = --$. To find its tails we must know $lcp(DR(--X^*))$. Since $DR(--X^*)$ includes elements $---$ and $+++$ (mapped from e.g. $---$ and $--+$, resp.), it follows that $lcp(DR(+-X^*)) = \lambda$. Therefore, $(-, ---) \in TL_{DR}(--)$. Observe that for all $n \neq 2$, it is the case that $(-, +^n) \notin TL_{DR}(--)$ since DR is length-preserving.

Next consider $x = ---$. To find its tails we must know $lcp(DR(---X^*))$. $DR(---X^*)$ includes elements $----$ and $++++$ (mapped from e.g. $----$ and $---+$, resp.), and so again the longest common prefix is λ. Therefore, $(-, ++++)$ belongs to $TL_{DR}(---)$ and $(-, +^n) \notin TL_{DR}(+-)$ for all $n \neq 3$ since DR preserves string lengths.

More generally for all distinct $n, m \in \mathbb{N}$ it is the case that $(-, +^{n+1}) \in TL_{DR}(-^n)$ and $(-, +^m) \notin TL_{DR}(+-^n)$ for all $m \neq n + 1$. Therefore there are infinitely many distinct sets of tails for functions conforming to (9), and thus no DR pattern is subsequential.

A similar argument (omitted) establishes that any DR pattern is not right subsequential. $\square$

On the other hand, DR is weakly deterministic. We establish this for the case when the alphabet contains only $\{+,-\}$. In fact, DR is simply the composition of progressive harmony and regressive harmony where only the dominant feature value spreads. Figure 4 shows a subsequential transducer T_{PHP} describing a progressive harmony function where only $[+]$ spreads. Observe
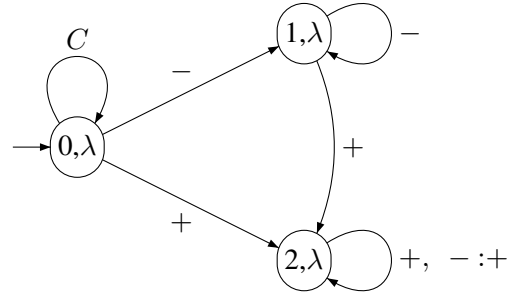

Figure 4: The subsequential transducer T_{PHP} which recognizes iterative, progressive harmony where only the $+$ value spreads.

that the transducer in Figure 4 is nearly identical to T_{PH} in Figure 2 without the opaque vowels and the C symbol. The important difference is that in Figure 2 there is a transition from state 1 to itself which reads a $[+]$ input and outputs a $[-]$, but in Figure 4, the transducer in state 1, upon reading input $[+]$ transitions to state 2 and writes out $[+]$.

Let PHP denote the function T_{PHP} computes.

Theorem 7 *DR is weakly deterministic.*

Proof We show that for all $w \in \{+,-\}^*$, $DR(w) = \overleftarrow{T_{PHP}} \circ \overrightarrow{T_{PHP}}(w)$.

Case 1. There exists $0 \leq i \leq |w|$ such that $w_i = +$. Then $DR(w) = +^{|w|}$. It follows from the definition of PHP that for all $j \geq i$, $w_j = +$. Letting $u = PHP(w)^r$, it follows that $u_1 = +$. Thus $PHP(u) = +^{|w|}$. The reverse of $+^{|w|}$ is clearly itself. Therefore, $PHP^r \circ PHP(w) = +^{|w|}$.

Case 2. Case 1 does not hold. Then $DR(w) = -^{|w|}$. By definition, $PHP(w) = -^{|w|}$. Clearly then $PHP^r \circ PHP(w) = -^{|w|}$.

Since $DR = \overleftarrow{T_{PHP}} \circ \overrightarrow{T_{PHP}}$, since $\overrightarrow{T_{PHP}}$ and $\overleftarrow{T_{PHP}}$ are alphabet-preserving, and since $\overrightarrow{T_{PHP}}$ is not length-increasing, the theorem is proved. $\square$

While, the proof of theorem 7 is limited to words in $\{+,-\}^*$, we believe the extension to words in $\{+, -, C, \boxminus, \boxplus\}^*$ is only challenging technically, and not conceptually. For example, the definition of DR harmony above in Equation 9 should be more articulated so that, for instance, $DR(--\boxminus- -+--\boxminus--) = --\boxminus+++++\boxminus--$.

Stem-controlled analyses are similar to dominant/recessive theories. Unlike dominant/recessive theories, however, the vowels which trigger harmony are the ones which belong to the morphological stem. Table 4 illustrates with stem boundaries indicated with #. What

59

	w	SC(w)
a.	$+\#-\#-$	$---$
b.	$-\#+\#+$	$+++$
c.	$-\#-\#-$	$---$
d.	$-\#+\#-$	$+++$

Table 4: Example mappings of underlying forms (w) with three vowels given stem control theories of vowel harmony. In each example, the middle vowel is the only vowel in the stem.

happens when there is more than one stem vowel? The stem precedence generalization (Baković, 2003) states that "an alternating affix vowel always agrees with the adjacent vowel in the stem to which the affix is attached." Therefore stem vowels themselves are not targets of the harmony process. Consequently, underlying $/--\#+-+-\#++/$ would surface as $[++\#+-+-\#--]$.

Since every underlying form is assumed to contain a stem, the domain of SC harmony is $X^*\#X^+\#X^*$. Then $\forall w\#u\#v \in X^*\#X^+\#X^*$, it is the case that

$$SC(w) =$$

$$\begin{cases} +^{|w|}\#u\#+^{|v|} & \text{if } u_1 = + \ \wedge \ u_{|u|} = + \\ +^{|w|}\#u\#-^{|v|} & \text{if } u_1 = + \ \wedge \ u_{|u|} = - \\ -^{|w|}\#u\#+^{|v|} & \text{if } u_1 = - \ \wedge \ u_{|u|} = + \\ -^{|w|}\#u\#-^{|v|} & \text{if } u_1 = - \ \wedge \ u_{|u|} = - \end{cases}$$

$$(10)$$

The analysis of SC is the same as DR; and so the proofs are omitted.

Theorem 8 *SC is neither left nor right subsequential.*

Theorem 9 *SC is weakly deterministic.*

We believe there is a difference between DR and SG: we conjecture that SG is not weakly deterministic. To get an intuition why, consider the two subsequential transducers A and B in Figure 5. Limiting our attention to the domain $\{+-^n\}\cup\{+-^n\boxminus\}\cup\{\lambda\}$, the composition $\overleftarrow{B}\circ\overrightarrow{A}$ equals SG. This is possible because these functions make use of an additional symbol $[\overset{?}{-}]$, indicating a minus value whose left context matches the environment to become $[+]$.

Table 5 illustrates the role the additional symbol plays in the derivation. We are doubtful that it is possible to decompose SG into a left and right deterministic function where the left function is pro-
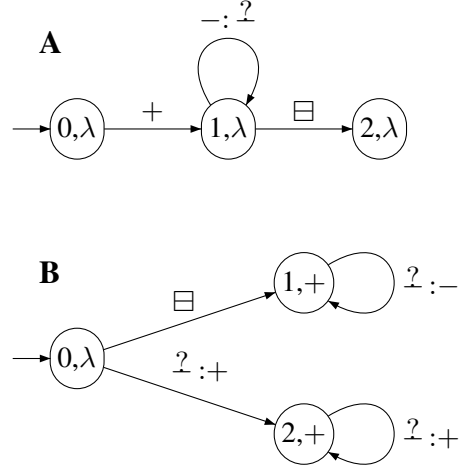

Figure 5: Two subsequential transducers such that $\overleftarrow{B}\circ\overrightarrow{A} = SG$. The symbol $[\overset{?}{-}]$ indicates a $[-]$ which would undergo harmony provided no opaque vowel occurs downstream.

w	$+---\boxminus$	$+---$
$\overrightarrow{A}(w)$	$+\overset{?}{-}\overset{?}{-}\overset{?}{-}\boxminus$	$+\overset{?}{-}\overset{?}{-}\overset{?}{-}$
$\overleftarrow{B}\circ\overrightarrow{A}(w)$	$+---\boxminus$	$++++$

Table 5: Illustrations of the role of $[\overset{?}{-}]$ in the deterministic decomposition of SG=$\overleftarrow{B}\circ\overrightarrow{A}$.

hibited from marking up its output in any way (either with extra symbols or with a length-increasing coding trick).

7 Conclusion

The first suggestion that phonological processes have a tighter computational bound than "being regular" may come from (Mohri, 1997), buried on page 279. He writes without elaboration or citation "Most phonological and morphological rules correspond to p-subsequential relations." This study suggests that Mohri's assessment is largely correct, though the complete picture is more complicated than Mohri's offhand comment indicates.

The more complicated picture with respect to vowel harmony is expressed in Figure 6, which summarizes this paper's contributions. Traditional directional theories of vowel harmony express simpler generalizations than dominant/recessive/stem-control theories. It is our opinion that future work will likely show that even the weakly deterministic boundary surely separates the pathological patterns from the attested ones.
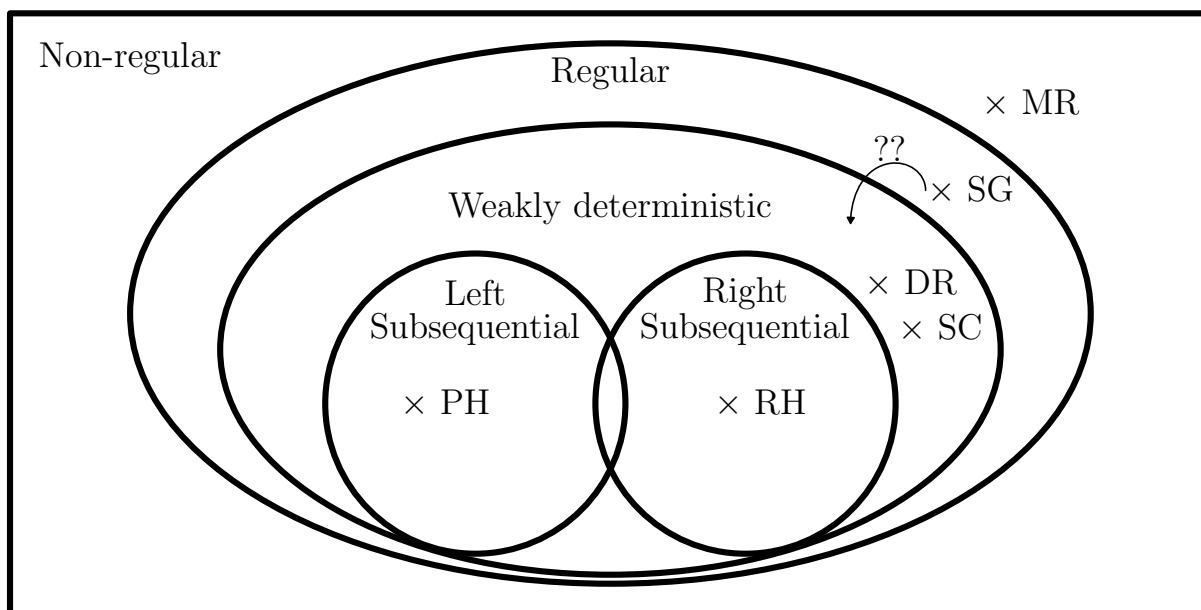
Figure 6: Hierarchies of transductions with the results of this paper shown. PH=progressive harmony, RH=regressive harmony, DR=dominant/recessive harmony, SC=stem control harmony, SG=sour grapes harmony, and MR=majority rules harmony.

Although the harmony patterns in this paper are all describable with same-length relations, we deliberately chose not to focus on the special properties same-length relations engender. This is largely because there are phonological processes such as epenthesis and deletion which are not same-length, and we would like our conclusions to hold for all phonological patterns. Nonetheless, future work which explores the same-lengthness aspect may lead to some interesting insights. One reviewer of this paper conjectured, for example, that if the same-length relations were coded as languages that they would then be k-reversible (Angluin, 1982).

With respect to learnability, total subsequential functions are identifiable in the limit from positive data (Oncina et al., 1993), though this algorithm appears to require data points unavailable in natural language corpora (Gildea and Jurafsky, 1996). Investigating subclasses of subsequential functions which cover attested phonological patterns may thus not only better characterize possible phonologies, but may also provide insights for learning (Chandlee and Koirala, 2013).

Finally, we believe Elgot and Mezei's theorem can shed new light on the old problem of abstractness in phonology (Hyman, 1970), and suspect a hierarchy of complexity depending on how much markup (either new symbols or with a length-

increasing function) needs to be introduced in the intermediate alphabet to order to decompose a regular function into left and right subsequential ones. Computational work whose results should be more carefully investigated with this in mind include Kempe (2000) and Crespi Reghizzi and San Pietro (2012).

Acknowledgments

References

Dana Angluin. 1982. Inference of reversible languages. *Journal for the Association of Computing Machinery*, 29(3):741–765.

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

Eric Baković. 2000. *Harmony, Dominance and Control*. Ph.D. thesis, Rutgers University.

Eric Baković. 2003. Vowel harmony and stem identity. In *San Diego Linguistic Papers*, pages 1–42. Department of Linguistics, UCSD, UC San Diego. http://escholarship.org/uc/item/7zw206pt.

Kenneth Beesley and Lauri Kartunnen. 2003. *Finite State Morphology*. CSLI Publications.

Michael Benedikt, Leonid Libkin, Thomas Schwentick, and Luc Segoufin. 2001. A model-theoretic approach to regular string relations. *Logic in Computer Science, Symposium on*, 0:0431.

Jean Berstel. 1979. *Transductions and Context-Free languages*. Teubner-Verlag.

Jane Chandlee and Jeffrey Heinz. 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*, pages 42–51, Montreal, Canada, June. Association for Computational Linguistics.

Jane Chandlee and Cesar Koirala. 2013. Learning local phonological rules. Penn Linguistics Colloquium, March.

Jane Chandlee, Angeliki Athanasopoulou, and Jeffrey Heinz. 2012. Evidence for classifying metathesis patterns as subsequential. In *The Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 303–309. Cascillida Press.

Noam Chomsky and Morris Halle. 1968. *The Sound Pattern of English*. Harper & Row, New York.

Stefano Crespi Reghizzi and Pierluigi San Pietro. 2012. From regular to strictly locally testable languages. *International Journal of Foundations of Computer Science*, 23(08):1711–1727.

C. C. Elgot and J. E. Mezei. 1965. On relations defined by generalized finite automata. *IBM Journal of Research and Development*, 9(1):47–68.

Sara Finley. 2008. *The formal and cognitive restrictions on vowel harmony*. Ph.D. thesis, Johns Hopkins University, Baltimore, MD.

Brian Gainor, Regine Lai, and Jeffrey Heinz. 2012. Computational characterizations of vowel harmony patterns and pathologies. In *The Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 63–71.

Daniel Gildea and Daniel Jurafsky. 1996. Learning bias and phonological-rule induction. *Computational Linguistics*, 24(4).

Bruce Hayes. 2009. *Introductory Phonology*. Wiley-Blackwell.

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

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

Jeffrey Heinz, Chetan Rawal, and Herbert G. Tanner. 2011. Tier-based strictly local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 58–64, Portland, Oregon, USA, June. Association for Computational Linguistics.

Larry Hyman. 1970. How concrete is phonology? *Language*, 46(1):58–76.

Larry Hyman. 1975. *Phonology: Theory and Analysis*. Holt, Rinehart and Winston.

C. Douglas Johnson. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.

Ronald Kaplan and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics*, 20(3):331–378.

André Kempe. 2000. Reduction of intermediate alphabets in finite-state transducer cascades. *CoRR*, cs.CL/0010030.

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

Kimmo Koskenniemi. 1983. Two-level morphology. Publication no. 11, Department of General Linguistics. Helsinki: University of Helsinki.

Martin Krämer. 2003. *Vowel Harmony and Correspondence Theory*. Berlin: Mouton de Gruyter.

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

Linda Lombardi. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory*, 17:267–302.

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

José Oncina, Pedro García, and Enrique Vidal. 1993. Learning subsequential transducers for pattern recognition tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:448–458, May.

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

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

Jason Riggle. 2004. *Generation, Recognition, and Learning in Finite State Optimality Theory*. Ph.D. thesis, University of California, Los Angeles.

Emmanuel Roche and Yves Schabes. *Finite-State Language Processing*. MIT Press.

James Rogers, Jeffrey Heinz, Margaret Fero, Jeremy Hurst, Dakotah Lambert, and Sean Wibel. to appear. Cognitive and sub-regular complexity. In *Proceedings of the 17th Conference on Formal Grammar*.

Jaques Sakarovitch. 2009. *Elements of Automata Theory*. Cambridge University Press. Translated by Reuben Thomas from the 2003 edition published by Vuibert, Paris.

Colin Wilson. 2003. Analyzing unbounded spreading with constraints: marks, targets, and derivations. Unpublished manuscript, UCLA.

Colin Wilson. 2004. Experimental investigation of phonological naturalness. In *Proceedings of WCCFL 22*, pages 534–546.

Anssi Yli-Jyrä and Kimmo Koskenniemi. 2006. Compiling generalized two-level rules and grammars. In *FinTAL*, volume 4139 of *Lecture Notes in Artificial Intelligence*, pages 174–185. Springer-Verlag, Berlin.