Stress and length in Hixkaryana

MORRIS HALLE AND WILLIAM J. IDSARDI

Abstract

An Optimality Theory analysis of the stress and length patterns of Hixkaryana is criticized. The analysis is shown to predict incorrect results for certain words, and is shown to be incompatible with the standard Optimality Theory account of acquisition. A rule-based account of Hixkaryana is offered and is shown to be superior to the Optimality Theory account.

1. Introduction

In this article we focus on the analysis of Hixkaryana stress in Chapter 4 of Kager 1999. The main conclusions of our discussion are:

1. The OT analysis offered by Kager is inadequate as given, and in order to work correctly it must be augmented with extra constraints of various types.
2. Kager's analysis is incompatible with other components of the OT research program, specifically the assumption of a universal constraint set, CON, and the learning algorithm of Prince and Tesar 1999. The universal set of constraints along with the learning algorithm force a much larger set of constraints to be active than appear in Kager's analysis. This means that Kager's attribution of effects to particular constraints is wrong, and that many constraints conspire to do the same work. The large number of constraints also leads to a vast space of constraint rankings, yielding far too much "delicacy" in ranking each constraint (see Cohn and McCarthy 1994; Idsardi 1995).
3. Kager's conceptual criticisms of rule-based analyses are misguided because they apply with equal force to his OT analysis and are, moreover, empirically incorrect.
4. There *are* conceptual and representational issues to be considered, but they favor the rule-based analysis, not the OT analysis.

The article below is organized as follows. In section 2, we show that Kager's OT analysis does not generate the correct forms in words beginning with a #LLH...

The Linguistic Review 17 (2000), 199–218 0167–6318/00/017-0199

sequence. In section 3, we investigate Kager's decomposition of the constraints involved in iambic lengthening, and show that they do not achieve Kager's aims. In section 4, we reconstruct an OT analysis that accounts for the facts of Hixkaryana, and conforms to the demands of the Prince-Tesar learning algorithm. We show that the algorithm leads to duplication of constraints and thereby renders the analysis needlessly cumbersome and inelegant. In section 5, we sketch an alternative rule-based account of the data. Section 6 outlines the conclusions of our investigations.

2. Kager's OT analysis

Kager devotes section 4.3.2 to the development of an OT analysis of the Hixkaryana facts. We have listed in (1) the set of forms considered by Kager.

(1)	a.	kʷa'ːja	'red and green macaw'
	b.	toro'ːno	'small bird'
	c.	atʃo'ːwowo	'wind'
	d.	nemo'ːkoto'ːno	'it fell'
	e.	a'kmata'ːrɨ	'branch'
	f.	to'hkurʲe'ːhona	'to Tohkurye'
	g.	to'hkurʲe'ːhona'ːhaʃa'ːha	'finally to Tohkurye'
	h.	na'kɲo'hja'tʃkena'ːno	'they were burning it'
	i.	kʰana'ːnɨ'hno	'I taught you'
	j.	mɨha'ːnanɨhno	'you taught him'

Stress and vowel length are predictable in Hixkaryana, in particular stressed vowels in open syllables are always long. Kager (p. 160) ends up with the constraint ranking in (2). The reader is referred to Kager's chapter for the definitions of the various constraints and their abbreviations.

(2) Non-Finality, GrWd = PrWd, Ft-Bin, Weight-to-Stress Principle (WSP)
≫ Uneven-Iamb
≫ Parse-Syl
≫ All-Ft-Left
≫ Dep-IO

We will follow Kager in evaluating the constraint Uneven-Iamb so that (LH) is preferred over (LL) and (H), which are in turn preferred over (HL) and (L) (see section 3, below, for more discussion). As shown in (3), the constraint ranking in (2) selects the wrong form ((3c) instead of (3a)) in words starting with #LLH... such as Kager's (16c) (= (1i)) (kʰan'aː)(nɨ'h)no. (Recall that stressed vowels in open syllables are lengthened.) All-Ft-Left is evaluated as a gradient constraint, counting the number of syllables from the left edge of the word to the left edge of each foot.

(3)

/k^hananɨhno/	Non-Fin	GrWd= PrWd	Ft-Bin	WSP	Uneven-Iamb	Parse-Syl	All-Ft-Left	Dep-IO
a. (k^ha.na:)(nɨh)no					*!	*	**	*
b. (k^ha:)(na.nɨh)no					*!	*	*	*
☞ c. k^ha(na.nɨh)no						**	*	

It is not so easy to correct the problem with the ranking in (2), for potential solutions interact with the decomposition of Uneven-Iamb, which we discuss in sec. 3. Moreover, we cannot promote Parse-Syl over Uneven Iamb because the form [aco:'wowo] (1c) requires the ranking Uneven Iamb ≫ Parse-Syl as shown by Kager (p. 154). We can improve matters by promoting an alignment constraint to require that words begin with a foot at the left edge, Align(PrWd, L, Ft, L) (Align-Left, p. 169), which must be ranked above Uneven-Iamb. In fact, Align-Left is true of the entire vocabulary, and so it can be ranked at the top of the constraint hierarchy. As we will see below, principles of OT in fact force Align-Left to be undominated if it is unviolated. The tableau for the new ranking is shown in (4).

(4)

/k^hananɨhno/	Non-Fin	GrWd= PrWd	Ft-Bin	WSP	Align-Left	Uneven-Iamb	Parse-Syl	All-Ft-Left	Dep-IO
a. (k^ha.na:)(nɨh)no						*	*	**!	*
☞ b. (k^ha:)(na.nɨh)no						*	*	*	*
c. k^ha(na.nɨh)no					*!		**	*	

However, this new ranking still chooses the wrong candidate. The high ranking of Align-Left, moreover, has the disadvantage of duplicating much of the work (and explanatory force) attributed to All-Ft-Left in Kager's original analysis. Having both Align-Left and All-Ft-L ranked high misses a generalization, as explained below, and gives a less parsimonious analysis. Moreover, once we have determined that Align-Left is undominated, there is nothing left for GrWd=PrWd to do in the analysis, since every word must have at least one foot.

Having ranked Align-Left to eliminate candidate (3c), we now observe that candidate (4b) will win with this ranking, which is still the wrong result. Again, it is still preferable to lengthen the first vowel rather than the second, as the form which lengthens the first vowel will better satisfy All-Ft-Left. What aspect of (4b) would render it worse than (4a)? It must be that (4b) is worse than (4a) because of the vowel that was lengthened in (4b). Following this line of reasoning, we could activate a positional faithfulness constraint (Beckman 1997, Kager pp. 407ff.) for initial syllables which would prevent them from lengthening (Dep-#_). However, bisyllabic words (1a) *do* lengthen the first syllable; so we need to rank this constraint appropriately. As Kager notes (p. 160), his analysis actually has two sources of lengthening: the ranking Ft-Bin ≫ Dep-IO, which supports lengthening in bisyllabic words, and the ranking Uneven-Iamb ≫ Dep-IO, which induces lengthening in other words. Therefore, we would need to rank Dep-#_ between Ft-Bin and Uneven-Iamb. This will correctly pick (4a) over (4b).

But we can construct longer words whose internal sequences include the same pattern – words of the form ...LLH..., which parse as ...(LL)(H)... . In such cases positional faithfulness for the first vowel will be of no help. This mode of parsing a "trapped" LL sequence into a foot is also observed in Chugach Alutiiq (Leer 1985). Notice that this parsing creates two non-canonical iambs instead of "looking ahead" to skip a single light syllable in order to form a single canonical iamb. This happens regardless of the position in the word in Chugach Alutiiq, and there is no evidence to suppose that Hixkaryana is any different on this issue. Thus, positional faithfulness to the first vowel is not the whole issue (though the availability of such a constraint exacerbates the learning problem, as we show below).

What other constraint or constraints does (4b) violate that (4a) doesn't? The only reasonable observation that connects the failure of (4b) with other issues is that (4a) has lengthened a vowel to create a "canonical" iamb (LH), whereas (4b) has lengthened a vowel to create a sub-canonical iamb (H). So the relevant constraint would be as in (5).

(5) Dep-(H): Don't add a mora to create a (H) foot.

This constraint should properly be understood as a conjoined constraint, putting together Dep-IO with Uneven-Iamb (but see below for the decomposition of Uneven-Iamb into separate constraints). The need for the conjoined constraint means that the ranked interaction of Uneven-Iamb ≫ Dep-IO (that is, the normal constraint interaction mechanism of OT) is not sufficient to do the required work. Rather we need a separate, additional mechanism to get the combined effect, namely constraint conjunction. But such an analysis is subject to the very objection Kager makes against Hayes's analysis – that the mechanisms for lengthening recapitulate the restrictions on foot parsing.

In sum, we have shown that Uneven-Iamb by itself is not sufficient to explain the lengthening pattern in Hixkaryana. This analysis misses the correct generalization about lengthening because it requires several conspiring constraints and constraint rankings to select the proper forms. In addition, we will need to properly rank Dep-(H) to allow bisyllabic words to lengthen the initial syllable so as to create (H) syllables in such words. The constraint ranking must be: Ft Bin, Non-Finality ≫ Dep-(H) ≫ All-Ft-Left, as in (6).

(6)

/kʰananɨhno/	...	Ft-Bin	Align-Left	Uneven-Iamb	Parse-Syl	Dep-(H)	All-Ft-Ft-Left	Dep-IO
☞ a. (kʰa.na:)(nɨh)no				*	*		**	*
b. (kʰa:)(na.nɨh)no				*	*	*!	*	*
c. kʰa(na.nɨh)no			*!		**		*	

It is also the case that Non-Finality is violated in (H) words, as in (7).

(7)	yoh	'chief of'	(Derbyshire 1985: 17)
	kay	'he did it'	(Derbyshire 1985: 23)
	mIn	'house of'	(Derbyshire 1985: 28)

Thus, Non-Finality must be dominated by at least one of GrWd=PrWd, WSP or Align-Left. The amended rankings for the constraints is given in (8).

(8) GrWd = PrWd, Ft-Bin, WSP, Align-Left
$\gg$ Non-Finality
$\gg$ Uneven-Iamb
$\gg$ Parse-Syl
$\gg$ Dep-(H)
$\gg$ All-Ft-Left
$\gg$ Dep-IO

The ranking in (8) differs from Kager's original account, (2), in having more active constraints, and being further stratified, both indications of a more complex analysis. Prince and Tesar (1999) give a formulation for markedness of constraint rankings, whereby the least-marked grammar has exactly two strata, with all of the Faithfulness constraints in the lower stratum. We discuss further consequences of the Prince-Tesar proposal in section 4, but we note here that one consequence is that in OT the preferred grammar would *not* activate the unusual constraint Dep-(H), and it would likewise be preferable not to rank Align-Left separately from All-Ft-Left. That is, in both an intuitive and a technical sense Kager's original analysis is less marked than the analysis which actually generates the Hixkaryana forms, (8). But Kager's analysis wrongly predicts that #LLH... words will be parsed *#L(LH).... The conclusion is inescapable that for OT the unattested parsings are more natural and less marked than those that actually occur for such words in Hixkaryana. We do not know of a single language that parses feet in such a fashion. As we show below, such analyses are correctly excluded by the rule-based account, but they are incorrectly allowed, even preferred, in OT.

The general moral here is that the vast space of possible rankings available in OT yields a grammar space of extreme "delicacy and specificity". This is highly problematic for the language learner, as Cohn and McCarthy explicitly argue, quoted in (9).

(9) "... linguistic parameters do not usually show this degree of delicacy and specificity, where natural changes ... lead to differences in just one form, with no repercussions elsewhere in the system." (Cohn and McCarthy 1994: 68)

Grammar delicacy is undesirable because it makes learning extremely difficult. If #LLH... words can have many different analyses while the rest of the grammar remains constant, the child has no alternative but to attend specifically to the #LLH... forms.

3. The decomposition of the Uneven-Iamb constraint

Kager says that Uneven-Iamb is a cover constraint combining the effects of more basic constraints, and in section 4.4.5 he decomposes Uneven-Iamb into the three rhythmic constraints, as in (10).

(10)
a. RhType=I: Feet have final prominence (p. 172)
b. RhType=T: Feet have initial prominence (p. 172)
c. RhContour: A foot must end with a strong-weak contour at the moraic level (p. 174)

The iambic part of Uneven Iamb is accomplished by ranking RhType=I ≫ RhType=T. The unevenness comes from RhContour. The immediate problem is that this does not actually achieve the goal set out for it: to derive the relative markedness scale LH > {LL, H} > {HL, L}. We show in (11) how the various foot types fare on these constraints. We have included the Stress-to-Weight Principle constraint (Kager: p. 172, fn. 9, 268) for completeness and clarity.

(11)

	RhType=I	RhType=T	RhContour	FtBin	WSP	SWP
(LH′)		*				
(LL′)		*	*			*
(HL′)		*	*		*	*
(HH′)		*			*	
(L′H)	*		*		*	*
(L′L)	*					*
(H′L)	*		*			
(H′H)	*		*		*	
(H′)						
(L′)			*	*		*
(H)	*	*	*		*	
(L)	*	*	*	*		

The problem is that (H′) is the best foot of all, not violating any of these constraints. Since all other feet violate at least one constraint, no matter what ranking of these constraints is chosen, (H′) is the best possible foot for these constraints. But what Kager wants and needs is for (LH′) to be superior to (H′) (at least in iambic systems, that is those for which RhType=I ≫ RhType=T). In order to accomplish this we need to raise the ranking of another constraint or to modify an existing constraint so that (H′) violates at least one constraint that (LH′) does not. Although this is a necessary condition, it is not sufficient, for on Kager's view (and Hayes's) the superiority of (LH′) to (H′) is a language universal, and is not just contingently true. This can be achieved in several ways. One is to modify Ft-Bin or to introduce a second type of Ft-Bin constraint so that there is a constraint forcing feet to be strictly bisyllabic. (LH′) would satisfy the bisyllabic requirement whereas (H′) would not. Another possibility would be to modify RhContour or activate an additional contour constraint so that feet must also start with a weak-strong contour at the moraic level. Choosing to modify RhContour to require both a beginning and an ending contour gives the results in (12).

(12)

	RhType=I	RhType=T	RhContour	FtBin	WSP	SWP
(LH′)		*				
(LL′)		*	*			*
(HL′)		*	*		*	*
(HH′)		*	*		*	
(L′H)	*		*		*	*
(L′L)	*		*			*
(H′L)	*		*			
(H′H)	*		*		*	
(H′)			*			
(L′)			*	*		*
(H)	*	*	*		*	
(L)	*	*	*	*		

Now no foot is universally the best, but (LH′) will be the best iambic foot, provided that RhContour ≫ RhType=T. In trochaic systems (those in which RhType=T ≫ RhType=I) the (H′) foot will still be the best, as it violates only RhContour, whereas the next best foot, (H′L), also violates RhType=I, an emergence of the unmarked phenomenon, showing iambic effects in trochaic systems. Note that this effect would disappear if foot type is a parametric choice rather than the relative ranking of RhType=I and RhType=T, because in that case there would be no constraint exerting pressure (however slight) for the opposite foot type.

In sum, in iambic systems we need to have RhContour ≫ RhType=T. But in trochaic systems RhContour must not dominate RhType=T, because then (LH′) would be preferred even though it violates RhType=T. Thus, RhContour and RhType=I seem to be bound together in their ranking possibilites, as we do not want to allow the ranking RhContour ≫ RhType=T ≫ RhType=I. That is, not all six possible rankings of the three constraints are allowed to occur, and specifically the ranking in (13) must be universally disallowed.

(13) No language can have the ranking:
RhContour ≫ RhType=T ≫ RhType=I

For trochaic systems so far we have (H) > (LL) > (LH). But (L) is not so easy to compare. It violates Ft-Bin but not FtType=I. Since Kager needs (L) to be worse than (LL), Ft-Bin must outrank FtType=I, and this can probably be declared a universal ranking. Once again the full factorial typology is not available, and the rankings of Ft-Bin, RhContour and RhType=I seem to be intimately connected in various ways.

If we now substitute the result in (13) for Uneven Iamb in the constraint ranking for Hixkaryana from (8) we obtain the constraint ranking in (14). This is only the first step in this process, we will determine the rankings of the rest of the constraints in (12) below.

(14) GrWd=PrWd, Ft-Bin, WSP, Align-Left
$\gg$ Non-Finality
$\gg$ RhType=I, RhContour
$\gg$ RhType=T
$\gg$ Parse-Syl
$\gg$ Dep-(H)
$\gg$ All-Ft-Left
$\gg$ Dep-IO

Since Uneven-Iamb has been separated into its component parts, we need to revise the constraint ranking once more. There is no reason to expect the component parts of decomposed Uneven-Iamb to cohere together in the ranking. We take up this question in the next section.

We have demonstrated that deriving the "iambic-trochaic law" in OT is not as easy as Kager claims. The RhContour constraint must enforce a "mountain-top" moraic structure of (wsw), and we cannot allow the full factorial typology, instead there must be universal restrictions on the constraint orderings allowed.

But there is still more compelling empirical evidence against the "iambic-trochaic law". Iambic footing and lengthening cannot be intrinsically connected, as they exist independently in human languages, so there is no necessary and intimate connection between them. For example, there are trochaic languages with stressed syllable lengthening (English, Halle 1998; Tiberian Hebrew, Prince 1975, and Rappaport 1984; Huariapano, Parker 1998), and there are iambic languages without stressed syllable lengthening (Creek, Haas 1977). Rather than try to enshrine these vague tendencies into Universal Grammar, we suggest instead that there are perceptual biases which nudge learners toward analyzing systems with alternating vowel length as iambic if they can do so (see Bell 1977). Instead of importing non-linguistic auditory perceptual biases into UG, we suggest that these biases provide preferred perceptions to the learners, which they try to maintain if they can. That is, if there are no interfering factors, children will hear ...LH′LH′LH′LH′LH′... as ...(LH′)(LH′)(LH′)(LH′)(LH′)... and will hear ...LL′LL′LL′LL′... as ...L(L′L)(L′L)(L′L)(L′.... If the stress system can be analyzed in accordance with this, then the child will do so. If other factors override this tendency (provide better cues to the system), then the child will end up with grammars that violate the "iambic-trochaic law". The moral here is that UG is not the only source of explanation, the perceptual system and its interaction with the language acquisition device can also explain cross-linguistic tendencies.

4. Universal CON and the attribution of explanation

We argued above in section 2 that activating Dep-(H) should be preferred to activating Dep-#_ because it will explain more cases in a more parsimonious fashion. But we will now show that this form of argumentation is at odds with other aspects of the OT program, in particular the universality of constraints (Kager: pp. 18, 21) and the Recursive Constraint Demotion (RCD) theory of learning constraint rankings (Kager: chapter 7).We adopt here the RCD as extended and modified by Prince and Tesar (1999) to reflect differences between markedness and faithfulness constraints. Following Prince and Tesar, we call the revised version the Biased Constraint Demotion algorithm (BCD).

Universality of the constraint set CON requires all constraints to be part of Universal Grammar (we put aside questions about morpheme specific constraints, which do not play any role in the analyses here under consideration). BCD starts with a minimally stratified constraint hierarchy where all markedness constraints are lumped together in a stratum dominating a stratum of faithfulness constraints, schematically $\{M\} \gg \{F\}$. As the learner acquires the grammar, it becomes progressively more stratified, while still adhering to an evaluation metric that ranks markedness constraints as high as possible, whereas faithfulness constraints are ranked as low as possible. In addition, conjoined constraints and positional faithfulness constraints must be ranked above their component or simple constraints (Kager: pp. 393, 409).

With this in mind, let us return to the question of the choice between Dep-#_ and Dep-(H). Prince and Tesar do not discuss this particular case, but they do discuss the general problem, and they draw the conclusion in (15).

(15) "It appears that the algorithm should choose the F constraint that is relevant to the narrowest range of structural positions. By this, any F/P for some position P should be favored over any F, given the choice, even if they constrain different elements." (Prince and Tesar 1999: 22)

What (15) says is that in order for the algorithm to work, markedness constraints must be demoted below the most specific faithfulness constraint which could choose the correct candidate ("the F constraint that is relevant to the narrowest range of structural positions"). Further, there is an intrinsic ranking among constraints (mirroring the Elsewhere Condition – see Halle and Idsardi 1998) such that more specific constraints outrank less specific ones. Together, (15) and the intrinsic ranking by specificity force the BCD to be conservative in its changes in the constraint ranking, choosing to demote a markedness constraint by the smallest amount possible. Prince and Tesar further show that positional faithfulness constraints are problematic for their algorithm in a way in which positional markedness constraints are not. However, in the present case the positional faithfulness constraint on initial vowels cannot be replaced by a positional markedness constraint, because it is simply not the case that non-initial positions are less marked positions for long vowels or heavy syllables. In fact, if there is a skewed distribution of weight, it seems more plausible that cross-linguistically initial syllables are more likely to be heavy than non-initial syllables (e.g., Korean – Sohn 1999: 157). So, in the present case it is Dep-#_ that will be chosen first by the principle in (15) rather than Dep-(H) precisely because Dep-#_ is more restricted, applying only to initial syllables (which if they lengthen, coincidentally violate Dep-(H)). Of course, upon receiving data from longer words which independently support Dep-(H) that constraint will rise in the ranking as well. So the answer is that the universality of the constraint set CON coupled with the BCD together require both Dep-#_ and Dep-(H) to be active. The ranking therefore is that in (16). (Actually, the acquired ranking is dependent on the order of presentation of forms to the learner. Learners can arrive at (14) as long as they are exposed to a #LLLLH... word and thus have to raise Dep-(H) in the rankings before they hear a #LLH word, which would otherwise motivate raising Dep-#_.)

(16) GrWd=PrWd, Ft-Bin, WSP, Align-Left, RhType=I
≫ Non-Finality
≫ RhContour
≫ RhType=T, Parse-Syl
≫ Dep-(H), Dep-#_
≫ All-Ft-Left
≫ Dep-IO

In (16), following the BCD, we have ranked the markedness constraints as high as possible, and the faithfulness constraints as low as possible. We return now to our #LLH... example in (17) to illustrate some further aspects of this ranking. Even though in Hixkaryana All-Ft-L is ranked fairly low, because it is a markedness constraint, it is ranked as high as it can be.

(17)

/kʰananɨhno/	...	Align-Left	Rh-Con	Rh-Type=T	Parse Syl	Dep-(H)	Dep-#_	All-Ft-Left	Dep-IO
☞ a. (kʰa.na:)(nɨh)no			*	*	*			**	*
b. (kʰa:)(na.nɨh)no			*	*	*	*!	*!	*	*
c. kʰa(na.nɨh)no		*!			**			*	

Notice that Align-Left still does crucial work in the analysis, for RhContour must be dominated by Align-Left, otherwise (17c) will win. In addition, Dep-(H) and Dep-#_ are ranked above All-Ft-Left to choose (17a) over (17b).

Let us now consider where the rest of the universal constraint set relating to metrical structure will be ranked. We start with the obvious alternative to (the now unpacked) Uneven-Iamb, the Stress-to-Weight Principle (SWP: 172, fn. 9, 268), "if stressed, then heavy". SWP is a markedness constraint and, therefore, by the BCD it must be ranked as highly as possible. In fact, all stressed syllables in Hixkaryana are heavy, and therefore we can conclude that SWP is undominated, so that we have the ranking in (18).

(18) GrWd=PrWd, Ft-Bin, WSP, SWP, Align-Left, RhType=I
$\gg$ Non-Finality
$\gg$ RhContour
$\gg$ RhType=T, Parse-Syl
$\gg$ Dep-(H), Dep-#_
$\gg$ All-Ft-Left
$\gg$ Dep-IO

But now, what do we attribute the preference of (LH) over (LL) to? Both RhContour and SWP favor (LH′) over (LL′), but SWP is unviolated (and undominated) whereas RhContour is violated in (17). So vowels lengthen in Hixkaryana because they are stressed, *not* because they are in iambic feet. This is precisely contrary to Kager's statements, but it is the result forced by the BCD. Since SWP is a member of the universal constraint set, and since it is a markedness constraint, it must be ranked as high as possible. Since it is never violated by any form in Hixkaryana, it must be in the undominated stratum. Therefore, it is SWP that selects the winning candidates with lengthening, not RhContour.

The next constraint to be considered is *Clash, which prohibits adjacent stressed syllables (p. 165). So far *Clash has played no role in the analysis, but being a markedness constraint, it must be ranked as high as possible by virtue of the BCD. We know that *Clash must be dominated by WSP, as adjacent heavy syllables receive stress. Avoiding clash also does not cause feet to flip to become trochaic, say in #LLH... words, and consequently RhType=I $\gg$ *Clash. And in fact, we can't skip parsing syllables to avoid clash in ...LLH sequences, so Parse-Syl $\gg$ *Clash.

In addition, we know that Hixkaryana does not delete any segments to achieve light syllables, and therefore Max-IO $\gg$ *Clash. Similarly, it doesn't delete a vowel in /LLLL/ to give (LH)L, so Max-IO $\gg$ Parse Syl. In sum we have the ranking in (19), with *Clash ranked as high as possible, and Max-IO ranked as low as possible.

(19) GrWd=PrWd, Ft-Bin, WSP, SWP, Align-Left, RhType=I
$\gg$ Non-Finality, Max-IO
$\gg$ RhContour
$\gg$ RhType=T, Parse-Syl
$\gg$ Dep-(H), Dep-#_

>> All-Ft-Left, *Clash
>> Dep-IO

But adding *Clash to the constraint ranking of Hixkaryana does no additional work. It is just a fifth wheel in the analysis, and even in Hixkaryana the work that it could do cannot replace the work done by any of the other constraints. Thus, *Clash cannot simplify the analysis at all. *Clash is simply not a factor in the stress assignment of Hixkaryana. Because of the doctrine of universal CON, however, OT has no choice in the matter, and moreover the BCD also insists that the *Clash constraint be ranked as high as possible. Consequently, the child must lower *Clash in the ranking during acquisition, as its position near the bottom of the hierarchy is far away from its original position.

There are still some remaining constraints to be considered. The constraint WSP-Ft (a constraint against spondees, *(HH) p. 184) is undominated. As Kager points out on pages 165–166, Prince and Smolensky's constraint "Non-Finality" can be interpreted in two ways – as a constraint against PrWd-final feet (the use of Non-Finality, above) or as a constraint against a final stressed syllable, ("No prosodic head is final in PrWd" p. 165). We will call the latter interpretation Non-Final-Stress. Non-Final-Stress is violated in monosyllables, and must therefore be ranked alongside (foot) Non-Finality. The constraints Align-Right and All-Ft-Right must be ranked below All-Ft-Left. Putting these requirements together, we have the ranking in (20).

(20) GrWd=PrWd, Ft-Bin, WSP-Ft, WSP, SWP, Align-Left, RhType=I
>> Non-Finality, Non-Final-Stress, Max-IO
>> RhContour
>> RhType=T, Parse-Syl
>> Dep-(H), Dep-#_
>> All-Ft-Left, *Clash
>> Align-Right, All-Ft-R, Dep-IO

Particular note is to be taken of the difference between (20) and (2). Kager's presentation makes it seem that (2) is all that is required to correctly assign stress and length in Hixkaryana. The facts of Hixkaryana would seem to require some modifications, but it is the conceptual structure of OT which forces (20) to be three times larger than (2), with additional stratification of the constraint hierarchy.

It is the twin doctrines of a universal constraint set CON and the BCD algorithm that cause the grammar to be cluttered with constraints that duplicate the efforts of other constraints. This is not constraint interaction, but a conspiracy of constraint duplication. The most obvious case we have already discussed – RhContour and SWP both mandate lengthening at the end of iambic feet, but since SWP is undominated and RhContour is not, we must ascribe lengthening effects to SWP and not to RhContour. Here it is not the iambic foot structure that causes lengthening in Hixkaryana, rather it is the fact that the syllable is stresssed. In other cases we cannot tell which constraint is responsible for the effects exhibited. For example, both GrWd=PrWd and Align-Left ensure that all words have at least one foot, but it is impossible to remove either one from the ranking. Is the phenomenon of culminativity to be ascribed to GrWd=PrWd, to Align Left or to both? Being undominated, these constraints do contribute to the selection of winners, but in a clumsy and uninsightful way.

Further examples of the same point are the relative contributions of WSP-Ft and WSP, which both penalize unstressed heavy syllables. WSP-Ft, being more specific, would be required to outrank WSP, but such universal ranking requirements only introduce extra structure into the undominated stratum of the hierarchy, to no empirical or conceptual advantage. Is it the case that the structure (HH') is demonstrably worse than H(H') in Hixkaryana? This is what the existence of WSP-Ft ≫ WSP entails. Likewise, since Non-Final-Stress and Non-Finality achieve the same ends in uniformly iambic systems (where trochaic feet are never allowed), but because they are markedness constraints they must be ranked as highly as possible, and therefore must duplicate each other in the selection of winning candidates. The coverage provided by Dep-#_ is subsumed by Dep-(H), but since positional faithfulness constraints are provided by UG, there is no reason to suppose that they won't be invoked by the learner, and ranked together, again doing the same work without providing any additional coverage. To pick an example from another part of the constraint hierarchy, Align-Left and All-Ft-Left overlap in both mandating a foot at the left edge of the word. It is empirically necessary to separate them in the grammar of Hixkaryana, but what is the conceptual advantage? Separating Align-Left from Align-Ft-Left means that the left alignment of the first foot remains unrelated to the left alignment of the rest of the feet, another constraint conspiracy.

We conclude that the Biased Constraint Demotion algorithm coupled with a univeral constraint set inevitably leads to constraint duplications and conspiracies. It is these two doctrines that ensure that OT is necessarily conspiratorial and inelegant.

5. Rule-based analyses of Hixkaryana stress

Kager compares his OT-based analysis of Hixkaryana stress with a rule-based alternative which is essentially that of Hayes (1995). This rule-based analysis is given in (21) (= Kager's (17), p. 149).

(21)
a. Step 1: Syllabify (open syllables are light, closed syllables are heavy)
b. Step 2: Mark the final syllable of each word as extrametrical.
c. Step 3: Assign iambs {(LH), (LL), (H)} iteratively from left to right.
d. Step 4: When the entire metrical domain is a single light syllable, assign (L) to it.
e. Step 5: Lengthen the vowel of each strong open syllable.

We review below each of Kager's criticisms of the rule-based approach and show either that they are untrue or that they apply with equal force to his own OT analysis.

Kager's first objection is that

(22) "the rule that assigns iambs (Step 3) is not intrisically connected to the rule that lengthens open syllables in strong positions of feet (Step 5), although both rules aspire toward a single output target: the canonical iamb (L**H**). This generalization is missed." (p. 150)

It is true that the rule-based analysis has distinct mechanisms for footing and lengthening, but, as we showed above, this is true of the OT analysis as well, as the SWP mandates length regardless of the foot type. In the OT account, it is the stress rather than the foot type that causes lengthening in Hixkaryana. In addition, as also pointed out above, iambic footing and lengthening cannot be intrisically connected, as they exist independently in languages, so there is no necessary and intimate connection between them. Additionally, in Hixkaryana the open syllable lengthening also takes place in the initial syllable of bisyllabic forms, which does not result in the "canonical iamb" (LH), but rather produces (H)L.

This points to a missed generalization in the OT analysis. As Kager remarks, his OT account does not intrinsically connect "iambic" lengthening (the cases mandated by RhContour $\gg$ Dep-IO) with initial bisyllabic lengthening (the cases due to FtBin $\gg$ Dep-IO). In fact, the OT account has at least three separate pressures for lengthening: Ft-Bin, SWP and RhContour. In contrast, the lengthenings are all captured by a single rule in the rule-based analysis. Thus, the rule-based analysis captures the generality of lengthening in Hixkaryana, which is not captured by the OT analysis, which has a conspiracy of three constraints mandating lengthening.

Kager's second objection is quoted in (23).

(23) "... this analysis relies on an intermediate stage in the derivation of disyllabic words in which a degenerate foot (**L**) is temporarily allowed (Step 4). This foot is repaired at the surface by lengthening its vowel into (**H**). This (**L**) foot is abstract since the output pattern contains only (**H**) and (L**H**). Again, a generalization is missed." (p. 150)

The rule-based analysis does have such an intermediate stage for bisyllabic words. Regardless of the syllable weight of the first syllable, such words are footed. Implicit in this analysis is the claim that such short words could – in the unmarked case – be unaffected by the general lengthening process of the language, that is, that the (L) feet could surface. Although we lack direct supporting data for this claim (say from alternative dialects of Hixkaryana), the claim itself does not seem to us wildly implausible. Indeed, it is possible to amend the OT constraint ranking so as to produce such an effect, as must be done for languages that do possess degenerate feet. In fact there are several ways to achieve such an effect. One obvious way is to raise the positional faithfulness constraint Dep-#_ above the constraints mandating length (SWP, Ft-Bin, RhContour). Likewise, if Parse-Syl, Dep-IO $\gg$ SWP, RhContour this will also select degenerate feet without lengthening. So OT does not ban a language that differs minimally from Hixkaryana by not lengthening the initial syllable in bisyllabic words. In fact, such a language can be created by promoting a single constraint (Dep-#_) to the top of the Hixkaryana ranking. Since both OT and rule-based systems can describe both kinds of languages, they are completely comparable in this regard, and this point cannot serve as crucial evidence to decide between the competing theories of stress.

Kager's third objection is quoted in (24).

(24) "... the analysis [in (21)] fails to explain why a degenerate foot is assigned in disyllabic words, whereas the prefinal syllable of the longer words ... remains unparsed. That is, Step 4 is just an ad hoc incarnation of 'culminativity' – it only applies in disyllabic words to satisfy the imperative that every word must have a foot." (pp. 150)

The treatment of degenerate feet to which Kager objects in (24) is not a property of rule-based accounts in general, but is rather a consequence of Kager's decision to adopt Hayes's 1995 views of foot structure. A proper criticism of Hayes's metrical theories would take us beyond the natural confines of this review of Kager's OT analysis of stress. We note, however, that a theory of metrical structure that on crucial points dissents from Hayes 1995 was advanced in Idsardi 1992 and has since been tested on empirical data of great variety and complexity in a number of studies, among them (on Russian, Lithuanian, and other Indo-European languages, in Halle 1997; on English, in Halle 1998; on tonal phenomena in Bantu languages and in Japanese, in Purnell 1998).

One of the crucial points on which this alternative theory dissents from that of Hayes 1995 and his followers is the assumption that the feet that define the metrical structure of words belong to a small inventory. Among the problems that this assumption is responsible for is the all but total neglect of unbounded feet, which play a major role in many of the stress systems examined in the studies cited above. Instead of taking Hayes's inventory of feet as primitives of the metrical system, the Idsardi alternative constructs feet by projecting the elements in the phoneme sequence that are capable of bearing stress onto a special stress plane. In the unmarked case all vowels are projected onto the metrical plane, whereas none of the consonants appears there. The elements projected onto the metrical plane are then chunked into feet by inserting foot boundaries at appropriate points in the sequence. Feet on this conception are composed not of syllables or moras, but of the stressable elements in the sequence which are projected on a plane that is orthogonal to the planes of syllable structure and morphological structure.

While in most languages, as noted, all vowels are stress-bearing, there are languages – and Hixkaryana is one of them – where that is not the case. In Hixkaryana, vowels in absolute word-final position are never stress-bearing. This fact is captured formally by the rule (25a), which has the effect rendering word-final vowels invisible to the metrical rules of the language, since these apply exclusively to elements projected onto the metrical plane. (In Hixkaryana the final segment of the word is never stressed even if it is a syllable head. We capture this fact by not projecting this segment on line 0 of the stress plane. Hixkaryana does have words of one closed syllable, e.g., (7); otherwise all words are vowel final.)

(25) a. Project nonfinal syllable heads

Hixkaryana moreover is subject to three footing rules, each of which inserts a parenthesis or foot boundary.

(25) b. Edge Marking:
LLL, Insert a *left* parenthesis to the *left* of the *leftmost* grid mark,
that is, Ø → (/ # _ *
c. Heavy Syllable Marking:
R, Insert a *right* parenthesis after a heavy syllable,
that is, Ø →) / * _

$$\begin{array}{cc} x & x \\ \setminus & / \\ \multicolumn{2}{c}{R} \end{array}$$

d. Iterative Footing:
R, L to R, Insert a *right* parenthesis from *right to left* for each grid mark pair
that is Ø →) / * * _

Rule (25d) assigns binary foot structure to a string. The deviations from this basic binary structure are due to the fact that the rules (25b, c) apply first and insert foot boundaries, which must be respected by (25d). We emphasize once again that footing is the result of boundary insertions into sequences of stressable elements, it is not the result of the imposition of structure on strings of phonemes.

Rule (25e) makes the rightmost element in each foot its head, and it is this head element to which stress (high tone) is assigned. Rule (25f) lengthens open syllables that are heads of feet.

(25) e. Head: R
f. Open syllable lengthening: 0 → x / * line 1
* line 0

$$\begin{array}{cc} x & \_ \\ \setminus & / \\ \multicolumn{2}{c}{R} \end{array}$$

This brings us back to Kager's objection (24). It is readily seen that the objection does not hold given the analysis in (25). No special pleading is required for disyllabic words. In such words only the nonfinal syllable is stressable and hence projected onto the metrical plane. Rule (25b) inserts a left foot boundary before the stressable element, which in this instance is also the head of the foot and hence subject to lengthening by (25f). This straightforward treatment is to be compared to Kager's OT analysis where the effect is a conspiracy between GrWrd=PrWrd and Align Left.

The ordering of the rules in (25) is fixed by UG, and is not subject to language particular specification. Some derivations are given in (26). There is no lengthening in the prefinal syllable of (26c) because the syllable is closed.

(26)

	a. /acowowo/	b. /akmatarɨ/	c. /khananɨhno/
(25a, b, c)	(* * * acowowo	(*) * * akmatarɨ	(* * *) k^hananɨhno
(25d)	(* *) * acowowo	(* * *) akmatarɨ	(* *)*) k^hananɨhno
(25e,f)	* (* *) * aco:wowo	* * (*) * *) akmata:rI	* * (* *) *) khana:nɨhno

Notice that the rules of (25) correctly generate the stresses in (26c), the problematic case in the OT account. Iterative Footing is prevented from "looking ahead" and therefore cannot see that a "canonical iamb" is available further on in the form. The strict left-to-right parsing in the Halle-Idsardi theory prevents the L(LH)L parsing from being available, and as noted above, we know of no languages that employ it. The ability to pick the best alternative, even if it requires "looking ahead" is one of the foundational assumptions of OT. OT is designed to provide explanations for problematic cases where derivations seem to require looking ahead in order to choose the correct alternative. We see from this case that there are also cases where it is detrimental to look ahead, and given the structure of OT these will be difficult to explain. Thus, all "best-first" strategies, including Harmonic Serialism (Prince and Smolensky 1993) are doomed to predict nonexistent systems for such cases, as they will allow the calculation to "look ahead" in the word to find the "best" available foot (the lurking canonical iamb). That languages do not use such parsings is a clear indication that parsing is done in a strict left-to-right fashion, and does not follow "best-first" principles.

In addition, the lack of choices available to the learner makes the rule-based framework more easily learnable than the OT alternative. The parsing of #LLH... words follows directly from the rules in (25), which can be determined without reference to this particular class of words. That is, in the rule-based analysis, the hard cases follow from the easy ones. In the OT analysis, the hard cases require recourse to additional constraints, even additional types of constraints (in this case, conjoined constraints and/or positional faithfulness constraints). In OT, the analysis of the hard cases thus does not follow from the analysis of the easy ones. In fact, given the BCD, the hard cases will have to be examined explicitly by the child, and corrective measures taken. Thus, we should expect languages to drift away from the Hixkaryana pattern of stressing #LLH... words, into the L(LH) pattern, a development similarly unattested from the languages of the world.

Furthermore, the imposition of GrWrd=PrWd, Align Left and other constraints at the top of the ranking is itself a stipulation. These constraints could be ranked elsewhere, in which case it would be possible to extend the "unstressable word syndrome" (Hayes 1995) to *bisyllabic* words, an unattested pattern in human languages. This is not possible in the analysis of (25) where either Edge Marking or Iterative Footing will provide foot structure to bisyllabic words. On general grounds of parsimony, the simpler explanation is to be preferred, and in this case the simplest explanation is given by (25).

In general, one aspect that *is* different between the two types of accounts is the number of different statements that must be invoked, and the degree of redundancy between the statements. The rule-based grammars require fewer statements with less overlap between the individual statements than does the OT account. Thus, as pointed out above, the OT grammars contain many cases where two constraints have significant (even complete) overlap in their coverage. For example, Align-Left and All-Ft-Left both penalize forms in which the first syllable is not parsed, WSP and Parse-Syl both penalize unparsed heavy syllables, WSP and RhType=I both penalize (LˈH), and everything that violates Dep-(H) also violates Dep-IO. The situation in OT gets worse still when we consider the full range of proposed universal constraints, which include SWP (stress to weight, p. 268) and *Clash (p. 165). For example, *Clash adds another reason to have binary feet, and SWP does the work orginally imputed to RhContour. So the combination of a universal constraint set and the BCD requires OT grammars to be inefficient and conspiratorial.

6. Conclusions

In this article, we have examined in some detail Kager's OT analysis of the Hixkaryana data and we have found it to be flawed in a great many respects. For example, as shown in section 2, Kager's analysis fails to deal correctly with words beginning with #LLH..., or as shown in section 3, Kager's Uneven-Iamb constraint does not achieve the effects that he attributes to it, et cetera. We have also shown that in order to repair these inadequacies in the analysis, the number of constraints actively involved must be radically increased. The direct effect of this increase is an increase in the indeterminacy of the grammar: there are now several alternative constraint rankings that relate a given underlying form to a particular surface string. As a result, one analysis is as good as several others and nothing in the theory chooses between them. In sum, the empirically-adequate OT account of the Hixkaryana facts raises severe conceptual problems.

In section 5 we examined Kager's objections to a rule-based account of the Hixkaryana data. We showed that two of his objections were invalid in that they held with equal force against the OT analysis. Kager's first objection involved the lack of intrinsic connection between stress and length in rule-based accounts. Kager's second objection concerned the fact rule-based analyses miss a generalization by allowing light syllables to constitute feet in intermediate representations. Since OT accounts are demonstrably subject to the same two objections, these objections are without force. Kager's third objection focused on the special treatment of bisyllabic words in a rule-based account which employed Hayes's inventory of feet. It was shown above that no special treatment is involved in an alternative rule-based account that, instead of Hayes's foot inventory adopts Idsardi's theory of metrical structure.

In sum, Kager's claims in favor of OT are invalid on purely empirical grounds: his analysis of Hixkaryana stress in (2) fails to deal with at least one of the ten examples in his data set (1). And when his analysis is repaired to conform with the data and the requirements of the BCD, the result is a cumbersome and unenlightening account involving twenty different (and yet overlapping) constraints (see (20)). Kager's objections to rule-based accounts are also invalid because the same objections hold with equal force against OT. Since Kager has failed to provide even the most minimal arguments for choosing OT over a rule-based account, the supposed superiority of OT to rule-based accounts remains a fiction.

MIT (MH)
University of Delaware (WJI)

References

Beckman, Jill N. (1997). Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14, 1: 1–46.

Bell, Alan (1977). Accent placement and perception of prominence in rhythmic structure. In Hyman, Larry M. (ed.) *Studies in Stress and Accent* (= Southern California Occasional Papers in Linguistics 4). Los Angeles: University of Southern California. 1–13.

Cohn, Abigail and John J. McCarthy (1994). Alignment and parallelism in Indonesian phonology. Ms. Cornell University and University of Massachusetts/Amherst.

Derbyshire, Desmond C. (1985). *Hixkaryana and Linguistic Typology*. Dallas: Summer Institute of Linguistics.

Haas, Mary (1977). Tonal accent in Creek. In Hyman, Larry M. (ed.), *Studies in Stress and Accent* (= Southern California Occasional Papers in Linguistics 4). Los Angeles: University of Southern California. 195–208.

Halle, Morris (1997). On stress and accent in Indo-European. *Language* 73, 2: 275–313.

— (1998). The stress of English words. *Linguistic Inquiry* 29, 4: 539–568.

Halle, Morris and William J. Idsardi (1995). General properties of stress and metrical structure. In John Goldsmith, John (ed.) *The Handbook of Phonological Theory*. Cambridge MA: Blackwell. 403–443.

— (1998). A response to Alan Prince's letter in issue 2–6. *GLOT International* 3, 1: 1–22.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Idsardi, William J. (1992). The computation of prosody. PhD. thesis, MIT.

— (1995). Grammar delicacy in stress theory. Talk presented at the University of Maryland Mayfest, College Park, Maryland, May (1995).

Kager, Rene (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Leer, Jeff (1985). Prosody in Alutiiq. In Krauss, Miachel (ed.) *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies*. Alaska Native Language Center, Fairbanks, Alaska. 25–45.

Parker, Steve (1994). Coda Epenthesis in Huariapano. *International Journal of American Linguistics*, 60, 2: 95–119.

Prince, Alan (1975). The phonology and morphology of Tiberian Hebrew. PhD. thesis, MIT.

Prince, Alan and Paul Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. Technical report, Center for Cognitive Science, Rutgers University, New Brunswick, N.J. and Computer Science Department, University of Colarado, Boulder.

Prince, Alan and Bruce Tesar (1999). Learning phonotactic distributions. Ms. Rutgers University.

Purnell, Thomas C. (1998). Principles and parameters of phonological rules: Evidence from tone languages. PhD. thesis, University of Delaware.

Rappaport, Malka (1984). Issues in the phonology of Tiberian Hebrew. PhD. thesis, MIT.

Sohn, Ho-Min (1999). *The Korean Language*. Cambridge: Cambridge University Press.

Tesar, Bruce, and Paul Smolensky (1998). The learnability of Optimality Theory. *Linguistic Inquiry* 29, 2: 229–268.