

Implications of Harmonic Serialism for Lexical Tone Association¹

John J. McCarthy, Kevin Mullin, and Brian Smith

University of Massachusetts Amherst

1. Introduction

One of the properties of the classic version of Optimality Theory in Prince & Smolensky (1993/2004) is *parallelism*: output candidates may show the effects of several different phonological changes at once. Harmonic Serialism (HS), a version of OT that Prince & Smolensky also briefly consider, is *serial* or *derivational* rather than parallel: output candidates show the effect of only one phonological change at a time, but the winning candidate is run through the grammar again so it can accumulate additional changes.

Although Prince & Smolensky put HS aside and it received little subsequent attention, the case for it was reopened in McCarthy (2000, 2002: 159-163, 2007). In these and other works (see McCarthy (to appear) for a summary), two main kinds of evidence for HS have been identified. One is that it permits generalizations to be stated on its intermediate representations, neither underlying nor surface. For example, the cross-linguistically common process of unstressed vowel syncope cannot be analyzed satisfactorily without access to a level of representation at which stress has been assigned but syncope has not yet occurred (McCarthy 2008). The other kind of evidence is typological: the same constraints can predict different language typologies under parallel OT (P-OT) and HS, and in cases examined so far these differences favor HS. For example, HS correctly predicts certain locality effects in stress assignment (Pruitt 2008).

Of course, an honest exploration of HS's consequences must include a search for potential problems. Most of the existing arguments for parallelism in the OT literature have been addressed (in, e.g., McCarthy 2008, McCarthy et al. 2010, Pater to appear). But new arguments have also emerged. This paper lays out one such argument, based on the phonology of tone in Kikuyu.

Section 2 describes the basic tonal phonology of Kikuyu, summarizing the analysis in Clements & Ford (1979). Section 3 then goes on to explain how this system can be analyzed in HS using standard tone constraints. A key feature of this analysis is the assumption that candidates can change by adding or removing at most one association line at a time. The argument for parallelism is presented in section 4. It consists of a demonstration that lexically linked tones disrupt the analysis in HS but not P-OT. The problem for HS is that removing an unwanted lexical tone association is sometimes impossible in Kikuyu-type systems. This problem does not arise in P-OT, where removing unwanted tone associations and inserting better ones take place simultaneously.

What conclusion can we draw from this result? One possibility is that HS is simply wrong and the parallel theory is right. This is not a very attractive option, however, because it disregards the large and growing body of evidence for HS over parallelism. The alternative, which we explore in section 5, is to return to a view that was prevalent in the early days of autosegmental phonology: tones are never associated in underlying representation.

¹ We are grateful for comments received from the McCarthy-Pater grant group at UMass Amherst. This research was supported by grant BCS-0813829 from the National Science Foundation to the University of Massachusetts Amherst.

Finally, section 6 concludes with some discussion of the convergence between our results and those in McCarthy & Pruitt (to appear), where it is shown that lexical foot structure is also incompatible with HS.

2. Tone association in Kikuyu

In Kikuyu, tone association is fully predictable (Clements & Ford 1979, Clements 1984): the first tone is associated with the first two syllables, and each subsequent tone is associated with the syllable immediately following the morpheme that supplied it.

(1) Example

a. Lexical representations

Segments	Tone	
to	L	‘we’
ma	H	‘them’
rɔr	L	‘look at’
ir-ε	H	current past tense

b. Surface form

L	H	L	H	
\				
tomarɔririε				‘we looked at them’

Kikuyu follows this pattern with remarkable regularity; indeed, Clements (1984: 290, 298) identifies just two non-conforming situations.

In Clements & Ford’s (1979) analysis, surface structures like (2) are obtained from underlying representations in which tones are lexically listed with morphemes but not autosegmentally associated with them. The underlying forms of the morphemes in (1) are therefore: /to, L/; /ma, H/; /rɔr, L/; and /ir-ε, H/. In the first step of the derivation, a language-particular rule associates the initial L with the second syllable:

(2) After initial tone association rule

L	H	L	H	
\				
tomarɔririε				

From this point forward, universal tone association conventions take over. They cause the remaining unlinked tones to be linked one-to-one, from left to right, to the toneless syllables:

(3) After one-to-one left-to-right association

L	H	L	H	
\				
tomarɔririε				

The universal tone association conventions also cause the initial L to spread to the toneless initial syllable, yielding the surface form in (1)b.

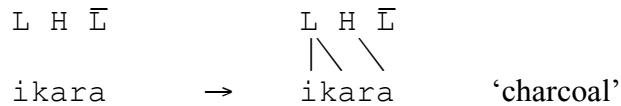
It sometimes happens that there are as many tones as syllables. Since the first tone takes up two syllables, a tone is left over after one-to-one left-to-right association. One possible disposition of the left-over tone is that it becomes part of a contour tone. Kikuyu allows rising tones on the final syllable, as in (4):

(4) Final rising tone (Clements & Ford 1979: 191)



But if the tonal contour is forbidden, then the left-over tone remains floating. For example, Kikuyu does not allow contour tones falling from H to \bar{L} (=extra low). As a result, the \bar{L} tone remains floating in (5), where its presence can be inferred from its role in the downstep system (Clements & Ford 1979: 203ff.):

(5) Final floating tone



The assumption that tones are lexically unassociated with their sponsoring morphemes is essential to Clements & Ford's analysis. It is precisely because the non-initial tones are unassociated in (2) that the universal tone association conventions produce (3). If, on the contrary, tones were associated in underlying representation, then the immediate output of the initial tone association rule would look like this:

(6) After initial tone association rule, underlying linked tones



The universal tone association conventions are no guide through the complex chain of operations that would be needed to get from (6) to the surface form in (1)b. Thus, the assumption that tones are unassociated in underlying representation is doing real work in Clements and Ford's analysis.²

3. Kikuyu tone association in Harmonic Serialism

In the view of Clements and Ford (1979), one-to-one left-to-right association of tones with syllables is the default pattern cross-linguistically. To analyze this pattern in HS (or OT generally), constraints of four types are required:

- (i) Constraints against floating tones and toneless syllables, which approximate the effects of Goldsmith's (1976a) Well-formedness Condition for autosegmental phonology. These constraints motivate tone-syllable association.
 - (7) NO-FLOAT (NO-FL)

Assign one violation mark for every tone that is not associated with a syllable.
 - (8) HAVE-TONE (HAVE-T)

Assign one violation mark for every syllable that is not associated with a tone.
- (ii) Constraints against skipping tones or syllables in the process of association. They ensure that association iterates directionally.

² We emphasize this point because non-association of tones in underlying representation is often posited for less weighty reasons, such as considerations of minimal redundancy.

- (9) NO-SKIP(tone) (NO-SK(t))
Assign one violation mark for every unlinked tone that is preceded and followed (at any distance) by linked tones.
- (10) NO-SKIP(syllable) (NO-SK(s))
Assign one violation mark for every unlinked syllable that is preceded and followed (at any distance) by linked syllables.
- (iii) A constraint requiring the initial tone to link to the initial syllable. This accounts for the default left-to-right direction.
- (11) LINK-INITIAL (LNK-INIT)
Assign a violation mark if the initial tone is not linked to the initial syllable.
- (iv) The faithfulness constraint IDENT(tone), which is violated once for each change in an association line. (Because deletion and insertion of tones are not in general at issue here, we assume that MAX(tone) and DEP(tone) are undominated.)

None of these constraints is really new; see Myers (1997) or Yip (2002), among others, for similar proposals.

In HS far more than in P-OT, the details of the candidate generator GEN are important, because GEN determines exactly how much a candidate can differ from its input. The natural assumption for autosegmental phonology is that GEN can insert or delete exactly one association line at a time. We also make the standard assumption that GEN cannot produce structures with crossing association lines.

We will illustrate this system with an HS analysis of a hypothetical language with left-to-right tone association. (This language is essentially Kikuyu without tone shift.) The underlying representation consists of segmental material for four syllables and an equal number of unassociated tones. The derivation in (12) is typical of what the analysis has to do:

(12) Derivation for left-to-right language

Underlying	Step 1	Step 2	Step 3	Step 4	Step 5 (convergence)
L H L H patakasa	L H L H patakasa	L H L H patakasa	L H L H patakasa	L H L H patakasa	L H L H patakasa

The candidates available at the first step of the derivation, some of which are shown in tableau (13), consist of all the ways of linking one tone to one syllable, plus the faithful candidate. The faithful candidate (13)b loses because the faithfulness constraint IDENT(tone) is ranked below NO-FLOAT and HAVE-TONE. The others (13)c,d lose because LINK-INITIAL favors candidate (13)a, where the first tone and first syllable are associated with one another.

(13) Step 1 of one-to-one left-to-right association

	L H L H	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	ID(t)
	patakasa						
a. →	<pre> L H L H patakasa </pre>	3	3				1
b.	<pre> L H L H patakasa </pre>	4 W	4 W			1 W	L
c.	<pre> L H L H \ patakasa </pre>	3	3			1 W	1
d.	<pre> L H L H patakasa </pre>	3	3			1 W	1

At step 2 (tableau (14)), the no-skipping constraints are important. NO-FLOAT and HAVE-TONE continue to compel tone association, and the no-skipping constraints determine that the next tone and syllable associated must be adjacent to the last tone and syllable associated.

(14) Step 2 of one-to-one left-to-right association³

	L H L H	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	ID(t)
	patakasa						
a. →	<pre> L H L H patakasa </pre>	2	2				1
b.	<pre> L H L H patakasa </pre>	3 W	3 W				L
c.	<pre> L H L H \ patakasa </pre>	2	2		1 W		1
d.	<pre> L H L H / patakasa </pre>	2	2	1 W			1

The derivation continues like this until it converges at step 5.

The derivation in (13)–(14) involves an example with equal numbers of tones and syllables. When the number of syllables is larger than the number of tones, as in (15), then the last tone will spread to any toneless syllables because of HAVE-TONE. When the number of tones is larger, as in (16), NO-FLOAT will force creation of a contour tone, unless some higher ranking constraint prevents it.

³ The violations of IDENT(tone) at step 2 are based on the input to that step rather than the underlying representation.

(15) Step 4 from /patakasa, LHL/

	L H L	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	ID(t)
	patakasa						
a. →	patakasa						1
b.	patakasa		1 W				L

(16) Step 4 from /pataka, LHLH/

	L H L H	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	ID(t)
	pataka						
a. →	pataka						1
b.	pataka	1 W					L

Kikuyu follows this pattern in all respects except one: it requires the first tone to be linked to two syllables. Although Clements (1984: 330--331) has an interesting proposal about why this should be so, in the interest of simplicity we will opt here for an ad hoc constraint:

(17) INITIAL-PLATEAU (INIT-PLAT) (ad hoc constraint)

Assign a violation mark if the initial tone is linked to fewer than two syllables.

We now have the resources we need to analyze Kikuyu. Since /tomarɔrɪɛ, LHLH/ contains five syllables and four tones, there are 20 ways of adding a single association line to it at step 1. That is the candidate set. Only one of these candidates satisfies LINK-INITIAL, which is the only constraint (other than faithfulness) that distinguishes among them, as tableau (18) shows.

(18) Step 1 from /tomarɔrɪɛ, LHLH/

	L H L H	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	NO- FL	ID(t)
	tomarɔrɪɛ							
a. →	tomarɔrɪɛ	4				1	3	1
b.	tomarɔrɪɛ	5 W			1 W	1	4 W	L
c.	tomarɔrɪɛ	4			1 W	1	3	1
d.	tomarɔrɪɛ	4			1 W	1	3	1

At the second step, the available options include spreading the first tone (19)a, doing nothing (19)b, one-to-one linking (19)c, creation of a contour tone (19)d, and delinking (19)e.

Spreading of the first tone is required by INITIAL-PLATEAU. For it to take precedence over one-to-one linking, INITIAL-PLATEAU has to dominate NO-FLOAT:

(19) Step 2 from /tomarɔrɪɛ, LHLH/

	L H L H tomarɔrɪɛ	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	NO- FL	ID(t)
a. →	L H L H \ tomarɔrɪɛ	3					3	1
b.	L H L H tomarɔrɪɛ	4 W				1 W	3	L
c.	L H L H tomarɔrɪɛ	3				1 W	2 L	1
d.	L H L H / tomarɔrɪɛ	4 W				1 W	2 L	1
e.	L H L H tomarɔrɪɛ	4 W			1 W	1 W	4 W	1

Observe that delinking of the previously associated tone, as in (19)e, is harmonically bounded by the winner (19)a. This will become important when we consider the effects of lexical associations.

One-to-one left-to-right association prevails at the next step of the derivation, where it is favored by the no-skipping constraints. In this respect, the Kikuyu step-3 tableau (20) is identical with the step-2 tableau of the hypothetical example in (14).

(20) Step 3 from /tomarɔrɪɛ, LHLH/

	L H L H \ tomarɔrɪɛ	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	NO- FL	ID(t)
a. →	L H L H \ tomarɔrɪɛ	2					2	1
b.	L H L H \ tomarɔrɪɛ	3 W					3	L
c.	L H L H \ tomarɔrɪɛ	2	1 W	1 W			2	1
d.	L H L H \ tomarɔrɪɛ	2		1 W			2	1
e.	L H L H \ tomarɔrɪɛ	2	1 W				2	1

The pattern of one-to-one left-to-right association continues at step 4:

(21) Step 4 from /tomarɔrɪɛ, LHLH/

	L H L H	HAVE-T	NO-SK(t)	NO-SK(s)	LNK-INIT	INIT-PLAT	NO-FL	ID(t)
	tomarɔrɪɛ							
a. →	tomarɔrɪɛ	1					1	1
b.	tomarɔrɪɛ	2 W					2 W	L
c.	tomarɔrɪɛ	1		1 W			1	1
d.	tomarɔrɪɛ	1	1 W				1	1

At step 5, the last unlinked tone associates with the last unlinked syllable:

(22) Step 5 from /tomarɔrɪɛ, LHLH/

	L H L H	HAVE-T	NO-SK(t)	NO-SK(s)	LNK-INIT	INIT-PLAT	NO-FL	ID(t)
	tomarɔrɪɛ							
a. →	tomarɔrɪɛ							1
b.	tomarɔrɪɛ	1 W					1 W	L

Finally, at step 6, the derivation converges on the correct surface form:

(23) Step 6 from /tomarɔrɪɛ, LHLH/ — Convergence

	L H L H	HAVE-T	NO-SK(t)	NO-SK(s)	LNK-INIT	INIT-PLAT	NO-FL	ID(t)
	tomarɔrɪɛ							
a. →	tomarɔrɪɛ							1
b.	tomarɔrɪɛ							1 W

The analysis just presented is strictly structure-building. The original inputs to the system have no tone-syllable associations, and the ultimate outputs have no toneless syllables and floating tones only when contours are impossible. Structure is built in a strictly monotonic fashion: at no point was it optimal to remove an association line. Kikuyu's monotonic structure-building derivations are, moreover, monotonically harmonically improving with respect to HAVE-TONE and NO-FLOAT. Each added association line improves performance on one or more of these constraints. Removing association lines only degrades performance, making the representation less harmonic rather than more.

4. Effect of lexical tone linking

This analysis, which was worked out under the assumption that tones are unassociated in the lexicon, fails when confronted with lexical associations. To show this, we will examine the threshold case, where every tone is lexically linked, and show that the HS analysis is unable to bring it into conformity with the with the observed regularities of tone association in Kikuyu.

An adequate grammar of Kikuyu needs to map the underlying representation in (24)a to the surface representation in (24)b:

(24) Effect of lexical linking

a. Underlying form

L H L H
| | | |
patakasa

b. Desired surface form

L H LH
| | |
patakasa

This requirement follows because the grammar has to capture some basic generalizations about tone in this language: the underlying tones of non-initial, non-final morphemes appear one syllable to their right in the surface form; the first two syllables have a level tone; and contour tones are confined to the final syllable. Section 3 presented a grammar that captures these generalizations when the underlying form has only unlinked tones. As we will now show, the same grammar is unable to accomplish the mapping in (24), which means that it is sometimes unable to capture these generalizations when the underlying representation has linked tones.

At step 1, INITIAL-PLATEAU prevails, and the first tone spreads to the second syllable, creating a LH contour tone on it, as in (25):

(25) Step 1 from /paL taH kaL saH/

L H L H patakasa	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	NO- FL	ID(t)
a. → L H L H pa ta ka sa							1
b. L H L H patakasa					1 W		L
c. L H L H patakasa	1 W	1 W	1 W		1 W	1 W	1

At step 2, the grammar converges on the output of Step 1:

(26) Convergence at step 2 from /paL taH kaL saH/

L H L H pa ta ka sa	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	NO- FL	ID(t)
a. → L H L H pa ta ka sa							
b. L H L H pa ta ka sa		1 W				1 W	1 W
c. L H L H pa ta ka sa							1 W

This is not a welcome result. Starting from an underlying representation where tones and tone-bearers are linked one-to-one, the grammar produces a result with a contour tone on the second syllables. This is at odds with the facts of Kikuyu, which allows contour tones only on the final syllable.

The source of this problem is evident from (26). To get from (26)a to a final result that has a contour tone on the final syllable, the H linked to the second syllable needs to shift to the third syllable, forcing the L on the third syllable to shift to the fourth and final syllable. (In longer words, obviously, more intermediate steps will be required). The only way to accomplish within our assumptions about GEN is for one of the losers in (26) to win. But both are harmonically bounded by the winner, which means that no ranking of these constraints can make winners of them.

The harmonic bounding in (26) can be broken by introducing another constraint that favors one of the losers over the unwanted winner. One plausible option is NO-MEDIAL-CONTOUR (NO-MC), which favors (26)b over (26)a. It has to be ranked below INITIAL-PLATEAU, since creating the plateau produces a medial contour tone. But if it dominates NO-SKIP(tone) and NO-FLOAT, it can force simplification of the contour tone created at step 1 by delinking the H part of the contour:

(27) Step 2 from /paL taH kaL saH/ with NO-MEDIAL-CONTOUR

		NO-SK(s)	INIT-PLAT	LNK-INIT	HAVE-T	NO-MC	NO-SK(t)	NO-FL	ID(t)
a. →							1	1	1
b.						1 W	L	L	L

At the next step, the floating H needs to reassociate with the third syllable, but that is impossible under this ranking. The problem (see tableau (28)) is that reassociating the floating H violates NO-MEDIAL-CONTOUR, while the floating H itself violates only NO-SKIP(tone) and NO-FLOAT. As we just saw in (27), these constraints have to be ranked below NO-MEDIAL-CONTOUR to get step 2 to work right.

(28) Convergence at step 3 on output of (27)

		NO-SK(s)	INIT-PLAT	LNK-INIT	HAVE-T	NO-MC	NO-SK(t)	NO-FL	ID(t)
a. →							1	1	
b.						1 W	L	L	1 W

Clearly, introducing NO-MEDIAL-CONTOUR does not in general help to map underlying representations with linked tones onto the actual surface tone pattern of Kikuyu. With or without this constraint, the HS analysis only works reliably when underlying representations are limited to unassociated tones.

Because analyses in HS depend on assumptions about GEN as well as CON, we need to check whether tweaking our assumptions about GEN will lead to a solution. A natural idea is to provide GEN with a “flop” operation that is capable of reassociating a tone from one syllable to another in a single step. If such an operation were available at step 2, it would add another

relevant candidate to those listed in (27). This candidate, which is shown in (29)c, is no improvement — in fact, it is harmonically bounded by (29)b, which has no changes at all.

(29) Step 2 from /paL taH kaL saH/ with flop

	NO-SK(s)	INIT-PLAT	LNK-INIT	HAVE-T	NO-MC	NO-SK(t)	NO-FL	ID(t)
						1	1	1
					1 W	L	L	L
					1 W	L	L	2 W

With further expansion of the power of GEN — minimally, allowing it to affect multiple tones and association lines in a single step — it would in principle be possible to get around this problem, simply by offering the final output as a candidate at step 1. Needless to say, we would then have blurred the distinction between HS's and P-OT's GEN components into invisibility. The results attributable to HS, particularly the desirable results noted in section 1, depend on maintaining a sharp distinction between these two theories of GEN.⁴

Yet another possibility is that underlying representations with linked tones map not to actual surface forms but rather to the null output. Prince & Smolensky (1993/2004) propose that the null output, which we will indicate here with the symbol \ominus , is a member of every candidate set. It has no phonological form, so it satisfies every markedness constraint. It also has an undefined relationship with the input, so it satisfies every faithfulness constraint (Wolf & McCarthy 2009).

In parallel OT, optimality of \ominus means that the underlying form has no corresponding surface form. (Hence, \ominus is often invoked in discussions of paradigmatic gaps.) In HS, optimality of \ominus at any step of the derivation terminates the derivation at that step, again with no surface form. It is a fact about HS that \ominus can only be optimal at the very first step of the derivation. Suppose a non-null candidate O is optimal at Step 1. Obviously, O is more harmonic than \ominus . At Step 2, O is still in the candidate set, as the faithful candidate. Since O is more harmonic than \ominus , \ominus cannot be optimal at Step 2 either. By the logic of harmonic improvement, the optima at succeeding steps are more harmonic than their predecessors, so it can never be the case that \ominus is optimal at step $n > 1$.

The only constraint that \ominus violates is called MPARSE, and no other candidate violates this constraint. Let \mathbf{C} denote the set of constraints that dominate MPARSE in the grammar of some language. The constraints in \mathbf{C} are effectively inviolable in that language, because \ominus beats any candidates that violates them. If all other candidates from some input violate constraints in \mathbf{C} , then \ominus is optimal and there is no output from that input.

⁴ Another undesirable move is to replace NO-MEDIAL CONTOUR with a gradient constraint ALIGN-R(contour tone, word). This constraint would break the harmonic bounding in (29) by favoring (29)c over (29)b, because (29)c's contour tone is one syllable further to the right than (29)b's. The problem with this constraint is that it makes implausible typological predictions in other grammars. For example, it predicts the existence of a language where the penult can have a contour tone only if the ultima does, where the antepenult can have a contour tone only if the penult and ultima do, and so on. In general, licensing constraints like NO-MEDIAL CONTOUR never seem to be evaluated gradiently (Zoll 1996: 141).

For \odot to be useful in solving the ROTB problem in Kikuyu, it must at a minimum be non-optimal at step 1 when the underlying representation has unlinked tones but optimal at step 1 when the underlying representation has linked tones. Specifically, we need to add the candidate \odot to tableaux (18) and (25), ranking MPARSE in such a way that \odot is optimal in the latter but not the former. This turns out to be impossible, as examination of tableaux (30) and (31) reveals.

When the underlying representation has unlinked tones, as in (30), \odot must not win at Step 1, since this underlying representation leads eventually to a surface form. Because \odot obeys all constraints except MPARSE, MPARSE must dominate every constraint that the intended winner violates. Among these constraints is INITIAL-PLATEAU:

(30) MPARSE dominates INITIAL-PLATEAU with unlinked input

	L H L H tomarɔrɪrɛ	MPARSE	INIT- PLAT
a. →	L H L H tomarɔrɪrɛ		1
b.	\odot	1 W	L

When the underlying representation has linked tones, as in (31), then \odot should be the optimum. But for \odot to be optimal, the opposite ranking of these two constraints is required:

(31) INITIAL-PLATEAU dominates MPARSE with linked input

	L H L H patakasa	INIT- PLAT	MPARSE
a. →	\odot		1
b.	L H L H patakasa	1 W	L

When the underlying representation is unlinked, as in (30), no non-null candidate satisfies INITIAL-PLATEAU. To prevent \odot from winning, then, MPARSE has to dominate INITIAL-PLATEAU. When the underlying representation is fully linked, as in (31), then the faithful candidate violates INITIAL-PLATEAU *and no other constraint*. To ensure that \odot wins, then, INITIAL-PLATEAU has to dominate MPARSE. So \odot offers no solution to the problem of accommodating lexically linked tones in an HS analysis of Kikuyu.

In contrast, underlying tone associations present no difficulties for a P-OT analysis. As long as IDENT(tone) is ranked low enough, underlying representations with linked and unlinked tones map to the same outputs. Precisely because IDENT(tone) is bottom-ranked, the markedness constraints overwhelm any faithfulness effects and thereby fully determine the tonal pattern of the output, regardless of whether or how tones are linked in the input. Tableaux (32) and (33) illustrate:⁵

⁵ The ranking in (32) and (33) is based on examples like (5), which show that INITIAL-PLATEAU has to dominate NO-FLOAT.

(32) Kikuyu in parallel OT: Underlying linked tones

	L H L H patakasa	NO- SK(s)	INIT- PLAT	LNK- INIT	HAVE- T	NO- MC	NO- SK(t)	NO- FL	Id(t)
a. →	L H L H \ \ \ patakasa								5
b.	L H L H patakasa		1 W						L
c.	L H L H \ patakasa					1 W			1 L
d.	L H L H \ patakasa						1 W	1 W	2 L

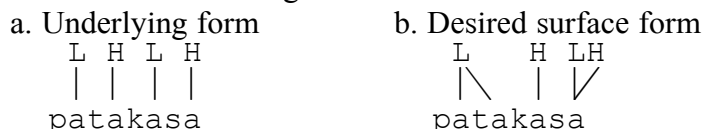
(33) Kikuyu in parallel OT: Underlying unlinked tones

	L H L H patakasa	NO- SK(s)	INIT- PLAT	LNK- INIT	HAVE- T	NO- MC	NO- SK(t)	NO- FL	Id(t)
a. →	L H L H \ \ \ patakasa								5
b.	L H L H patakasa				4 W			4 W	L
c.	L H L H patakasa		1 W						4 L
d.	L H L H \ patakasa					1 W			5
e.	L H L H \ patakasa						1 W	1 W	4 L

Why do P-OT and HS analyses differ so sharply in their ability to accommodate underlying representations with linked tones? Why is it that one is able to accomplish the mapping in (34) (repeated from (24)), while the other is not? The answer to this question is that the relationship between an underlying form and its surface form is different in P-OT and HS. In P-OT, a surface form is the most harmonic candidate derived from its underlying form. In HS, a surface form is the most harmonic candidate derived from the form that preceded it in the derivation, which is the most harmonic candidate derived from its predecessor, and so on, all the way back

to the underlying form. Whether a particular surface form is accessible from some underlying form in HS depends on whether there is a series of intermediate optima linking them by a succession of small changes. The mapping in (34) is impossible for the HS analysis because there is no such series of optima, as we saw when we attempted to construct one in (25)–(29).

(34) Effect of lexical linking



There is clearly a problem for HS here. One obvious solution to this problem is to prohibit linked tones in the Kikuyu lexicon. This is in fact the solution we adopt, but it cannot be true of the lexicon of Kikuyu alone — it must be true of every language or not at all. A fundamental premise of OT is that languages differ in constraint ranking. The null hypothesis is that languages differ *only* in constraint ranking. Lexica are language-particular, of course, but all systematic differences between languages are expressed by the grammar. This thesis is called richness of the base (ROTB), because it holds that the base (= original input to the grammar) is rich in the sense that it is not subject to any language-particular restrictions.

ROTB rules out a fairly standard pre-OT analytic move of saying that tones are underlyingly unlinked in languages where tone association is fully predictable, such as Kikuyu, but they are underlyingly linked in languages where tone association is unpredictable, such as Japanese (discussed in the next section). If we are to say that tones are underlyingly unlinked in the lexicon of Kikuyu, as the argument just presented shows, then we are obliged under ROTB to say that they are underlyingly unlinked in *all* languages.

Although our reasons for taking this position are novel, the position itself is not. According to Odden (1995: 468), “The view that tones and segments are underlyingly separated was pursued in early autosegmental phonology to the point that in Goldsmith (1976b) there are no lexical linkings between tones and vowels.” Lexical tone association was seen as unnecessary, and hence excluded by Occamite reasoning. The assumptions that made it unnecessary are also relevant to our proposal, as the next section will show.

5. Lexical tone association: Evidence and response

5.1. Introduction

Tone association is not always entirely predictable from simple phonological principles like those operative in Kikuyu. A language with left-to-right one-to-one tone association might have medial contour tones in certain words. Often, observations like these are just an indication that the analyst has oversimplified: the medial contour tones might have arisen from merged syllables.

It will sometimes happen, though, that the familiar phonological resources are insufficient to account for certain irregularities. In such cases, a surface contrast in tone association has to be encoded in the lexicon. We have argued that direct encoding is not an option. What alternatives are available to us?

In the early autosegmental literature, apparent contrasts in tone association were sometimes attributed to contrasts in the tone melody. In Etung, for example, there are contrasting tone patterns like *ákpùgà* ‘money’ versus *ésébè* ‘sand’.⁶ If these words are assumed to have the same

⁶ High tone is indicated by an acute accent and low by a grave accent.

HL tone melody, then underlying tone association would seem unavoidable. In Goldsmith's (1976b: 221ff.) analysis of Etung, however, these words are claimed to have different melodies: HL in *ákpùgà* and HHL in *ésébè*. If these are the underlying representations, then tone association can be determined by the grammar with no assistance from the lexicon. In 5.2, we analyze Venda in these terms, showing how ROTB actually forces this sort of analysis.

There is also a tradition of work on tone in Bantu and Japanese that posits diacritic accent (Goldsmith 1976b, 1982, 1984b, Haraguchi 1977, Hyman & Byarushengo 1980, Hyman 1981, 1982, Odden 1982, 1985). A diacritic accent is a phonetically uninterpretable lexical feature of a vowel (or mora). It attracts a particular tone, H or L in the analyses cited. In the literature on Bantu, the primary purpose of the diacritic accent is to allow for a modular system in which the locations of tones can be manipulated by rules prior to insertion of tone melodies (Goldsmith 1982: 48). In the literature on Japanese, however, diacritic accent is primarily used to mark contrasts in tone association, so Japanese will receive our attention in 5.3.

A third possible source of tone-association contrasts will be mentioned here for completeness but will not be discussed further. When contrasts emerge because different parts of the lexicon are subject to different regularities, lexical indexation of ranking and/or constraints may be the best analysis (Inkelas 1999: 143, Pater 2000, 2006, and others).

5.2. Venda

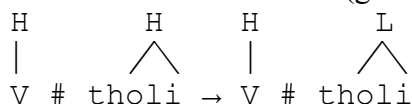
Venda has been cited as a Bantu language with an unusually wide range of contrasting tone patterns (Kisseberth & Odden 2003: 59). It has also provided the basis for an explicit argument in support of underlying tone association (Cassimjee 1992: 77--80). It therefore merits our close attention as representative of the kinds of challenges a theory without underlying tone association must face.

Here, we reanalyze the Venda evidence in HS under ROTB. ROTB not only forces HS to ban underlying tone associations, but it also rules out a key assumption that made underlying tone association necessary in the first place: the assumption that underlying representations are subject to the Obligatory Contour Principle (OCP), which prohibits adjacent identical tones (Goldsmith 1976b, Leben 1973). We argue instead that, although the OCP controls aspects of surface form, it does not and cannot control underlying forms. Once this argument is in place, the need for underlying tone association in Venda disappears.

Venda has two tones, H and L. In noun stems, every combination of H and L is possible (Cassimjee 1992). Thus, monosyllabic nouns are H or L, disyllabic nouns are HH, LL, HL, or LH, and so on. There is complete attestation of the eight possible patterns in trisyllables, and nearly complete attestation of the 16 possible patterns in quadrisyllabic stems.

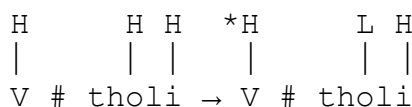
Cassimjee (1992: 77--80) uses this system of contrasts to argue for underlying tonal association. Her argument starts with a claim about the representation of high tone in Venda: tautomorphic sequences of H autosegments are prohibited. The evidence for this claim comes from the process known as "Meeussen's Rule" (Goldsmith 1984a and much subsequent work). Meeussen's Rule is a dissimilatory process that changes the second of two adjacent H tones into L. When a stem containing several high-toned syllables in a row undergoes Meeussen's Rule, all of the high-toned syllables become low:

(35) Effect of Meeussen's Rule (gloss: 'spy')



From this observation, Cassimjee infers that sequences of H autosegments are prohibited in underlying representation. If such sequences were allowed in the input to Meeussen's Rule, then only the first H in the sequence would lower:

(36) Not the effect of Meeussen's Rule



Cassimjee goes on to consider whether this result entails lexical tonal association under two different assumptions about L tone, underspecification and full specification. Since Venda presents no good evidence for underspecification of L, we focus on the stronger argument for lexical associations that is based on full specification.

The possible contrasts in quadrisyllables are relevant to this issue. Specifically, in her analysis HLH melodies can contrast in the choice of which H is multiply linked in quadrisyllables:⁷

(37) Contrast in H linking (with L specification)

Near-surface form	Underlying form	
	No lexical association	Lexical association
$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ \wedge & & \\ \text{gokoshomba} \\ \text{'Cape grape'} \end{array} $	$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ \text{gokoshomba} \end{array} $	$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ \wedge & & \\ \text{gokoshomba} \end{array} $
$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ & & \wedge \\ \text{sudzungwane} \\ \text{'shrub species'} \end{array} $	$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ \text{sudzungwane} \end{array} $	$ \begin{array}{ccc} \text{H} & \text{L} & \text{H} \\ & & \wedge \\ \text{sudzungwane} \end{array} $

Since HLH can be associated with a stem in at least two distinct, contrasting ways, lexically specified tone association would appear to be indispensable.

This argument rests on the assumption that the lexical representations of morphemes do not contain sequences of H autosegments. In other words, OCP(H) is a morpheme-structure constraint of Venda (Cassimjee 1992: 122). But OCP(H) is also active dynamically, in the course of phonological derivations. In Cassimjee's analysis, it is responsible for fusion of adjacent Hs between some prefixes and the stem in verbs (Cassimjee 1992: 124, 324). Furthermore, Meeussen's Rule has also been attributed to OCP(H), because it eliminates sequences of adjacent H tones dissimilatorily (Myers 1987, 1997 and others). When a phonological constraint has these dual roles — as a passive restriction on underlying morphemes and as an active participant in derivations — we have what is known as the Duplication Problem (Clayton 1976, Kenstowicz & Kisseberth 1977). The Duplication Problem is a problem because the analysis misses a generalization: two different grammatical components have the same constraint, with only its functions differing.

⁷ The ultimate surface forms are *gókóshômbá* and *súdzúngw!áné*, as a result of a process that spreads the first H rightward, creating a contour tone on the (long) penult or a floating L otherwise.

ROTB is OT's solution to the Duplication Problem (see McCarthy 2002: 71--76 for an explanation). By denying the existence of language-particular constraints on the lexicon, ROTB eliminates the possibility of this duplication. In the present instance, ROTB entails the possibility of lexical items with underlying tone sequences like HHLH and HLHH. Indeed, those can be exactly the tones of the lexical items in (37). In other words, a putative contrast in tonal association is reanalyzed as a contrast in tones themselves. Lexical association of tones is not necessary.

The principal challenge to this analysis is the Meeussen's Rule evidence in (35). If [gókóshòmbá] has separate H tones on its first two syllables, why does Meeussen's Rule lower both of them? The answer is that there are two H tones in the lexicon, but there is only one H by the time Meeussen's Rule comes into effect because sequences of H tones on adjacent syllables have already fused to satisfy OCP(H), defined as follows:

(38) OCP(H) (after Myers 1987: 154)

Assign one violation mark for every pair of H autosegments that are adjacent on their tier and linked to adjacent syllables.

Although it compels fusion later on, OCP(H) does not affect one-to-one left-to-right association at the initial stages of the derivation. That is because it is ranked below all of the constraints that are responsible for that pattern of association, as shown in tableau (39):⁸

(39) Step 2 from /gokoshomba, HHLH/

	H H L H gokoshomba	LNK- INIT	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	ID(t)	OCP(H)
a. →	H H L H gokoshomba		2	2			1	1
b.	H H L H gokoshomba		3 W	3 W			L	L
c.	H H L H ^ gokoshomba		3 W	2			1	L
d.	H H L H gokoshomba		2	2		1 W	1	L
e.	H H L H / gokoshomba		2	2	1 W		1	L

This derivation continues in the same vein, eventually producing [gókóshòmbá], with one-to-one association between tones and syllables:

⁸ Also see Reynolds (1997) for a parallel OT analysis of Venda nouns in a headed domains theory of tone representation.

(40) /gokoshomba, HHLH/ after step 4



Now that all of the constraints ranked higher than OCP(H) have been satisfied, attention turns to it. The OCP evokes various responses in the world's languages, and among them is fusion of the offending elements (Boersma 1998, Keer 1999, Myers 1987, 1997, Yip 1988). Fusion or coalescence of identical elements violates no faithfulness constraints except UNIFORMITY (McCarthy & Prince 1995: 371). We propose that OCP(H) dominates UNIFORMITY in Venda, so (40) is transformed into a representation where a single H autosegment is associated with the first two syllables:

(41) Step 5 from /gokoshomba, HHLH/

<table style="margin: 0 auto; border-collapse: collapse;"> <tr> <td style="text-align: center; padding: 0 5px;">H</td> <td style="text-align: center; padding: 0 5px;">H</td> <td style="text-align: center; padding: 0 5px;">L</td> <td style="text-align: center; padding: 0 5px;">H</td> </tr> <tr> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> </tr> <tr> <td colspan="4" style="text-align: center; padding-top: 5px;">gokoshomba</td> </tr> </table>	H	H	L	H					gokoshomba				LNK- INIT	NO- FL	HAVE- T	NO- SK(t)	NO- SK(s)	ID(t)	OCP(H)	UNIF
H	H	L	H																	
gokoshomba																				
a. → <table style="margin: 0 auto; border-collapse: collapse;"> <tr> <td style="text-align: center; padding: 0 5px;">H</td> <td style="text-align: center; padding: 0 5px;">L</td> <td style="text-align: center; padding: 0 5px;">H</td> </tr> <tr> <td style="text-align: center;">^</td> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> </tr> <tr> <td colspan="3" style="text-align: center; padding-top: 5px;">gokoshomba</td> </tr> </table>	H	L	H	^			gokoshomba										1			
H	L	H																		
^																				
gokoshomba																				
b. <table style="margin: 0 auto; border-collapse: collapse;"> <tr> <td style="text-align: center; padding: 0 5px;">H</td> <td style="text-align: center; padding: 0 5px;">H</td> <td style="text-align: center; padding: 0 5px;">L</td> <td style="text-align: center; padding: 0 5px;">H</td> </tr> <tr> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> <td style="text-align: center;"> </td> </tr> <tr> <td colspan="4" style="text-align: center; padding-top: 5px;">gokoshomba</td> </tr> </table>	H	H	L	H					gokoshomba										1 W	L
H	H	L	H																	
gokoshomba																				

All stem-internal sequences of H autosegments will fuse in this fashion. For example, the underlying representation of *bólóngóndó* is /bolongondo, HHHH/ with three OCP(H) violations at step 4 of its derivation. These violations are removed in three steps, each of which effects fusion of one pair of adjacent Hs. The order in which these tones fuse is indeterminate, but the final result is a single H linked to four syllables.

Meeussen's Rule — that is, dissimilation of H to L — competes with fusion as a way of satisfying OCP(H). In nouns, there is a clear division of responsibility, with fusion intramorphemically and dissimilation intermorphemically.⁹ The key to the analysis is that fusion is preferred to dissimilation because IDENT(H) dominates UNIFORMITY(H). But heteromorphemic fusion is blocked by the constraint MORPHDIS (abbreviated MDIS), which prohibits morphemes from sharing segments, tones, or other phonological elements (Keer 1999: 53, McCarthy & Prince 1995: 310). Tableau (42) illustrates with a schematic example of a heteromorphemic sequence of high tones. Dissimilation in (42)a competes against fusion in (42)c. Since this H sequence arises across word boundary, merger is blocked by MORPHDIS and dissimilation prevails.¹⁰

⁹ The situation in verbs is less clear; the choice between fusion and Meeussen's Rule in verbs is perhaps the most important incompletely resolved question in Cassimjee (1992).

¹⁰ Since it is not really relevant to our concerns, we set aside the question of why Meeussen's Rule affects the second H rather than the first. See Myers (1997: 887) for one proposal.

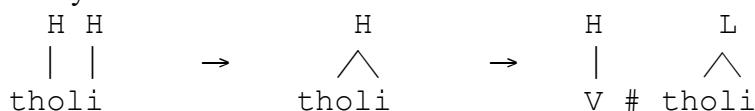
(42) Meeussen's Rule

At step 3 from /SS,HL/ after H-final word

	H # H L	LNK-INIT	NO-FL	HAVE-T	NO-SK(t)	NO-SK(s)	MDis	OCP(H)	NO-LNK(t)	ID(H)	UNIF(H)
a. →	$\begin{array}{c} H \# H L \\ \quad \\ S \# S S \end{array}$									1	
b.	$\begin{array}{c} H \# H L \\ \quad \\ S \# S S \end{array}$							1 W		L	
c.	$\begin{array}{c} H \quad L \\ \diagdown \quad \diagup \\ S \# S S \end{array}$							1 W		L	1 W

A final detail. Some sort of cyclic analysis is required to account for cases like (35) where a H-final word is followed by a noun whose underlying representation begins in an H sequence. The stem-internal H sequence is merged on the first or lexical cycle, and then it dissimilates when the preceding H becomes visible on the second or postlexical cycle:

(43) Cyclicity in H#HH...



See Wolf (2008) for a demonstration of how the cycle can be incorporated into a serial version of OT similar to HS.

As we saw in the last section, ROTB forces the conclusion that tone associations are universally absent from the lexicon. This conclusion is challenged by the complex tonal patterns of nouns in Venda, which seem to require lexical associations. But ROTB forces the analysis to include tone melodies that eliminate the need for lexical associations. In short, ROTB not only creates the problem but also solves it.

5.3. Japanese¹¹

Tokyo Japanese has a well-known surface contrast in tone association. For example, *kokoro-ga* 'heart-Subj' and *atama-ga* 'head-Subj' have the same HL tone melody, but differently aligned:¹²

(44) Surface contrast in tonal association



According to Poser (1984), this surface contrast is derived from an underlying contrast in where the H is associated:

¹¹ The material in this section is drawn from McCarthy & Pruitt (to appear).

¹² The basic tone melody of Tokyo Japanese was previously identified as LHL, not HL. Subsequent research showed that the initial L is an utterance-level boundary tone (Pierrehumbert & Beckman 1988: 135--136).

(45) Lexical contrast in tone association



If correct, this would obviously put paid to our claim that lexical tone associations are prohibited universally.

There is a competing hypothesis. According to Haraguchi (1977), this contrast is represented lexically with a diacritic accent, indicated here with a superimposed *x*: /kokō̃ro/ versus /atamã̃/. Unlike a tone, a diacritic accent is a phonetically uninterpretable feature. Its effect on the surface representation is thus always indirect, mediated by the grammar. The grammar of Tokyo Japanese associates H with moras that bear this feature.

Because a diacritic accent is not a tone but rather only a possible harbinger of a tone, it does not present the same difficulties for our HS analysis of Kikuyu as lexically associated tones do. Suppose that there is a constraint ACCENT→H that requires accents to be realized as high tones:

(46) ACCENT→H (ACC→H)

Assign one violation mark for every accent that is not associated with a high tone.

Moras that bear accents lexically receive H from the grammar because of this constraint. That is why Japanese has contrasts in tone association like (44).

Because of ROTB, Kikuyu can also have lexical items with diacritic accents. But this does not lead to contrasts in tone association in Kikuyu because ACCENT→H is ranked too low to matter. For example, imagine that the suffix /ɾ-ε/ is accented on the first syllable, and consider whether this would interfere with selecting the correct winner in tableau (20). Tableau (47) includes (20)'s winner and two candidates that ACCENT→H favors. It also shows the ineffectiveness of ACCENT→H when it is ranked low.

(47) Kikuyu with diacritic accent

$\begin{array}{c} L \quad H \quad L \quad H \\ \diagdown \quad \diagup \\ \text{tomar} \quad \text{r} \quad \text{ire} \end{array}$	HAVE- T	NO- SK(t)	NO- SK(s)	LNK- INIT	INIT- PLAT	ACC →H	NO- FL	ID(t)
a. → $\begin{array}{c} L \quad H \quad L \quad H \\ \diagdown \quad \\ \text{tomar} \quad \text{r} \quad \text{ire} \end{array}$	2					1	2	1
b. $\begin{array}{c} L \quad H \quad L \quad H \\ \diagdown \quad \diagdown \\ \text{tomar} \quad \text{r} \quad \text{ire} \end{array}$	2		1 W			L	2	1
c. $\begin{array}{c} L \quad H \quad L \quad H \\ \diagdown \quad \diagup \\ \text{tomar} \quad \text{r} \quad \text{ire} \end{array}$	2	2 W	1 W			L	2	1

Tableau (47) shows that HS can deal with lexical accent, even though it cannot deal with lexical tone association. The difference is that lexical accents can simply be ignored when ACCENT→H when is ranked low, but even ranking IDENT(tone) low is not enough to ensure that lexical tone associations will be ignored, as we saw in section 4. A lexical accent and ACCENT→H constitute a request for a particular kind of structure; they are not the structure itself. This is an important distinction.

Before embracing diacritic accent as the lexical representation of (some) contrasts in tone association, we need to address the arguments in Poser (1984) and Pulleyblank (1986) against a diacritic treatment of accent.

Poser (1984: 37) presents an argument from parsimony. Tokyo Japanese should be analyzed with lexically linked tones, as in (45), rather than diacritic accents because lexically linked tones are required in other languages to represent contrasts that cannot be analyzed with diacritics. Since lexically linked tones are sufficient to analyze Japanese, so the argument goes, parsimony demands that diacritic accents be dispensed with. Obviously, this argument has no force if underlying tonal associations are disallowed universally.

Another argument is that positing an accentual diacritic entails a sharp typological distinction between accentual and tonal languages (Pulleyblank 1986: 158--161): accentual languages like Japanese have the diacritic and tonal languages like Kikuyu do not. In reality, there is no sharp typological split between accentual and tonal languages (see Hyman 2009 for a recent review of the evidence). According to this argument, then, the accentual diacritic predicts a nonexistent typological difference.

Although this type of argument may be valid in the rule-based framework in which it was first made, it does not carry over to OT or by extension HS. In OT, typology emerges from ranking permutation. Typology does not come from differences in underlying representations; that is the point of ROTB. All languages have the possibility of accentual diacritics, but their grammars differ in the ranking of constraints that control overt realization of the diacritics, such as ACCENT→H. Depending on how those constraints are ranked, a language might seem more “accentual”, like Tokyo Japanese, or more “tonal”, like Kikuyu or Venda, but there will be no sharp typological difference. Whenever a phenomenon is controlled by multiple interacting constraints, as tone is, the typological space will contain many possibilities rather than just two.

Another objection to accentual diacritics is that they enjoy no independent support from the existence of stress-marking diacritics (Pulleyblank 1986: 156). Stress diacritics usually mark the heads of metrical constituents. The objection is that tonal diacritics do not mark heads. But there is a body of recent work arguing that the constituents defined by autosegmental spreading are headed (Cassimjee & Kisseberth 1997, 1998, Cole & Kisseberth 1995a, b, McCarthy 2004, Smolensky 1997, 2006). The notion that a diacritic can mark metrical *or* tonal heads does not seem so outlandish after all.

Perhaps the most compelling objection to accentual diacritics is also the one that is most clearly an empirical claim. Pulleyblank (1986: 154--157) describes a prediction that follows from diacritic accent. Since the accent diacritic and the tone that realizes it are distinct entities, it is in principle possible for a tone coming from one morpheme (A in (48)) to be realized on an accented syllable in a different morpheme (B in (48)). This process could be straightforwardly analyzed in a theory with diacritic accent, but not in a theory where accent is represented by a lexically linked tone. If this process is unattested, as Pulleyblank assumes, then this is a strong argument against the diacritic accent theory.

(48) A truly accentual process (after Pulleyblank 1986: 157)

$$[\dots \underset{*}{T} \dots]_A \quad [\dots]_B$$

\

$$[\dots]_A \quad [\dots \underset{*}{\sigma} \dots]_B$$

In fact, this supposedly impossible process actually occurs in Kansai Japanese compound nouns. To show this, we first need to present some background information.

Kansai nouns fall into two lexically specified tonal classes, HL and LHL. The location of accent — in Kansai, the rightmost syllable with which H is associated — must also be specified

lexically for each noun. Here are some examples in surface form with the accented syllable underlined:

(49) Tonal and accentual contrasts in Kansai-ben (Haraguchi 1999: 16--17)

LHL melody, accent on 2 nd syllable	L H L 	bitāmin	‘vitamin’
LHL melody, accent on 3 rd syllable	L HL \	nokogiri	‘saw’
HL melody, accent on 2 nd syllable	H L /	otōko	‘man’
HL melody, accent on 3 rd syllable	H L / \	kamināri	‘thunder’

When two nouns are joined in a compound, the entire compound has just a single HL or LHL melody. The underlying melody of the *first* member is preserved and that of the second member is lost (Haraguchi 1977: 95). But, as the $N_1 + N_2$ column in (50) shows, this melody is realized on the lexically-specified accent of the compound’s *second* member (unless that would put H on the final syllable).¹³ This is exactly the scenario in (48), a scenario that is consistent with marking the accented syllables by a diacritic, but not with marking them by a prelinked H.

(50) Kansai noun compounds

N_1	N_2	$N_1 + N_2$	
H L 	L H L ^	H L ^ ^ ^	
kāmi	hikōoki	kāmi hikōoki	‘paper + airplane’
L H L 	H L ^	L H L ^ ^	
harū	yāsumi	harū yāsumi	‘spring + holiday’
H L 	H L ^	H L ^ ^	
nātu	yāsumi	nātu yāsumi	‘summer + holiday’
L H L ^	H L ^	L H L ^ ^	
yakyuū	kūrabu	yakyuū kūrabu	‘baseball + club’
H L ^	H L ^	H L ^ ^	
sākkaa	kūrabu	sākkaa kūrabu	‘soccer club’

In summary, we have argued that diacritic accent is a viable alternative to underlying tone-TBU association. Indeed, if Kansai compound accent is an authentic example of the phenomenon in (48), diacritic accent may even be superior.¹⁴ This is important to our overall

¹³ This follows Kubozono’s (2008, p.c.) statement of the generalization. Kubozono also provided the data in (50).

¹⁴ Pulleyblank (1986: 157) says that (48) is “inherently impossible” in a theory that represents accent position with a prelinked tone. This is perhaps an overstatement, since the tonal theory might succeed in analyzing (48)

result because, as we showed in section 4, lexical tone linking is problematic for HS under our assumptions about GEN and CON, but diacritic accent is not.

6. Conclusion

Any interesting linguistic theory has unanticipated consequences. Harmonic Serialism is interesting because it has led to a body of interesting results (summarized in McCarthy to appear). Among its unanticipated consequences is its incompatibility with lexical tone linking in languages like Kikuyu where tone association is completely predictable.

As we noted in the introduction, the results in this paper converge with the conclusions about lexical marking of metrical feet in HS in McCarthy & Pruitt (to appear). This convergence is no accident. In both cases, we are dealing with phonological distinctions that are represented structurally. Furthermore, in both cases there are languages where the properties of that structure are fully predictable, though there are also others where they are not. Most importantly, in both cases there are constraints favoring full parsing of the structure: with metrical feet, the primary full-parsing constraint is PARSE-SYLLABLE, and with tone the constraints are HAVE-TONE and NO-FLOAT. Lexical structure is a parse, and removing it degrades performance on these constraints. In HS, where the removal of unwanted structure and the introduction of new structure cannot occur simultaneously, lexical structure may prove itself invulnerable precisely because of these pro-parsing constraints.

Dealing with this consequence of HS led us to explore alternative ways of lexically encoding unpredictable tone association, and this brought us back to ideas that were prevalent in earlier work on tone: lexical violation of the OCP and diacritic accent. We argued that these alternatives are not only viable but may be preferable to lexical tone linking. Obviously, there is much more to be said on this topic, and there is a vast literature on accent and its relationship to tone. But that is for future research.

References

- Boersma, Paul (1998). The OCP in functional phonology. *Proceedings of the Institute of Phonetic Sciences, University of Amsterdam* 22. 21-45.
- Cassimjee, Farida (1992). *An Autosegmental Analysis of Venda Tonology*. New York: Garland Pub.
- Cassimjee, Farida & Charles Kisseberth (1997). Optimal Domains Theory and Bantu tonology: A case study from Isixhosa and Shingazidja. In Larry Hyman & Charles Kisseberth (eds.) *Theoretical Aspects of Bantu Tone*. Stanford: CSLI Publications, 33-132.
- Cassimjee, Farida & Charles Kisseberth (1998). Tonal variation across Emakhuwa dialects. Benedict College, Columbia, SC, and Tel Aviv University.
- Clayton, Mary L. (1976). The redundancy of underlying morpheme-structure conditions. *Language* 52. 295-313.
- Clements, G. N. & Kevin C. Ford (1979). Kikuyu tone shift and its synchronic consequences.

with a sufficiently rich apparatus of ad hoc rules or constraints. The analysis of Kansai noun compounds sketched in Poser (1984: 242) is a case in point: it requires rules that delete only H in N_1 but H or L in N_2 , and it works only if the position of accent in the compound is assigned by rule (which is not generally true, according to Kubozono (2008)). It would be better to say that (48) is a natural and expected consequence of the diacritic theory that would require elaborate contortions in the tonal theory.

- Linguistic Inquiry* 10. 179-210.
- Clements, G. N. (1984). Principles of tone association in Kikuyu. In G. N. Clements & John Goldsmith (eds.) *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris, 281-339.
- Cole, Jennifer S. & Charles Kisseberth (1995a). An Optimal Domains theory of harmony. Urbana: University of Illinois.
- Cole, Jennifer S. & Charles Kisseberth (1995b). Nasal harmony in Optimal Domains Theory. Urbana: University of Illinois.
- Goldsmith, John (1976a). An overview of autosegmental phonology. *Linguistic Analysis* 2. 23-68.
- Goldsmith, John (1976b). *Autosegmental Phonology*. PhD dissertation, MIT.
- Goldsmith, John (1982). Accent systems. In Harry van der Hulst & Norval Smith (eds.) *The Structure of Phonological Representations*. Dordrecht: Foris, 47-63.
- Goldsmith, John (1984a). Meeussen's rule. In Mark Aronoff & Richard T. Oehrle (eds.) *Language Sound Structure*. Cambridge, MA: MIT Press, 245-259.
- Goldsmith, John (1984b). Tone and accent in Tonga. In G. N. Clements & John Goldsmith (eds.) *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris, 19-51.
- Haraguchi, Shosuke (1977). *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.
- Haraguchi, Shosuke (1999). Accent. In Natsuko Tsujimura (ed.) *The Handbook of Japanese Linguistics*. Malden, MA, and Oxford, UK: Blackwell, 1-30.
- Hyman, Larry & Ernest Byarushengo (1980). Tonal accent in Haya: An autosegmental approach. University of Southern California.
- Hyman, Larry (1981). Tonal accent in Somali. *Studies in African Linguistics* 12. 169-203.
- Hyman, Larry (1982). Globality and the accentual analysis of Luganda tone. *Journal of Linguistic Research* 2. 1-40.
- Hyman, Larry (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31. 213-238.
- Inkelas, Sharon (1999). Exceptional stress-attracting suffixes in Turkish: representation vs. the grammar. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press, 134-187.
- Keer, Edward (1999). *Geminates, the OCP, and the Nature of CON*. PhD dissertation, Rutgers University.
- Kenstowicz, Michael & Charles Kisseberth (1977). *Topics in Phonological Theory*. New York: Academic Press.
- Kisseberth, Charles & David Odden (2003). Tone. In Derek Nurse & Gérard Philippon (eds.) *The Bantu Languages*. London & New York: Routledge, 59-70.
- Kubozono, Haruo (2008). Japanese accent. In Shigeru Miyagawa & Mamoru Saito (eds.) *The Oxford Handbook of Japanese Linguistics*. New York: Oxford University Press, 163-189.
- Leben, Will (1973). *Suprasegmental Phonology*. PhD dissertation, MIT.
- McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *University of Massachusetts Occasional Papers in Linguistics* 18. Amherst, MA: GLSA Publications, 249-384.
- McCarthy, John J. (2000). Harmonic serialism and parallelism. In Masako Hirotsu (ed.) *Proceedings of the North East Linguistics Society* 30. Amherst, MA: GLSA

- Publications, 501-524.
- McCarthy, John J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, John J. (2004). Headed spans and autosegmental spreading. Amherst, MA: University of Massachusetts Amherst.
- McCarthy, John J. (2007). Restraint of analysis. In Sylvia Blaho, Patrik Bye & Martin Krämer (eds.) *Freedom of Analysis*. Berlin and New York: Mouton de Gruyter, 203-231.
- McCarthy, John J. (2008). The serial interaction of stress and syncope. *Natural Language & Linguistic Theory* 26. 499-546.
- McCarthy, John J., Wendell Kimper & Kevin Mullin (2010). Reduplication in Harmonic Serialism. Amherst: University of Massachusetts Amherst.
- McCarthy, John J. (to appear). An introduction to Harmonic Serialism. *Language and Linguistics Compass*.
- McCarthy, John J. & Kathryn Pruitt (to appear). Sources of phonological structure. In Ralf Vogel & Hans Broekhuis (eds.) *Linguistic Derivations and Filtering: Minimalism and Optimality Theory*. London: Equinox Publishing.
- Myers, Scott (1987). *Tone and the Structure of Words in Shona*. PhD dissertation, University of Massachusetts Amherst.
- Myers, Scott (1997). OCP effects in Optimality Theory. *Natural Language & Linguistic Theory* 15. 847-892.
- Odden, David (1982). Separating tone and accent: The case of Kimatuumbi. In D. Flickinger, M. Macken & N. Wiegand (eds.) *The Proceedings of the West Coast Conference on Formal Linguistics 1*. Stanford: Stanford Linguistic Association, 219-230.
- Odden, David (1985). An accentual approach to tone in Kimatuumbi. In D. Goyvaerts (ed.) *African Linguistics: Studies in memory of M. W. K. Seminenke*. Amsterdam: John Benjamins, 345-419.
- Odden, David (1995). Tone: African languages. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, MA, and Oxford, UK: Blackwell, 444-475.
- Pater, Joe (2000). Nonuniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17. 237-274.
- Pater, Joe (2006). The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In Leah Bateman, Adam Werle, Michael O'Keefe & Ehren Reilly (eds.) *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory III*. Amherst, MA: GLSA.
- Pater, Joe (to appear). Serial Harmonic Grammar and Berber syllabification. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya & Mariko Sugahara (eds.) *Prosody Matters: Essays in Honor of Lisa Selkirk*. London: Equinox Publishing.
- Pierrehumbert, Janet & Mary Beckman (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Poser, William J. (1984). *The Phonetics and Phonology of Tone and Intonation in Japanese*. PhD dissertation, MIT.
- Prince, Alan & Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA, and Oxford, UK: Blackwell.
- Pruitt, Kathryn (2008). Iterative foot optimization and locality in stress systems. Amherst, MA: University of Massachusetts Amherst.
- Pulleyblank, Douglas (1986). *Tone in Lexical Phonology*. Dordrecht: D. Reidel.

- Reynolds, William T. (1997). Post-High Tone Shift in Venda Nominals. Johannesburg: University of the Witwatersrand.
- Smolensky, Paul (1997). Constraint interaction in generative grammar II: local conjunction, or random rules in Universal Grammar. Baltimore, MD: Johns Hopkins University.
- Smolensky, Paul (2006). Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature-domain markedness. In Paul Smolensky & Géraldine Legendre (eds.) *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press/Bradford Books, 585-720.
- Wolf, Matthew (2008). *Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model*. PhD dissertation, University of Massachusetts Amherst.
- Wolf, Matthew & John J. McCarthy (2009). Less than zero: Correspondence and the null output. In Curt Rice (ed.) *Modeling Ungrammaticality in Optimality Theory*. London: Equinox Publishing.
- Yip, Moira (1988). The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19. 65-100.
- Yip, Moira (2002). *Tone*. Cambridge: Cambridge University Press.
- Zoll, Cheryl (1996). *Parsing below the Segment in a Constraint-based Framework*. PhD dissertation, University of California, Berkeley.