Iterative foot optimization and locality in stress systems[1]

Kathryn Pruitt
University of Massachusetts Amherst

Abstract. This paper proposes a model of stress assignment in which metrical structure is built serially – one foot at a time – through a series of OT-style evaluations. Iterative foot optimization (IFO) is made possible in the framework of Harmonic Serialism, which defines the path from an input to an output with a series of gradual changes in which each form is more harmonic than its predecessor, relative to a constraint ranking. Stress assignment in IFO is compared to parallel OT and it is found that they predict different classes of languages even when the same standard stress constraints are considered. IFO makes the strong prediction that decisions about metrical structure are made locally, while parallel OT predicts stress systems with non-local interactions. The interactions of stress with syllable weight, vowel shortening, and edge restrictions are considered, and in all cases it is shown that attested languages exhibit local interactions, while parallel OT predicts non-local counterparts which are not clearly attested.

1 Introduction

Harmonic Serialism (HS) is a variant of Optimality Theory that combines optimization with a derivation. HS was discussed briefly by Prince and Smolensky (1993/2004:6-7, 94-96) but was at the time put aside in favor of now-standard parallel Optimality Theory (henceforth "parallel OT" or simply "OT"). McCarthy (2000, 2006, 2007ab, to appear-ab) has recently demonstrated HS's potential to address some of parallel OT's well-known problems and has shown that the typological predictions of HS – that is, the range of phonological grammars that it predicts – may differ importantly from those of parallel OT.[2,3]

[1] Many thanks are due to John McCarthy and Joe Pater for help on this work. Thanks also to the members of the fall 2007 3rd-year seminar at UMass Amherst and the audience of HUMDRUM 2008 at Rutgers for helpful comments and suggestions.

[2] McCarthy (2007a) proposes a variant of HS called Optimality Theory with Candidate Chains (OT-CC) as a way of modeling phonological opacity, which HS alone handles no better than OT. I will continue to refer to the framework I adopt in this paper as HS, though the claims in this paper are largely consistent with the additional architectural assumptions of OT-CC.

[3] See also Pater (2007, 2008ab), who suggests adopting HS in a framework with constraint weighting (rather than ranking), as well as Wolf (2008), who proposes a theory of the phonology-morphology interface based on OT-CC (McCarthy 2007a), and Elfner (2008), Jesney (2008), and Kimper (2008) who adopt variants of HS to deal with prosodic circumscription, positional faithfulness, and local optionality, respectively. See Jacobs (2008) for an application of OT-CC to stress.

This paper proposes a HS-based model of stress assignment, in which foot-placement is determined by a series of optimizations, and finds that its typological predictions indeed differ significantly from those of parallel OT when the same constraints are used. Importantly, a class of largely unattested grammars is predicted by the parallel OT analysis of stress using standard stress constraints from Prince and Smolensky (1993/2004) and McCarthy and Prince (1993a), and this class of grammars is correctly predicted not to exist by the HS-based model. The common feature of these unattested stress systems is that they contain non-local interactions between stress/foot placement and other phonological properties and processes in the word. Parallel OT makes these incorrectly non-local predictions about what can and cannot affect stress assignment because of its ability to parse a word into feet in one swoop. The present model prevents these non-local interactions by proceeding iteratively with foot-building, effectively limiting the properties of the word that can affect foot placement at each iteration. This paper thus concludes that HS with serial foot-building provides clear advantages over parallel OT in the domain of rhythmic stress.

The organization is as follows: In section 2 I discuss the details of the present proposal and illustrate how stress is assigned in HS with these assumptions. In section 3 we turn to cases in which HS and parallel OT have different predictions – in all cases it will be shown that HS is superior to parallel OT as a model of stress and its interactions with other processes because it effectively enforces a locality restriction on foot placement decisions, and this has positive typological consequences. In section 4 I discuss some formal consequences of adopting this view of stress assignment, particularly for the proper constraints for describing metrification. Section 5 concludes.

2 Iterative foot optimization

This section outlines the specifics of the present proposal for stress assignment, in which word stress is assigned by iteratively building the best foot with a series of OT-style optimizations. I will refer to this proposal as HS with iterative foot optimization (HS/IFO or IFO). The substantive assumptions for HS/IFO are laid out below. First, some background on HS is provided, followed by an explanation the assumptions of IFO and an illustration of how it works. This section concludes with a brief look at how HS/IFO compares with two of its predecessors in serial stress assignment.[4]

Before moving on, the sense in which I mean 'iterative' should be made clear. In many rule-based theories of stress, "iterative" refers to the directional application of rules that build metrical structure. This is often distinguished from non-iterative parsing, in which a stress rule applies only once, yielding non-rhythmic stress (Hayes 1995: 113). Here, "iterative" simply refers to the fact that in HS, OT-style evaluations occur repeatedly until convergence (as defined below). The entire grammar is iterative in this sense, and I will refer to a particular derivational step or evaluation as an "iteration." Not coincidentally I will also be primarily concerned with stress systems which are iterative in the traditional sense, but the specifics of IFO will also extend to other stress systems, for which the constraint ranking, rather than a binary parameter, determines whether multiple feet are built.

2.1 Harmonic Serialism

HS shares with parallel OT several trademark features, including the characterization of a language as a ranking of some presumed universal set of constraints, but it departs from parallel OT in its assumptions about how this ranking is accessed as an input becomes an output. In parallel OT any and all mutations of an input are performed in parallel and the resulting candidates are evaluated once with a single winner chosen relative to the ranking of the constraints. In contrast, an input to a HS grammar is gradually altered to reach an output, with each step in the derivation chosen by optimization over a set of alternatives derived by "doing one thing" to the previous iteration's output (Prince and Smolensky 1993/2004: 6). The same ranking is used at each iteration and the derivation terminates when the input to an iteration is returned as its own best output, as this indicates that no additional changes result in a more harmonic output relative to the constraint hierarchy. An illustration of this basic difference between HS and parallel OT is shown in (1) and (2).

[4] For more extensive background on HS, see the references cited in the introduction.

(1) Parallel OT: /in/ → [out]

One step, where /in/ and [out] may differ in any and all ways.

(2) HS: /in/ → int_1 → int_2 → … → [out]

Series of steps, where there may be intermediate forms, and each form differs by exactly one change from, and is more harmonic than, its predecessor.

In formal terms, CON and EVAL work the same in HS and parallel OT, while GEN and the relationship between GEN and EVAL are different. The GEN of HS does not have the 'freedom of analysis' which is familiar from parallel OT's GEN (Prince and Smolensky 1993/2004). Instead, it only provides candidates differing in at most one respect from the input at each iteration. What counts as one difference is a matter for investigation, an enterprise to which this paper is intended to contribute. The GEN-EVAL relationship also differs in HS, in that GEN and EVAL are in a loop until convergence: GEN produces a set of single-change candidates, EVAL chooses the most optimal one relative to the constraint hierarchy of the language, and this intermediate output is passed back to GEN for another iteration in which all single-change candidates from this intermediate input are computed and then compared.

Because of the nature of optimization, it is true in HS as well as in parallel OT that any form that wins over a faithful candidate is less marked with respect to the constraint hierarchy. This is because a form which violates a faithfulness constraint can only win when it better satisfies a higher ranked markedness constraint. The result is that derivations are always harmonically improving. Each form is more harmonic than its predecessor and less harmonic than any subsequent form; this is guaranteed by the architecture of the system.

The winner at any of the iterations is the *locally* optimal candidate (terminology following McCarthy 2007a, et seq.); it wins when compared to the members of a limited candidate set, which includes only the input to that iteration and forms that are one change away from the input. Each intermediate form in a derivation is a local optimum, and when one iteration's local optimum is returned again in the following iteration, the derivation converges because no other single change is harmonically improving.

Parallel OT, in contrast, evaluates in one swoop candidates which differ in any number of ways from the input, and thus always finds the *global* optimum relative to that input and constraint ranking. Sometimes a HS derivation converges at a global optimum, but sometimes not, and in the latter case we find that the predictions of parallel OT and HS may differ even when the same constraint ranking is employed. Failure to reach the hypothetical global optimum in a HS derivation occurs when it is not reachable in a series of gradual, harmonically improving steps, however one step has been defined. Importantly, as we will see in this paper and has been shown in other recent work (e.g., McCarthy to appear-b), parallel OT's global optimum may not be the typologically desirous output. Thus, HS's failure to reach such a candidate is not a failure at all.

2.2 Gradualness in IFO

Because of the requirement that GEN produce only candidates with one difference relative to the input, derivations are gradual. However, it is necessary to assume some theory of gradualness (that is, of GEN) in order to make claims about the predictions of HS, and the present case is concerned with what might constitute a gradual way to build metrical structure. I will adopt the assumption that gradualness in the domain of metrical structure-building is instantiated by construing 'one difference' as the addition of one headed (that is, stressed) metrical foot. Thus, at each iteration GEN produces candidates corresponding to all possible ways of adding one foot to that input (in addition to candidates representing other kinds of single changes). EVAL selects among these candidates based on the constraint hierarchy, and a stress derivation proceeds by building the 'best' next foot each time.

For the purposes of this paper I will follow most work in metrical theory in assuming that predictable stress (our focus here) is not underlying, and the job of the grammar is to account for stress patterns by building metrical feet from scratch. The relationship between metrical structure and Richness of the Base (Prince and Smolensky 1993/2004) is complicated by moving to a derivational theory, as the question of whether it is feasible to gradually remove unwanted structure from hypothetical input forms becomes important. Interesting though it is, this question is left to future work. As we will see, adopting this simplifying assumption allows us to easily compare HS/IFO and parallel OT methods of foot-building.

2.2.1 Additional assumptions

I will also assume that GEN is restricted to producing candidates with feet that are maximally disyllabic and have exactly one syllable designated as a head.[5] An exhaustive list of the candidates for the first iteration of stress assignment in a five syllable word is shown in (3). The candidate set includes the faithful candidate (in bold), while the other candidates exemplify all possible ways of building a single foot given these assumptions: the five possible monosyllabic feet, the four possible trochaic (left-headed) feet, and the four possible iambic (right-headed) feet. The ranking of markedness constraints on metrical structure will determine which of these is optimal, since we are concentrating on cases in which stress is not present in the original input.

(3) Candidates for first foot in a five-syllable word

σσσσσ	(ˈσ)σσσσ	σ(ˈσ)σσσ	σσ(ˈσ)σσ	σσσ(ˈσ)σ
σσσσ(ˈσ)	(ˈσσ)σσσ	σ(ˈσσ)σσ	σσ(ˈσσ)σ	σσσ(ˈσσ)
(σˈσ)σσσ σ	(σˈσ)σσ	σσ(σˈσ)σ	σσσ(σˈσ)	

A further assumption adopted in this paper is that foot structure can be built by GEN, but not altered or removed. Any foot that is built in the course of deriving stress is inherited by every member of the candidate set for subsequent iterations, an assumption that I will call *strict inheritance*. Thus, any foot structure chosen as optimal at any iteration is kept throughout the derivation, and the derivation is monotonic, adding exactly one foot each time until convergence. One consequence of strict inheritance is that subsequent foot building can only parse 'free' syllables, i.e., those that are not yet in a foot. This requirement is familiar from Prince (1985) as the Free Element Condition, given in (4). Similar notions have been adopted and argued for in other work (e.g., Steriade 1988).

---

[5] The disyllabic maximum may eventually need to be relaxed for unbounded or ternary stress systems, but it will suffice as a restriction on GEN for the analyses in this paper.

(4) Free Element Condition (FEC; Prince 1985:479)

Rules of primary metrical analysis apply only to Free Elements – those that do not stand in the metrical relationship being established.

To give an example of this assumption at work, if the candidate with a left-aligned disyllabic trochee, (ˈσσ)σσσ, is the most harmonic candidate among all those in (3) at the first iteration, then the set of candidates for the second iteration will be those shown in (5). This set includes a faithful candidate (again in bold), which inherits the structure from the previous output but adds no more, and it also includes all the candidates derivable from this input by building one licit foot on its remaining free syllables, since, by the assumption of strict inheritance, subsequent iterations are barred from altering previous feet in any way. By assumption we also impose a ban against the incorporation of free syllables into previously built feet. Thus, if (ˈσ)σσσσ emerges as the winner in the first iteration, (ˈσσ)σσσ would not be a candidate in the second iteration.

(5) Candidates for second foot in a five-syllable word, given input (ˈσσ)σσσ

(ˈσσ)σσσ	(ˈσσ)(ˈσ)σσ	(ˈσσ)σ(ˈσ)σ	(ˈσσ)σσ(ˈσ)
(ˈσσ)(ˈσσ)σ	(ˈσσ)σ(ˈσσ)	(ˈσσ)(σˈσ)σ	(ˈσσ)σ(σˈσ)

Given the assumptions laid out in this section, (6) illustrates examples of derivations that are not permitted. The derivation in (a) is not allowed because it is insufficiently gradual. GEN does not produce candidates with more than one foot added in a single step. The derivations in (b) through (f) are not permitted because the form on the right in each case violates strict inheritance with respect to the preceding forms. The derivations in (b) and (c) violate the subtype of strict inheritance covered by the FEC. The cases in (d)-(f) show that strict inheritance is sometimes a redundant assumption, because the last step in these derivations would not be harmonically improving even if GEN permitted them as candidates. For instance, in (d), if input /σσσσσ/ becomes (ˈσσ)σσσ, then (ˈσσ)σσσ must have been more harmonic than its competitors, including σσσσσ, at the first step. Thus, a subsequent iteration accessing the same constraint ranking could not possibly judge σσσσσ to be more harmonic than (ˈσσ)σσσ, assuming all else is equal. The same goes for the derivations in (e) and (f). The FEC violations in (b) and (c), on the other hand, could potentially result in harmonic improvement, and in these cases the assumption is not redundant.[6]

(6) Illicit derivations

Hypothetical derivation	Reason disallowed
a. σσσσσ ↛ (ˈσσ)(ˈσσ)σ	Insufficiently gradual
b. (ˈσσ)σσσ ↛ (ˈσ)(ˈσσ)σσ	Strict inheritance / FEC
c. (ˈσσ)σσσ ↛ (ˈσ)(σˈσ)σσ	Strict inheritance / FEC
d. σσσσσ → (ˈσσ)σσσ ↛ σσσσσ	Strict inheritance / (Not improving)
e. σσσσσ → (ˈσσ)σσσ ↛ (ˈσ)σσσσ	Strict inheritance / (Not improving)
f. σσσσσ → (ˈσσ)σσσ ↛ (σˈσ)σσσ	Strict inheritance / (Not improving)

2.2.2 Summary

In sum, the proposal here is that feet are built through a series of optimizations which always choose the best foot to build. GEN produces candidates with feet that are mono-syllabic or disyllabic, with exactly one syllable designated as head. The additional assumption of strict inheritance dictates that which syllables constitute a foot's membership and which syllable is designated as head are not alterable in the course of deriving word stress.[7]

Importantly, while deriving stress patterns is the topic of this paper, other changes which are sufficiently gradual will also compete at each foot-building step when the entire grammar is set into motion. That is, at each iteration, both the best *kind* of change is considered (whether to delete a segment, or to build a metrical foot, for example), and also the best *instance* of the best change (e.g., the *best* way to delete one segment/feature, or the *best* way to build one foot). For the purposes of exploring the consequences of the present proposal, I will generally only consider steps in which building a foot is the best possible kind of change, and I will not include in my illustrations candidates which have

[6] Strict inheritance also becomes less redundant if other processes, such as deletion, can be interleaved with stress assignment. I leave to future investigation the question of whether the predictions of strict inheritance are correct in these cases, and simply adopt the assumption here.

[7] It is conceivable that morphologically complex words will appear to require metrical reanalysis in the course of the derivation, but this is non-problematic. Either we assume that morphology is cyclic and that post-lexical strata can treat faithfulness to the lexical phonology's input as a violable constraint, or we assume that we have access to the morphology right away, and thus build metrical structure only once with access to morpheme boundaries which may or may not play a role in stress assignment.

undergone other operations, unless directly relevant. One should bear in mind that the presence of only stress-addition candidates in these illustrations may be considered an abstraction.[8]

2.3 Illustration

In this section I illustrate the workings of HS/IFO with a stress derivation for the language Pintupi (Hansen and Hansen 1969, 1978), which has been analyzed as having a quantity-insensitive trochaic stress system (Hayes 1995: 62). Quantity-insensitive stress systems do not respect syllable weight, either because vowel length is not contrastive in the language or because the language simply ignores it. Pintupi falls into the latter category – vowel length is indeed contrastive, but stress appears to be indifferent to it: main stress is initial, with secondary stresses falling on odd-numbered non-final heavy syllables. I analyze Pintupi with left-aligning (that is, left-to-right) syllabic trochees, adapting the standard analysis from Hayes (1995).

The constraints employed to describe this and subsequent stress systems are familiar constraints for defining metrical structure in OT. First, we will assume the constraints PARSESYLL, defined in (7), and ALLFTL/R, defined in (8). PARSESYLL provides the impetus for foot-building, while gradiently defined ALLFTL/R, from the generalized alignment family (McCarthy and Prince 1993a), prefer feet to be aligned as far as possible to the edge of the word, simulating directionality. These constraints are defined exactly as they usually are in parallel OT analyses, though they will ultimately result in different predictions because of the difference between parallel and serial evaluation.

(7) **PARSESYLL**: Assign one violation mark for each syllable that is not a member of some foot.

(8) **ALLFTL/R**: For each foot in a word assign one violation mark for every syllable separating it from the left/right edge of the word.

[8] Choosing the best kind of change at each iteration is a property that is not directly encoded in the HS-variant OT-CC (McCarthy 2007a). In OT-CC the grammar only chooses the best order of changes indirectly, by allowing outputs of chains (derivations) to compete. Whether process competition at each iteration indeed exists is an open question, though I have assumed here that it does.

In order to enforce left-headed feet (i.e., trochees) over right-headed feet (i.e., iambs), the constraints in (9) will be necessary. These are equivalent to the RHTYPE constraints of Prince and Smolensky (1993/2004:63). Monosyllabic feet are assumed to satisfy both constraints.

(9) **IAMB/TROCHEE**: Assign one violation mark for a foot whose head is not aligned with the right/left edge of the foot.

Finally, to account for the strict disyllabicity of feet in Pintupi, a constraint preferring disyllabic feet is necessary. In OT, foot binarity constraints are generally called upon to dictate a minimum foot size, and I adopt this strategy as well. However, FTBIN as standardly defined ("Feet are binary under syllabic or moraic analysis"; Prince and Smolensky 1993/2004: 56) effectively puts a bimoraic minimum on feet and cannot force disyllabic feet. Therefore, I will assume following Hewitt (1994) that standard FTBIN is actually two constraints, one preferring feet with at least two syllables, as in (10), and the other preferring feet with at least to moras, as in (11). Not surprisingly, FTBIN(μ) assigns violation marks exactly like traditional FTBIN. We might then view this proposal as simply adding the constraint FTBIN(σ).

(10) **FTBIN(σ)**: Assign one violation mark for a foot with fewer than two syllables.

(11) **FTBIN(μ)**: Assign one violation mark for a foot with fewer than two moras.

For a syllabic trochee stress pattern these constraints must be ranked so that a five-syllable input /σσσσσ/ is ultimately parsed as (ˈσσ)(ˈσσ)σ in a series of gradual foot-building steps.[9] I will assume the derivation proceeds as shown in (12) with a five-syllable word from Pintupi. At the first foot-building step a left-aligned disyllabic trochee must be optimal, such that *(ˈpuɭiŋ)kalatʲu* wins. At the second step, another disyllabic trochee is built adjacent to the first, yielding *(ˈpuɭiŋ)(ˈkala)tʲu*. A subsequent step will signal convergence when this form is again returned as the optimal candidate, which indicates that additional foot-building does not improve harmony.

[9] I ignore degrees of stress.

(12) Syllabic trochee derivation for Pintupi *puɭiŋkalatʲu* ‘we (sat) on a hill’

puɭiŋkalatʲu → (ˈpuɭiŋ)kalatʲu → (ˈpuɭiŋ)(ˈkala)tʲu

In the first step, a left-aligned disyllabic trochee beats every other possible candidate (see (3) for an exhaustive list). The tableau in (13) shows this step and indicates which rankings must obtain to achieve this outcome. This and subsequent tableaux in this paper are in a modified comparative format (Prince 2002). The intended winner in (b) is indicated with an arrow, and the numbers in the cells correspond to the number of violation marks incurred by that candidate on that constraint (replacing the familiar *’s). The violation profiles for each of the losing candidates may also include a W or an L. A W in a cell indicates that the winning candidate, in this case (b), satisfies the constraint in that column better than the loser represented by that row (i.e., the constraint is “winner-favoring” in that comparison); an L indicates the winning candidate does worse (has more violations) than the loser on that constraint (i.e., the constraint is “loser-favoring” in that comparison). If the winner and loser receive the same number of violation marks for a particular constraint, neither a W nor an L is indicated. Ranking arguments can easily be made on the basis of the location of the W’s and L’s. For the intended winner to win, every L in the tableau must be preceded in the same row by a W. This follows the familiar requirement that the highest-ranked constraint which can distinguish between a winner and a loser must favor the winner. Thus, returning to our example in (13), ALLFTL must dominate ALLFTR according to rows (c), (d), and (e), and TROCHEE must dominate IAMB according to row (f). PARSESYLL must also be high enough ranked to compel building this foot even though doing so causes violations of IAMB and ALLFTR, as row (a) indicates.

(13) Ranking arguments from 1st iteration:

PARSESYLL >> IAMB, ALLFTR; ALLFTL >> ALLFTR; TROCHEE >> IAMB

/puɭiŋkalatʲu/ ‘we (sat) on a hill’ **1st iteration**	PARSESYLL	TROCHEE	ALLFTL	IAMB	ALLFTR
a. puɭiŋkalatʲu	**W** 5			**L**	**L**
→ b. (ˈpuɭiŋ)kalatʲu	3			1	3
c. pu(ˈɭiŋka)latʲu	3		**W** 1	1	**L** 2
d. puɭiŋ(ˈkala)tʲu	3		**W** 2	1	**L** 1
e. puɭiŋka(ˈlatʲu)	3		**W** 3	1	**L**
f. (puˈɭiŋ)kalatʲu	3	**W** 1		**L**	3

To complete the analysis it is necessary to continue through the derivation and confirm that the rankings we need at the first iteration are consistent with those required at subsequent iterations. Additional ranking arguments can also be made as the derivation proceeds. At the second iteration, *(ˈpuɭiŋ)(ˈkala)tʲu* wins over its competitors. As row (a) of the tableau in (14) shows, this requires that one additional ranking be assumed: PARSESYLL >> ALLFTL. A familiar requirement from stress analyses in parallel OT is that this ranking must hold to get rhythmic stress; otherwise, a single foot remains at the left edge and additional feet are not built (McCarthy and Prince 1993a). The same holds in HS/IFO as well.

(14) Ranking argument from 2nd iteration: PARSESYLL >> ALLFTL[10]

⟨(ˈpuɭiŋ)kalatʲu⟩ **2nd iteration**	PARSESYLL	TROCHEE	ALLFTL	IAMB	ALLFTR
a. (ˈpuɭiŋ)kalatʲu	**W** 3		**L**	**L** 1	**L** 3
→ b. (ˈpuɭiŋ)(ˈkala)tʲu	1		2	2	4
c. (ˈpuɭiŋ)ka(ˈlatʲu)	1		**W** 3	2	**L** 3
d. (ˈpuɭiŋ)(kaˈla)tʲu	1	**W** 1	2	**L** 1	4

The third iteration requires that *(ˈpuɭiŋ)(ˈkala)tʲu* again be returned as the best output, indicating that additional foot-building is not harmonically improving, and signaling convergence. Because the input to this iteration, *(ˈpuɭiŋ)(ˈkala)tʲu*, contains only one free syllable, there is only one other candidate for adding foot structure,

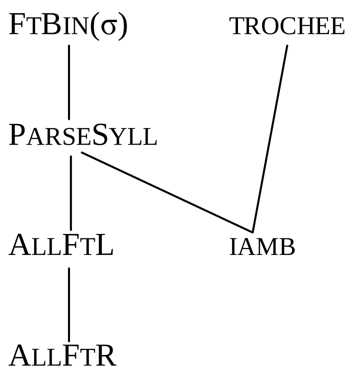
[10] I will use angled brackets ⟨ and ⟩ to mark intermediate forms in a derivation when they appear in the input cell of a tableau to avoid confusion with / and / which are reserved for original phonological inputs.

(ˈpuɭiŋ)(ˈkala)(ˈtʲu). This candidate loses on the grounds that it contains a monosyllabic foot, even though it features perfect satisfaction of PARSESYLL. Thus, the ranking FTBIN(σ) >> PARSESYLL is required, as shown by row (b) in (15). The only remaining constraint among those we defined in (7) through (11) is FTBIN(μ), which does not crucially decide any winners. The diagram in (16) illustrates the Pintupi ranking.

(15) Ranking argument from 3rd iteration

⟨(ˈpuɭiŋ)(ˈkala)tʲu⟩ **3rd iteration**	FTBIN(σ)	PARSESYLL
→ a. (ˈpuɭiŋ)(ˈkala)tʲu		1
b. (ˈpuɭiŋ)(ˈkala)(ˈtʲu)	**W** 1	**L**

(16) Pintupi ranking (left-to-right syllabic trochees)



In (17) is a summary tableau that shows the full derivation of this form with this ranking in place. Candidate (b) wins in the first iteration because it satisfies TROCHEE and FTBIN(σ), and is left-aligned to the edge of the word in accordance with ALLFTL. This form is passed to the second iteration where it appears in (e) as the faithful candidate. At this pass, candidate (f) is chosen as optimal because it adds a foot of the proper form (disyllabic trochee), satisfies PARSESYLL better than not adding a foot, and best satisfies ALLFTL among the remaining candidates. This form is then passed to the third iteration, at which point the faithful candidate in (i) is compared to the only remaining stress candidate, which has added a monosyllabic foot on its final syllable. The fully parsed candidate fatally violates FTBIN(σ), which outranks PARSESYLL, and thus the derivation converges by choosing candidate (i), the third iteration's input.

(17) Pintupi derivation summary for input /puɭiŋkalatʲu/[11]

/puɭiŋkalatʲu/ ‘we (sat) on the hill’ **1st iteration**	FTBIN(σ)	TROCHEE	PARSESYLL	ALLFTL	IAMB	ALLFTR
a. puɭiŋkalatʲu			**W** 5		**L**	**L**
• b. (ˈpuɭiŋ)kalatʲu			3		1	3
c. puɭiŋka(ˈlatʲu)			3	**W** 3	1	**L**
d. (puˈɭiŋ)kalatʲu		**W** 1	3		**L**	3
2nd iteration						
e. (ˈpuɭiŋ)kalatʲu			**W** 3	**L**	**L** 1	**L** 3
• f. (ˈpuɭiŋ)(ˈkala)tʲu			1	2	2	4
g. (ˈpuɭiŋ)ka(ˈlatʲu)			1	**W** 3	2	**L** 3
h. (ˈpuɭiŋ)(kaˈla)tʲu		**W** 1	1	2	**L** 1	4
3rd iteration						
• i. (ˈpuɭiŋ)(ˈkala)tʲu			1	2	2	4
j. (ˈpuɭiŋ)(ˈkala)(ˈtʲu)	**W** 1		**L**	**W** 6	2	4

Output: (ˈpuɭiŋ)(ˈkala)tʲu

This example shows how familiar OT constraints combined with the architecture of HS and the assumptions of IFO derive a simple stress pattern such as the one exemplified by Pintupi.

2.4 Serial predecessors to HS/IFO

Before moving on it is useful to mention briefly two prominent predecessors of HS/IFO and point out the ways in which the theories differ from one another. First, rule-based stress assignment is discussed, followed by Prince’s (1990) theory of Harmonic Parsing.

2.4.1 Standard rule-based accounts

Serial, rule-based versions of metrical theory differ from one another in the various details of their execution, but they share notions of parameterization and rule-

[11] In this tableau format a winning candidate at a particular iteration is indicated by a bullet symbol that begins an arrow down to the next iteration. The faithful candidate at an iteration (the original input or the previous iteration’s winner) is listed first before others. Lettering is continuous within the derivational tableau to avoid ambiguity in referring to candidates.

ordering (see e.g., Halle and Vergnaud 1987; Kager 1989; Hayes 1995). Parameterization accounts for things like directionality and foot type, while ordering determines when foot-building should take place relative to other processes. Although HS/IFO is a serial model, it differs from these rule-based models by employing neither of these things.

Parameterization does not explicitly exist in models with optimization, which model preferences as emerging from violable constraints on well-formedness and their ranking with respect to each other and to faithfulness constraints, rather than from the setting of binary switches. This much is carried over from parallel OT into HS/IFO, as both share optimization with violable constraints as the method of determining output forms. With respect to directionality for example, although the serial parse in the previous section achieved the same result as a directional foot-building rule beginning at the left, the process was actually quite different. In HS/IFO, as in OT, directionality is emergent; the best foot is built at each iteration, and this only looks like a directional parse when the constraints are ranked to prefer contiguous feet aligning toward one edge or the other. In fact, the constraints might instead be ranked to allow non-contiguous foot building. If a constraint preferring heavy syllables to be stressed, e.g. WSP as in (18), is ranked above ALLFTL, then it will be more harmonic to build a foot that places stress on a heavy syllable than it will be to build feet contiguously from the left edge. This is illustrated in the hypothetical derivation in (19). Here, although ALLFTL is assumed to dominate ALLFTR, it cannot enforce its preference for left-aligning feet at the first iteration, because higher-ranked WSP requires heavy syllables to be stressed before anything else. Once the only heavy syllable in the input is stressed, the subsequent footing of light syllables is sensitive to ALLFTL'S preferences, and such feet are filled in by apparent left-to-right parsing.

(18) **WEIGHT-TO-STRESS (WSP)**: Assign one violation mark for every unstressed heavy syllable.

(19) Non-contiguous foot-building in HS/IFO

/LLHLL/ **1st iteration**	WSP	PARSESYLL	ALLFTL
a. (ˈLL)HLL	**W** 1	**L** 3	**L**
→ b. LL(ˈH)LL		4	2
2nd iteration			
e. LL(ˈH)LL		**W** 4	2
→ f. (ˈLL)(ˈH)LL		2	2
g. LL(ˈH)(ˈLL)		2	**W** 5
3rd iteration			
i. (ˈLL)(ˈH)LL		**W** 2	**L** 2
→ j. (ˈLL)(ˈH)(ˈLL)			5

Output: (ˈLL)(ˈH)(ˈLL)

The other feature of rule-based analyses of stress, ordering, is a non-issue in parallel OT but becomes relevant again in a discussion of how serial derivations in HS/IFO are determined. Unlike in many rule-based theories, processes are not explicitly ordered with respect to one another in HS, but instead the constraint ranking is responsible for choosing an optimal candidate and thus, indirectly, an optimal process at each iteration. The definition of constraints also plays a role in determining the ordering of processes, and can give rise to 'intrinsic ordering'. In such cases, a particular process does not result in harmonic improvement until another process has applied; we will see an example of this in section 3.2 (see also McCarthy to appear-a for another case of intrinsic ordering in HS).

These two main differences are important for setting HS/IFO apart from standard rule-based accounts. Although more work is needed, it appears that many of the results of Prince and Smolensky (1993/2004) show the need for violable constraints rather than for parallelism necessarily, allowing many of the original pro-OT arguments to also apply to HS. Thus, we retain the virtues of a model with violable constraints in HS/IFO and do not explicitly employ ordered rules and parameters in describing stress.

2.4.2 Prince (1990)'s Harmonic Parsing

Prince (1990) discusses a serial theory of metrification, which he calls Harmonic Parsing (HP), that has some similarities to the framework adopted here. One of the main contributions of Prince (1990) was to begin formalizing the notion of relative well-formedness of foot structure and to allow this to be a guiding force in foot building. Although he retains a directional mode of parsing, the notion of serially building the 'best foot' was already present in HP, and it should therefore be recognized as a significant predecessor to HS/IFO.

There are a number of differences between HP and HS/IFO however. Because HP uses directional parsing, a moving window effectively limits the number of feet it considers at once, compared to the larger set of foot candidates of HS/IFO which are not constrained by location within the word. As illustrated in the previous section this can lead to non-contiguous foot-building in HS/IFO since constraints preferring apparent directional parsing are violable. Thus, HP differs from HS/IFO with respect to directional foot-building in the same way as many traditional rule-based accounts.

The main point of similarity between HP and HS/IFO is in HP's use of a metric for determining the 'best' foot, foreshadowing notions of relative harmony and optimization. The formalization is not in terms of ranked constraints however, and leaves little room for language-particular subversions of the relative harmony of foot shapes. In addition, the metric for determining the best foot relies exclusively on foot form. In HS/IFO, as in parallel OT, violable constraints are responsible not only for foot form but also for foot placement within the word, further honing the notion of what it means to be the 'best possible' foot.

Prince (1990) stands out as a work in a similar vein as the proposal made in this paper, though the present proposal integrates iterative foot building with subsequent theoretical developments in Optimality Theory. In the next section I turn to an illustration of what sets HS/IFO apart from its immediate relative, parallel OT, in the domain of rhythmic stress. In particular, I show in detail how HS/IFO makes a distinction between local and non-local interactions with stress, allowing the former and not the latter, while parallel OT does not.

3 Locality vs. globality in stress

Adopting HS/IFO for stress assignment makes a number of predictions about the typology of stress and its interactions with syllable weight, shortening and lengthening for metrical conformity, and constraints on edges of prosodic domains. In particular, all these interactions must be local and must not require the derivation to 'look ahead'. In this section I show that in each of these cases, local interactions are attested, but non-local counterparts are not, and thus, consistent with the predictions of HS/IFO, only interactions which do not require derivational look-ahead are attested. Parallel OT, however, freely predicts non-local counterparts to each of the examples, and in each case we find that such languages are not unequivocally attested. The conclusion is that HS/IFO is to be preferred over parallel OT as a model of stress assignment.

I begin in section 3.1 with an example of an unattested non-local prediction of standard stress constraints in parallel OT first noticed by Hyde (2007) and I show that HS/IFO does not make the same errant prediction. In section 3.2 I examine the well-known process of trochaic shortening, arguing that it is found only as a local interaction; I illustrate that parallel OT predicts a non-local counterpart and that such a system is not attested. In section 3.3 I turn to the edge restriction non-finality, and show that languages respond locally to it, though parallel OT predicts a non-local counterpart in which the effects of non-finality permeate the word. Section 3.4 concludes this section with a summary of these cases and reiterates the theme of attested local processes vs. unattested non-local ones.

3.1 Stress and weight

There are some otherwise quantity insensitive languages that allow monosyllabic feet at the 'end' of a parse in an odd-parity word if and only if the final syllable is heavy. These have been called generalized trochee (GT) languages (Kager 1992ab, citing Hayes 1991 for the term). This stress pattern is shown schematically in (20).[12]

[12] Hayes (1995) does not include the generalized trochee as a separate class of stress systems, but instead subsumes such systems under the syllabic trochee name, positing (1995:102) that (ˈL) is the only truly degenerate foot, regardless of the type of language and thus allowing languages that employ (ˈH) to nonetheless be syllabic trochees if they normally require (ˈσσ). Hayes (1995:103) uses Estonian to argue for this updated conception of syllabic trochees in addition to pointing out that syllabic trochee languages

(20) Generalized trochee pattern

a. Odd-parity words

σσσσH → (ˈσσ)(ˈσσ)(ˈH)

σσσσL → (ˈσσ)(ˈσσ)L

b. Even-parity words

σσσσ → (ˈσσ)(ˈσσ)

Some of the languages that are claimed to have this pattern or something like it are Estonian (Hint 1973; Prince 1980; Kager 1992ab), Weṛgaia (Hercus 1986; Hyde 2007), and some other Victorian Australian languages (Hercus 1986).

This kind of limited weight sensitivity is clearly intended to preserve alternating rhythm except when it would create a degenerate foot, i.e., (ˈL). The languages with this pattern do not show a preference for stressing heavy syllables generally, but they take advantage of a syllable's heaviness in final position in an odd-parity word to preserve a regular rhythmic alternation. GT languages, then, present a kind of local weight sensitivity, which we will see is analyzable in both HS and parallel OT. However, parallel OT additionally predicts the existence of a class of unattested GT-like languages with non-local weight sensitivity, as shown by Hyde (2007). I show in this section that opting for serial evaluation and a restricted GEN reins in OT's power and rules out this class of unattested languages.

3.1.1 Local weight-sensitivity in HS/IFO

To show how IFO handles generalized trochee languages we will illustrate with derivations for the Australian language Weṛgaia (Hercus 1986). Weṛgaia has a canonical generalized trochee stress pattern, and has been recently analyzed by Hyde (2007), with whom the observation of the inadequacy of parallel OT discussed in the following section originates. The forms in (21) illustrate the pattern.

often employ a bimoraic, rather than the expected disyllabic, word minimum. I will continue to refer to this pattern as the generalized trochee, however, since it differs from a language like Pintupi (section 2.3).

(21) Weṛgaia data[13]

a. Even parity: ˈwiṛimˈbulinj

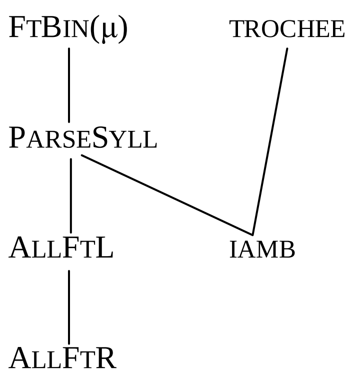
b. Odd with final heavy: ˈbunaˈḍug

c. Odd with medial heavy: ˈdaguŋga

d. Odd with initial heavy: ˈdelguna

To account for the pattern shown in (21) (which is identical to the schematic illustration in (20)) in HS/IFO, we will employ the same constraints used in the illustration of syllabic trochees in section 3.2, and nearly the same ranking will be necessary. Generalized trochees are much like syllabic trochees in terms of parsing, but while syllabic trochee languages like Pintupi leave the final syllable unparsed in an odd-parity word, GT languages like Weṛgaia will parse the syllable if it is heavy, that is, bimoraic. Simply switching the two FTBIN constraints in the hierarchy accounts for this difference (providing further justification to the proposal that they are two different constraints (Hewitt 1994)). Thus, the ranking from Pintupi is brought over but with FTBIN(μ) replacing FTBIN(σ) at the top of the hierarchy, as shown in (22).

(22) Weṛgaia ranking in HS/IFO



The reason that making this switch achieves the desired outcome is simple. Recall that the ranking FTBIN(σ) >> PARSESYLL was motivated in Pintupi in the final iteration of foot-building, in which the decision is made to forego additional parsing order

[13] Weṛgaia data are taken from Hyde (2007:306-7). According to Hyde CV:, CVV, and CVC count as heavy, but CV: never occurs in final position; Hercus (1986) indicates that only CVC induces final stress in odd-parity words. I ignore distinctions between degrees of stress (leftmost stress is primary).

to avoid a monosyllabic foot. Weṛgaia faces the same decision at the right edge of odd-parity words, but instead of requiring a disyllabic foot as in Pintupi or allowing any kind of monosyllable to become its own foot, it chooses to parse only if it would create a bimoraic foot. Thus, FTBIN(μ) dominates PARSESYLL (and PARSESYLL >> FTBIN(σ), in turn) to account for why some but not all monosyllables are parsed into feet in GT languages. I will note here in passing that high-ranking FTBIN(σ) is not required for the building of disyllabic feet throughout the word (as can be confirmed by reviewing the tableaux in section 2.3), and thus its low rank does not interfere with the ability of a GT language to parse syllabic trochees until the end of the word. Rather, it is PARSESYLL that effectively prefers disyllabic feet when possible; this point will be taken up again in section 4.

With the ranking we have motivated for GT languages, the tableaux in (23) through (26) illustrate how stress is derived in Weṛgaia using the data in (21). The tableau in (23) shows a derivation of an even-parity word, in which /LHLH/ becomes (ˈLH)(ˈLH). In the first iteration, the candidate with a disyllabic left-aligned trochee, *(ˈwi.ṛim)bu.linj*, is chosen as optimal because it meets the bimoraic minimum required by undominated FTBIN(μ), while parsing as much as possible (given that feet are maximally disyllabic), and being left-aligned in accordance with ALLFTL. This form is passed to the second iteration, where the addition of another disyllabic foot is found to be harmonically improving, so *(ˈwi.ṛim)(ˈbu.linj)* emerges as optimal. The third iteration shows convergence, which is trivial in this case because all syllables have been parsed into feet by this point. This derivation shows that, in general, parsing is insensitive to syllable weight in generalized trochee languages (and always in even-parity words).

(23) Weṛgaia /LHLH/ → (ˈLH)(ˈLH)

/wiṛimbulinj/ 'spider' **1st iteration**	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈwi)ṛim.bu.linj	**W** 1	**W** 3		**W** 3	**W** 1
b. wi(ˈṛim)bu.linj		**W** 3	**W** 1	2	**W** 1
c. (ˈwi.ṛim)bu.linj		2		2	
2nd iteration					
d. (ˈwi.ṛim)bu.linj		**W** 2	**L**	2	
e. (ˈwi.ṛim)(ˈbu.linj)			2	2	
3rd iteration					
f. (ˈwi.ṛim)(ˈbu.linj)			2	2	

Output: (ˈwi.ṛim)(ˈbu.linj)

In the tableau in (24) we see that an input with the shape /LLH/ becomes output (ˈLL)(ˈH). At the first iteration, the candidate *(ˈbu.na)ḍug* with a disyllabic left-aligned trochee wins. At the second iteration this candidate is compared with a candidate that has the last syllable footed, *(ˈbu.na)(ˈḍug)*. The latter wins because it better satisfies PARSESYLL while also satisfying the bimoraic minimum on feet, since the last syllable is heavy. The third iteration shows convergence, which occurs because no more metrification is possible, and thus the output is *(ˈbu.na)(ˈḍug)*.

(24) Weṛgaia /LLH/ → (ˈLL)(ˈH)

/bunaḍug/ 'broad-leaved mallee' **1st iteration**	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈbu)na.ḍug	**W** 1	**W** 2		**W** 2	**W** 1
b. (ˈbu.na)ḍug		1		1	
c. bu(ˈna.ḍug)		1	**W** 1	**L**	
d. bu.na(ˈḍug)		**W** 2	**W** 2	**L**	**W** 1
2nd iteration					
e. (ˈbu.na)ḍug		**W** 1	**L**	1	**L**
f. (ˈbu.na)(ˈḍug)			2	1	1
3rd iteration					
g. (ˈbu.na)(ˈḍug)			2	1	1

Output: (ˈbu.na)(ˈḍug)

The tableaux in (25) and (26) below show derivations for odd-parity words ending in light syllables. In both, the first iteration chooses the disyllabic left-aligned trochee as the best foot, just as in the first iterations of (23) and (24). In (25) and (26), the second iteration compares the faithful single-foot candidate with a candidate in which the final syllable is parsed into a foot, and in both cases footing the last syllable fails to improve harmony because FTBIN(μ) outranks PARSESYLL. Thus, the output in (25) is *(ˈda.guŋ)ga*, and the output in (26) is *(ˈdel.gu)na*.

(25) Weṛgaia /LHL/ → (ˈLH)L

/daguŋga/ ‘to punch’ **1st iteration**	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈda)guŋ.ga	**W** 1	**W** 2		**W** 2	**W** 1
b. (ˈda.guŋ)ga		1		1	
c. da(ˈguŋ.ga)		1	**W** 1	**L**	
2nd iteration					
d. (ˈda.guŋ)ga		1		1	
e. (ˈda.guŋ)(ˈga)	**W** 1	**L**	**L** 2	1	**W** 1

Output: (ˈda.guŋ)ga

(26) Weṛgaia /HLL/ → (ˈHL)L

/delguna/ ‘to cure’ **1st iteration**	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈdel)gu.na		**W** 2		**W** 2	**W** 1
b. (ˈdel.gu)na		1		1	
c. del(ˈgu.na)		1	**W** 1	**L**	
2nd iteration					
d. (ˈdel.gu)na		1		1	
e. (ˈdel.gu)(ˈna)	**W** 1	**L**	**L** 2	1	**W** 1

Output: (ˈdel.gu)na

A summary of the pattern that is produced in HS/IFO, given the ranking in (22), is shown in (27). As the tableaux above indicate, this stress pattern is handled straightforwardly in HS with standard stress constraints and the assumptions of IFO adopted here.

(27) Even-parity words: parsing insensitive to weight

LHLH → (ˈLH)(ˈLH)

Odd-parity words: H is footed as (ˈH) at the right edge only

LLH → (ˈLL)(ˈH)   LHL → (ˈLH)L   HLL → (ˈHL)L

One characteristic of all these derivations is that they all begin in the same way. At each first iteration, the candidate with a disyllabic left-aligned trochee is chosen as optimal. We will return to this point in section 3.1.3, where it will be important for explaining why HS/IFO does not predict the unattested language that parallel OT does.

3.1.2 Stress and weight in parallel OT

Hyde (2007) observes that parallel OT cannot account for the Weṛgaia stress pattern using just the standard parsing constraints (ALLFTL, PARSESYLL, FTBIN) but it can do so with a high-ranked rhythm constraint, *CLASH, as defined in (28) (Prince 1983, Selkirk 1984, Kager 1994).

(28) ***CLASH**: Assign 1 violation mark for every adjacent pair of stressed syllables.

To see why *CLASH is necessary we can consider the input /HLL/. The tableau in (29) shows that this input is incorrectly parsed as (ˈH)(ˈLL) in parallel OT using only the parsing constraints. The reason we get this outcome in parallel OT is that the ranking which permits (ˈH) feet, namely FTBIN(μ) >> PARSESYLL, is needed in order to account for mappings like the one in (24), /LLH/ → (ˈLL)(ˈH), and the optimal candidate in (29) capitalizes on this allowance to achieve greater parsing. In fact, no ranking of these constraints will derive the Weṛgaia pattern – that is, no ranking simultaneously prefers /LLH/ → (ˈLL)(ˈH) and /HLL/ → (ˈHL)L (Hyde 2007).

(29) /HLL/ ↛ (ˈHL)L – Wrong result

/delguna/ 'to cure'	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈdel.gu)na		**W** 1	**L**	**L** 1	**L**
☠ → b. (ˈdel)(ˈgu.na)			1	2	1
c. (ˈdel.gu)(ˈna)	**W** 1	**L**	**W** 2	1	**W** 1
d. del(ˈgu.na)		1	**W** 1	**L**	

Since the undesired winner in (29) has a stress clash, which no words of Weṛgaia ever show on the surface, we could reasonably posit that *CLASH is undominated. This will rule out the mapping /HLL/ → (ˈH)(ˈLL) that we found in (29), and instead get the desired result for this input, as illustrated in (30). This ranking would continue to allow the other attested mappings in Weṛgaia since no outputs in this language violate *CLASH.

(30) /HLL/ → (ˈHL)L in parallel OT with high-ranked *CLASH

/delguna/ 'to cure'	*CLASH	FTBIN(μ)	PARSESYLL	ALLFTL
→ a. (ˈdel.gu)na			1	
b. (ˈdel)(ˈgu.na)	**W** 1		**L**	**W** 1
c. (ˈdel.gu)(ˈna)		**W** 1	**L**	**W** 2

But, as Hyde (2007) points out, considering the constraints needed here reveals a serious over-generation problem in parallel OT. When *CLASH is low-ranked, a class of languages that have a very peculiar kind of weight sensitivity is predicted to exist. With the ranking FTBIN(μ) >> PARSESYLL >> ALLFTL >> *CLASH, for instance, parallel OT generates a language with the stress system summarized in (31) (paraphrasing Hyde 2007: 312).

(31) Language predicted when FTBIN(μ) >> PARSESYLL >> ALLFTL >> *CLASH

- Parsing is *in*sensitive to the weight of a heavy syllable when it occurs in an even-numbered syllable counting from the left or in any syllable of a word with even parity.
- Parsing *is* sensitive to the weight of a heavy syllable when it occurs in an odd-numbered syllable of an odd-parity word, and is the closest heavy to the left edge of the word among those heavy syllables with these properties.

Tableaux (32)-(35) schematically illustrate this generalization with larger words.

Tableau (32) shows that a word with an even-number of light syllables followed by a heavy syllable is parsed maximally similar to parsing in Weṛgaia and other GT languages.

(32) Maximal parsing with final heavy syllable

/LLLLH/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈLL)(ˈLL)(ˈH)			6	
b. (ˈLL)(ˈLL)H		**W** 1	**L** 2	

However, in (33) and (34) we see that a medial or initial heavy syllable with an even number of light syllables on either side is also parsed into a monosyllabic foot, as permitted by FTBIN(μ), in order to satisfy PARSESYLL. The low ranking of *CLASH ensures it will not be respected.

(33) Maximal parsing with medial heavy syllable

/LLHLL/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈLL)(ˈH)(ˈLL)			5	1
b. (ˈLL)(ˈHL)L		**W** 1	**L** 2	**L**
c. (ˈLL)(ˈHL)(ˈL)	**W** 1		**W** 6	**L**

(34) Maximal parsing with initial heavy syllable

/HLLLL/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈH)(ˈLL)(ˈLL)			4	1
b. (ˈHL)(ˈLL)L		**W** 1	**L** 2	**L**
c. (ˈHL)(ˈLL)(ˈL)	**W** 1		**W** 6	**L**

Finally, tableau (35) shows that if there is more than one heavy syllable in an odd-numbered syllable of an odd-parity word, the first (left-most) is parsed as a monosyllabic foot, while parsing remains insensitive to the other's weight.[14] Again, *CLASH is too low-ranked to exert any influence.

[14] As Hyde points out this is a variant of the problem with alignment and monosyllabic feet first noticed by Crowhurst and Hewitt (1995a). In section 4 I show why this particular behavior of alignment does not apply in HS/IFO.

(35) Maximal parsing with initial and final heavy syllables

/HLLLH/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈH)(ˈLL)(ˈLH)			4	1
b. (ˈHL)(ˈLL)H		**W** 1	**L** 2	**L**
c. (ˈHL)(ˈLL)(ˈH)			**W** 6	**L**

Meanwhile, the tableaux in (36) illustrate that even-parity words are parsed into disyllabic feet and that heavy syllables do not exhibit weight sensitivity in this case.

(36) Even-parity words – no weight sensitivity

(a) /HLLL/ → (ˈHL)(ˈLL)

/HLLL/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈHL)(ˈLL)			2	
b. (ˈH)(ˈLL)L		**W** 1	**L** 1	**W** 1

(b) /HLLH/ → (ˈHL)(ˈLH)

/HLLH/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈHL)(ˈLH)			2	
b. (ˈH)(ˈLL)(ˈH)			**W** 4	**W** 1

(c) /LHLH/ → (ˈLH)(ˈLH)

/HLLH/	FTBIN(μ)	PARSESYLL	ALLFTL	*CLASH
→ a. (ˈLH)(ˈLH)			2	
b. L(ˈHL)(ˈH)		**W** 1	**W** 4	
c. (ˈLH)L(ˈH)		**W** 1	**W** 3	

This stress system is schematically summarized in (37).

(37) Schematic summary of predicted language:

a. Even parity words

σσσσ → (ˈσσ)(ˈσσ)

where σ = any weight (H or L)

b. Odd parity words

Hσσσσ → (ˈH)(ˈσσ)(ˈσσ)

LσHσσ → (ˈLσ)(ˈH)(ˈσσ)

LσLσH → (ˈLσ)(ˈLσ)(ˈH)

etc.

The language described in (31) and illustrated in (37) is unattested, as no known stress system matches this description. Hyde also shows that these constraints predict several other variations on the stress pattern in (31)/(37), none of which is attested. Given the strange character of these patterns, this predicted class of languages is not plausibly an accidental gap in the typology of stress but rather appears to be an example of undesirable over-generation in parallel OT.

Hyde (2007) argues that what is responsible for this prediction is the behavior of the constraints PARSESYLL and FTBIN as well as traditional structural assumptions, and he provides an analysis in parallel OT using a different set of stress constraints and structural assumptions. However, in the next section I show that because of HS/IFO's local decision-making (that is, its lack of foresight), it cannot reproduce the prediction of the unattested language under any ranking of the standard constraints. It will thus be argued that the alternative culprits for this pathological prediction are parallelism and global evaluation.

3.1.3 Local vs. global sensitivity to weight

In the cases considered so far, parallel OT and HS/IFO choose different optima for the same inputs even under the same constraint ranking. With the ranking FTBIN(μ) >> PARSESYLL >> ALLFTL >> *CLASH, parallel OT will choose the global optimum (ˈH)(ˈLL) for input /HLL/ because it is the most harmonic among all possible metrical parses, while HS/IFO converges on (ˈHL)L after building the (ˈHL) foot in the first step. Meanwhile, both OT and HS/IFO parse even-parity words into disyllabic trochees, regardless of syllable heaviness, e.g. /HLLL/ → (ˈHL)(ˈLL). In HS/IFO this set of optima is achieved in the following way: both (ˈHL)L and (ˈHL)LL win on the first iterations of their respective derivations because PARSESYLL and ALLFTL prefer a maximal left-aligned foot, and the assumption of strict inheritance ensures that (ˈHL)L will never lead to (ˈH)(ˈLL). The derivation lacks the ability to look ahead and see that the global optimum for input /HLL/, namely (ˈH)(ˈLL), could be reached by first creating non-optimal (ˈH)LL; it simply chooses the locally optimal candidate instead. This is a clear case in which parallel OT and HS make different predictions because the latter converges on a local optimum and rightly fails to find the global optimum.

The previous section discussed the odd nature of the language parallel OT predicts under this ranking. From the standpoint of iterative parsing, the real weirdness of this language lies in its ability to treat an initial heavy syllable differently depending on whether an odd or even number of syllables follow it.[15] For input /HLL/, which has an even number of syllables after the heavy, we get (ˈH)(ˈLL) with the heavy syllable parsed as a monosyllabic foot, while for /HLLL/, which has an odd number of syllables after the heavy, we get (ˈHL)(ˈLL) with the heavy syllable parsed into a disyllabic foot with the immediately following light syllable. If no heavy syllables exist in a word, the parsing is unambiguously left-to-right, (ˈLL)(ˈLL)L. This unattested language thus utilizes a kind of weight sensitivity that can only be described as non-local. This contrasts with the kinds of local weight sensitivity that we actually see in languages such as Weṛgaia, in which it is only at the 'end' of the metrical parse where stress is sensitive to the weight of a final stray syllable in an odd-parity word.

Importantly, HS/IFO not only fails to generate the unattested parity-counting language under this ranking – it is not possible to analyze this class of unattested languages in HS/IFO under *any* ranking. The reason is that the ranking that would be necessary to get the derivation /HLL/ → (ˈH)LL → (ˈH)(ˈLL) could not also produce /HLLL/ → (ˈHL)LL → (ˈHL)(ˈLL). At each first stage, the ranking that would prefer /HLL/ → (ˈH)LL would also prefer /HLLL/ → (ˈH)LLL, while the ranking that would prefer /HLLL/ → (ˈHL)LL would also prefer /HLL/ → (ˈHL)L. That is, these inputs must be treated the same at the first iteration, since the grammar computes only local optimality and does not know when a global optimum could be achieved by choosing a locally non-optimal form along the way. The tableaux in (38) and (39) make this point explicit.

The tableaux in (38) show that the constraints we have been considering prefer the disyllabic trochee over the monosyllable heavy at the first iteration for both inputs (*CLASH is omitted because it is ranked too low to exert influence in this case.)

[15] Actually not just an initial heavy syllable, but any heavy syllable in an odd-numbered syllable of an odd-parity word provided it is the first of such syllables counting from the left.

(38) Disyllabic trochee preferred in both derivations – 1st iteration

/HLL/	FTBIN(μ)	PARSESYLL	ALLFTL	ALLFTR	FTBIN(σ)
a. (ˈH)LL		**W** 2		**W** 2	**W** 1
→ b. (ˈHL)L		1		1	
/HLLL/					
a. (ˈH)LLL		**W** 3		**W** 3	**W** 1
→ b. (ˈHL)LL		2		2	

The tableaux in (39) show that when we insert a constraint in the hierarchy which favors (ˈH) over (ˈHL) and outranks PARSESYLL – e.g., *(ˈHL) – it favors (ˈH) in *both* derivations, not just the top one.[16]

(39) Monosyllabic heavy preferred in both derivations – 1st iteration

/HLL/	FTBIN(μ)	*(ˈHL)	PARSESYLL	ALLFTL	ALLFTR	*CLASH
→ a. (ˈH)LL			2		2	
b. (ˈHL)L		**W** 1	**L** 1		**L** 1	
/HLLL/						
→ a. (ˈH)LLL			3		3	
b. (ˈHL)LL		**W** 1	**L** 2		**L** 2	

In other words, the combination of outputs in the unattested GT-like language(s), though all globally optimal under the same ranking in parallel OT, cannot be modeled as outputs in the same language using HS/IFO with these constraints.

The fact that languages of this sort do not exist suggests that a more traditional view of metrical parsing is more correct than the standard OT analysis. OT predicts that whether an HL sequence is parsed as (ˈHL)… or as (ˈH)L… depends on properties of the word that would not yet be evident if we are parsing from left to right (as /LLL/ → (ˈLL)L indicates). The tableaux in (38) and (39) illustrated why this is not a possible pattern in a serial model in which feet are built incrementally. In an actual HS derivation a directional parse proceeding from left to right will cause each next foot to be affected by the location of the previously placed feet, but a syllable cannot be treated differently depending on whether or not an even or odd number of syllables *follow* it.[17] Odd and

[16] Actually, the constraints ranked as in (39) will favor building (ˈLL) feet before monosyllabic feet, though the eventual output will be the same as that intended here.
[17] Vice versa for right to left stress, of course.

even parity words are predicted in HS to be treated the same until the ‘end’ of the parse, where it is clear that odd-parity words have a stray syllable, for instance, while even-parity words do not. Thus, both /HLL/ and /HLLL/ are first footed as (ˈHL)L and (ˈHL)LL in GT languages, or both as (ˈH)LL and (ˈH)LLL in, e.g., a moraic trochee language.

This prediction relies on the assumption that constraints with definitions such as “A foot should be followed by an even number of syllables” do not exist. If such constraints were admitted into the theory this prediction would not hold, as it would then be possible in a left-to-right derivation to asses the consequences of the choice of a foot for the potential metrical parses of the remaining syllables into feet. Such a constraint could thus simulate derivational look-ahead in certain circumstances. Notice however, that this constraint would only set up the derivation for subsequent disyllabic footing indirectly, by referencing parity, but the constraint itself has no metrical characteristics. Instead, it would enforce an output preference that happens to be important for leading the way to the global optimum in this example. In fact, the implicit assumption of work in metrical theory is that parity counting is carried out exclusively via metrical representations, namely feet, and thus we would not generally expect a constraint to be afforded this power. Notice that in parallel OT the comparison of fully-specified metrical parses with one another permits exactly such parity counting. That is, allowing the comparison of candidates with diverse, fully-specified metrical constituents effectively transfers the parity-counting power of feet beyond their normal purview to create the strange prediction outlined in the previous section.

The theory of IFO proposed here lacks any source of derivational foresight, and for good reason. Although it can ‘see’ the whole word in each of its local evaluations, it does not envision the possible paths that each local optimum might lead to and does not know that some local optima may lead to a global optimum while others do not. Instead, it chooses based only on the relative harmony of the candidates at that iteration. As I have attempted to illustrate in this section, this is a positive prediction of HS/IFO when compared with parallel OT, since it more accurately reflects the typology of stress systems and how and when properties of a word can affect metrification.

This observation recalls a similar argument made by Prince (1990) with respect to Harmonic Parsing (which was discussed in section 2.4.2). Prince encounters an example of a similar type in his analysis of Cairene Arabic stress. In the analysis of input /ʔadwiyatu/ 'drug nom.sg', adding the 'best foot' in a left-to-right parse yields *(ˈʔad)wiyatu*, according to the metric of foot well-formedness proposed in HP, and this is the right choice, as *(ˈʔad)(ˈwiya)tu* is the ultimate output. But if the derivation had foresight it would see that building the suboptimal *(ˈʔadwi)yatu* would ultimately lead to a more complete parse of the word. Disregarding the fact that in this example what counts as 'best' is different from that we assumed with generalized trochees (since ranking will account for this), we have another case in which locality plays an important role. Prince (1990) has this to say: "Notice that the parsing decision is made locally: even though the incorrect *(ˈʔadwi)(ˈyatu)* has the not-inconsiderable virtue of avoiding the unparsed sequence…at the end of the word, there can in fact be no anticipatory admission of the second-best HL foot." Precisely.

It is due to the fundamental character of serial evaluation combined with the assumptions of IFO that we do not predict the unattested GT-like languages when optimization is performed iteratively with restricted candidate sets. This is something an HS/IFO derivation can only do with the power of look-ahead, because the relevant properties are influencing parsing at a distance. In parallel OT this kind of global maximization is par for the course, but since no language seems to stress words in this way, the standard theory is too permissive. In the next two sections I show that in fact this is not an isolated case, and that parallel OT predicts a host of other non-local interactions that are not attested. In each case HS/IFO correctly predicts they should not occur.

3.2 Stress and vowel shortening

Syllable weight is important for quantity sensitive stress systems so it is not surprising that there are processes of quantity adjustment that seem to be motivated by metrical structure. An example is the process of 'trochaic shortening,' wherein a stressed heavy syllable becomes light, usually by a process of vowel shortening, before an unstressed light syllable. Motivations for this process within metrical theory generally cite the preference for trochaic feet to group elements of equal weight (Hayes 1985, 1987, 1995; Prince 1990), ruling out (ˈHL) as a parse.

This section first shows that trochaic shortening is attested as a local interaction, illustrated by data from Fijian. Both HS/IFO and parallel OT can easily account for local trochaic shortening. However, parallel OT additionally leads us to expect a nominally similar but non-local variant that is not attested in known natural language stress systems and has strange characteristics, paralleling the results of the previous section. The finding is ultimately the same as the previous section: the mechanisms commonly employed in parallel OT to account for the well-known local process can again not fail to predict unattested non-local interactions as well.

3.2.1 Local shortening in HS/IFO

An example of a language with standard trochaic shortening is Fijian (Hayes 1995: 142-9; Schütz 1985; Dixon 1988), in which a long vowel in the penultimate syllable shortens when the vowel in the final syllable is short. Hayes (1995: 145) gives the description in (40). Fijian has a right-to-left moraic trochee stress system in which long vowels count as heavy.

(40) Fijian shortening

V: →V / __CV#

Trochaic shortening in Fijian is a local process, which both parallel OT and HS/IFO can account for. It qualifies as local because shortening is dependent only on the local context – whether the following syllable is light (and word-final) – and does not require look-ahead to know whether it should apply.

A serial analysis of trochaic shortening can be implemented in HS/IFO along the lines of Prince (1990)'s analysis of trisyllabic shortening in English, which is similar to Fijian shortening, but with final syllable extrametricality, so it targets the antepenult rather than the penult.[18] By this account, an (ˈHL) foot is built at the first step, and then shortening to (ˈLL) occurs in the second. An input /HL/ cannot become output (ˈLL) in

[18] The moniker "trisyllabic shortening" is potentially misleading, since as Myers (1987) shows and Prince (1990) acknowledges, the process is foot-based.

one step in an HS/IFO analysis because two changes, shortening and foot building, must take place. This follows from the gradualness requirement in HS and the claim in this paper that building a foot constitutes its own step. There are two logical derivations from /HL/ to (ˈLL) given my assumptions: either shortening first, HL → LL → (ˈLL), or foot building first, HL → (ˈHL) → (ˈLL). Since it is assumed that shortening is a consequence of preferring (ˈLL) over (ˈHL), it cannot be motivated to occur until foot building has taken place, and thus the latter derivation is the one I have assumed, essentially following Prince (1990). The constraints that will be relevant for deriving trochaic shortening include the metrical constraints familiar from the analyses in the previous two sections (PARSESYLL, ALLFTL/R, FTBIN(μ), FTBIN(σ)) and some additional constraints defined below.

In our analysis, (ˈHL) feet are built and then the heavy syllable is shortened, and therefore a constraint disfavoring (ˈHL) feet must dominate a faithfulness constraint penalizing shortening. The (ˈHL)-dispreferring constraint assumed here penalizes trochees which contain more than two moras; we will call this constraint BALANCED, since it prefers (ˈLL) and (ˈH) as trochaic feet while penalizing both (ˈLH) and (ˈHL). Though note that this constraint, defined in (41), also penalizes (ˈHH) feet.[19] The relevant faithfulness constraint against shortening we will assume is MAX-μ, defined as in (42).

(41) **BALANCED (BAL)**: Assign one violation mark for a trochee that contains more than two moras. E.g., (ˈLL) and (ˈH) receive zero marks; (ˈHL), (ˈLH), and (ˈHH) receive one.

(42) **MAX-μ**: Assign one violation mark for a mora in the input that does not have a correspondent in the output.

Clearly, the constraint dispreferring (ˈHL) feet must itself be dominated by a constraint that is satisfied by building one; otherwise we could not motivate the proposed derivation. The process of trochaic shortening is obligatory when the light syllable is

[19] The implications of this constraint's formulation are not a central issue here; any constraint or set of constraints which penalizes (ˈHL) (and later (ˈLH)) would suffice. (Though see fn. 26).

word-final, but optional otherwise, so we will first concentrate on accounting for the pattern of final trochaic shortening only. Since an (ˈHL)# foot is created in the derivation before the H is shortened, (ˈHL)# must win over competitors such as (ˈH)L#. To ensure this we will posit the constraint in (43), ALIGNWDR, which prefers a word to have a foot at its right edge. Although this constraint is nominally from the alignment family, it is defined here categorically (see also McCarthy 2003: 109). When this constraint dominates BAL, (ˈHL)# will be preferred over (ˈH)L#. To ensure that the language otherwise prefers (ˈH)L word-medially, BAL must dominate PARSESYLL, which would otherwise exert its preference for the larger (ˈHL) foot throughout.

(43) **ALIGNWDR**: Assign one violation mark for a word that does not have a foot at its right edge.

Recall that in HS, unlike in rule-based theories, processes are not explicitly ordered with respect to one another, but instead an optimal process is chosen at each iteration by the constraint ranking. This point bears mentioning because this section explicitly considers not only metrical parsing at each iteration but also the process of heavy syllable shortening. In this case, trochaic shortening will not be harmonically improving until foot building has taken place because the constraint that is satisfied by shortening references foot shape, and thus these two processes have a necessary ordering under this ranking.

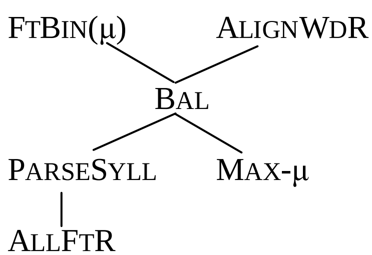
The derivation in (44) shows an input of the form /LLHL/, which is ultimately parsed as (ˈLL)(ˈLL) with vowel shortening in the penultimate syllable. Ranking arguments can be determined from this derivation. Working backwards, the third iteration informs us that PARSESYLL must dominate ALLFTR in order to ensure that feet are built beyond the right edge, i.e., for (ˈLL)(ˈLL) to beat LL(ˈLL). This is again the familiar ranking for iterative foot building (McCarthy and Prince 1993a). Given this ranking, the second iteration requires that BAL dominate both MAX-μ and PARSESYLL. The ranking BAL >> MAX-μ ensures that an unbalanced foot (e.g., (ˈHL)) can be corrected by shortening the vowel (i.e., deleting a mora), while BAL >> PARSESYLL means that correction of the (ˈHL) foot will occur before subsequent foot-building. In other derivations the ranking BAL >> PARSESYLL will also ensure that word medial HL sequences are not parsed as (ˈHL).[20] Finally, the first iteration shows that both FTBIN(μ) and ALIGNWDR must dominate BAL. The ranking FTBIN(μ) >> BAL is required to rule out H(ˈL)#, while ALIGNWDR >> BAL is required to make sure an (ˈHL) foot is built at the right edge of the word at the beginning of the derivation, even though it violates BAL. The ranking is summarized in (45).

(44) Fijian trochaic shortening; input /LLHL/

/LLHL/ **1st iteration**	FTBIN(μ)	ALIGNWDR	BAL	MAX-μ	PARSESYLL	AllFtR
a. LLH(ˈL)	**W** 1		**L**		**W** 3	
b. LL(ˈHL)			1		2	
c. LL(ˈH)L		**W** 1	**L**		**W** 3	**W** 1
d. LLLL		**W** 1	**L**	**W** 1	**W** 4	
2nd iteration						
e. LL(ˈHL)			**W** 1	**L**	2	
f. LL(ˈLL)				1	2	
g. (ˈLL)(ˈHL)			**W** 1	**L**	**L**	**W** 2
3rd iteration						
h. LL(ˈLL)					**W** 2	**L**
i. (ˈLL)(ˈLL)						2

Output: (ˈLL)(ˈLL)

(45) Fijian penultimate trochaic shortening ranking



[20] It is not so important for our purposes that shortening happens before *subsequent* foot building as long as the shortening eventually does happen. But because we independently need the BAL >> PARSESYLL ranking so that medial HL sequences are not parsed as (ˈHL), this ranking will force this ordering. This is another example of process ordering emerging from constraint ranking.

The tableau in (46) shows a derivation similar to the one in (44), but with input /HLHL/, which has both a final and a non-final HL sequence. As this derivation confirms, the word-final HL sequence is parsed as (ˈLL), while the non-final HL sequence is parsed as (ˈH)L.

(46) Fijian input /HLHL/

/HLHL/ **1st iteration**	FTBIN(μ)	ALIGNWDR	BAL	MAX-μ	PARSESYLL	AllFtR
a. HLH(ˈL)	**W** 1		**L**		**W** 3	
b. HL(ˈHL)			1		2	
c. HL(ˈH)L		**W** 1	**L**		**W** 3	**W** 1
d. HLLL		**W** 1	**L**	**W** 1	**W** 4	
2nd iteration						
e. HL(ˈHL)			**W** 1	**L**	2	
f. HL(ˈLL)				1	2	
g. (ˈHL)(ˈHL)			**W** 2	**L**	**L**	**W** 2
h. (ˈH)L(ˈHL)			**W** 1	**L**	**L** 1	**W** 3
3rd iteration						
i. HL(ˈLL)					**W** 2	**L**
j. (ˈHL)(ˈLL)			**W** 1		**L**	**L** 2
k. (ˈH)L(ˈLL)					1	3

Output: (ˈH)L(ˈLL)

Under this ranking we also get the following derivations, which accurately reflect Fijian's stress system. Thus, the analysis in HS/IFO captures the Fijian pattern of stress and penultimate syllable trochaic shortening straightforwardly.

(47) /LLLH/ → LLL(ˈH) → L(ˈLL)(ˈH)
/LHLL/ → LH(ˈLL) → L(ˈH)(ˈLL)
/HLLL/ → HL(ˈLL) → (ˈH)L(ˈLL)

As noted above, Fijian also variably exhibits trochaic shortening throughout the word, though the shortening process in the penult is always obligatory. We can account for HL shortening throughout the word by assuming that when this variation obtains, ParseSyll is ranked over Bal, with the derivation proceeding as shown in (48). The ranking of AlignWdR is no longer crucial because ParseSyll is now higher ranked than Bal and is best satisfied by building (ˈHL) feet everywhere and not just at the word edge. This constraint is shown for illustration at the bottom of the hierarchy in (48), though there is now no ranking evidence to support its position in this language. We should also note that in the third iteration candidates (j) and (k) tie because they both repair a Bal violation by deleting a mora; the choice of (j) as the winner here is arbitrary, and the important thing is that both HL sequences are shortened, in some order.

(48) Fijian input /HLHL/ with medial and penultimate trochaic shortening

/HLHL/ **1st iteration**	FtBin(μ)	ParseSyll	Bal	AllFtR	Max-μ	AlignWdR
a. HLH(ˈL)	**W** 1	**W** 3	**L**			
b. HL(ˈHL)		2	1			
c. HL(ˈH)L		**W** 3	**L**	**W** 1		**W** 1
d. HLLL		**W** 4	**L**		**W** 1	**W** 1
2nd iteration						
e. HL(ˈHL)		**W** 2	**L** 1	**L**		
f. HL(ˈLL)		**W** 2	**L**	**L**	**W** 1	
g. (ˈHL)(ˈHL)			2	2		
h. (ˈH)L(ˈHL)		**W** 1	**L** 1	**W** 3		
3rd iteration						
i. (ˈHL)(ˈHL)			**W** 2	2	**L**	
j. (ˈHL)(ˈLL)			1	2	1	
k. (ˈLL)(ˈHL)			1	2	1	
4th iteration						
l. (ˈHL)(ˈLL)			**W** 1	2	**L**	
m. (ˈLL)(ˈLL)				2	1	

Output: (ˈLL)(ˈLL) (5th iteration not shown)

Some comments are in order on these analyses. First, they require that (ˈHL) feet be built in the course of the derivation, though such feet arguably do not occur on the surface in Fijian. This difference between intermediate and surface forms is expected however. Because of the ranking Bal >> Max-μ in Fijian, *any* (ˈCVː.CV) foot built at an intermediate stage in the derivation will be shortened to (ˈCV.CV) in a subsequent step. That is, it will always be harmonically improving to shorten this configuration, and therefore the absence of (ˈHL) feet on the surface is not evidence that (ˈHL) feet cannot form part of a licit derivation in the language.[21]

In addition, the intermediate stages in the derivation reflect well-attested output preferences. The building of an intermediate (ˈHL) foot in the obligatory penult shortening case reflects the need to satisfy ALIGNWDR at the expense of foot form, and although Fijian later corrects the suboptimal foot, other languages tolerate a suboptimal foot on the surface in order to satisfy an ALIGNWD constrain. Some examples are German (Alber 1997, 2005), and Finnish (Hanson and Kiparsky 1996; Alber 1997, 2005).[22] In Finnish, a disyllabic foot appears at the left edge of the word, satisfying ALIGNWDL, no matter the weight of the initial two syllables. Elsewhere in a word subsequent foot building will obey foot form by sacrificing violations of general ALLFTL in order to avoid (ˈLH) feet, and the result is occasional ternary rhythm. Thus, feet not satisfying foot form constraints are tolerated on the surface to avoid violating ALIGNWDL in Finnish (and German), just as an otherwise illicit foot in Fijian is built at an intermediate stage of our derivation in order to satisfy ALIGNWDR (or PARSESYLL, in the HL-shortening-throughout case).

The shortening step in the derivation also reflects a well-attested output preference. The shortening itself is motivated by the desire for balanced trochees, a preference reflected cross-linguistically in trochaic languages (Hayes 1985, 1987, 1995; McCarthy and Prince 1986). This preference is often observed by failing to build (ˈHL) trochees at all, preferring when possible (ˈH) or (ˈLL), though shortening to create balanced trochees is also relatively common, occurring, according to Hayes (1995:148), in Hawaiian, Tongan, (Middle) English, and some Italian dialects, in addition to Fijian.[23]

Thus, although our analysis requires the intermediate building of a foot that otherwise is not part of the foot inventory of Fijian, this is not in itself a problem since it is always harmonically improving to subsequently correct the foot. And furthermore, the

---

[21] See McCarthy (to appear-a) for a related argument regarding intermediate derivational stages in Tongan.

[22] In most attested cases of this kind there is a confound since the single-edge foot is also the main stress foot. This could be captured in the present theory by relativizing ALIGNWDR/L to main stress only. The analysis of Fijian presented here would remain the same, since the rightmost stress is the main stress.

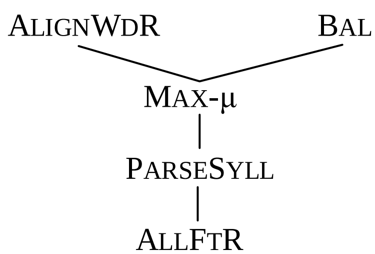
[23] Latin exhibits lightening of the heavy syllable in some LH sequences to create a balanced (ˈLL) foot (Mester 1992, Prince 1990), which might be argued to be related to trochaic shortening, if a constraint like BALANCED is indeed what compels length adjustments in trochees.

derivation provided here is made of component processes which are well-attested repairs to common output preferences in stress systems. This is what we expect from a HS derivation.

3.2.2 Stress and shortening in parallel OT

We have just seen that Fijian is easily analyzed in HS/IFO. To illustrate how parallel OT handles Fijian penultimate trochaic shortening, the same constraints from the previous section are needed, though their ranking will be slightly different. I will first assert the ranking and then point out the ranking arguments in the tableau presented for illustration. The ranking that is necessary is shown in (49), assuming additionally that Trochee and FtBin(μ) are undominated and that AllFtL is ranked below AllFtR.

(49) Fijian penultimate trochaic shortening ranking in parallel OT



The tableau in (50) shows the input /HLHL/, which becomes output (ˈH)L(ˈLL), illustrating how HL sequences are treated word-finally and word-medially. In (50) the candidate in (e) wins over its competitors because it obeys proper foot form (Bal) while satisfying the need to have a foot at the right edge of the word (AlignWdR). This tableau contains the necessary information for determining ranking arguments. ParseSyll must dominate AllFtR, as losing candidate (g) shows. Candidate (f) illustrates that Max-μ dominates ParseSyll. Candidates (a) through (d) show that AlignWdR and Bal must dominate Max-μ, justifying the hierarchy in (49).[24]

[24] In order to account for the optional case in which HL is shortened throughout the word we must invert ranking of the constraints Max-μ and ParseSyll, but as we will see shortly, this move alone has undesirable consequences.

(50) Fijian /HLHL/ → (ˈH)L(ˈLL) – Penultimate trochaic shortening in parallel OT

/HLHL/	ALIGNWDR	BAL	MAX-μ	PARSESYLL	ALLFTR
a. HL(ˈHL)		**W** 1	**L**	**W** 2	**L**
b. (ˈHL)(ˈHL)		**W** 2	**L**	**L**	**L** 2
c. (ˈH)L(ˈH)L	**W** 1		**L**	1	**W** 4
d. (ˈH)L(ˈHL)		**W** 1	**L**	1	3
→ e. (ˈH)L(ˈLL)			1	1	3
f. (ˈLL)(ˈLL)			**W** 2	**L**	**L** 2
g. HL(ˈLL)			1	**W** 2	**L**

The ranking required for the parallel OT analysis is slightly different from the one required for the HS/IFO analysis. In HS/IFO ALIGNWDR must dominate BAL in order for (ˈHL)# to win over a misaligned competitor that satisfies BAL (namely, (ˈH)L#), but in parallel OT, ALIGNWDR and BAL are not able to be ranked with respect to one another because all surface forms satisfy both constraints. Additionally, in parallel OT MAX-μ must dominate PARSESYLL in order to rule out shortening heavy syllables throughout the word for more complete parsing, but in HS/IFO these constraints cannot be ranked because they do not conflict in any one step of the derivation. Such differences between rankings are the result of the fact that GEN restricts the candidates that compete with one another in the serial model while no such restrictions are present in the parallel one.

From these illustrations it is evident that parallel OT can account for the obligatory penult shortening process in Fijian. However, as with the generalized trochee illustration in section 3.1, the problem with parallel OT is over-generation. With the power of a parallel theory we predict, in addition to Fijian, languages that are not attested though they undergo nominally similar shortening processes. An example of such a prediction can be found with the ranking in (51), assuming also that FTBIN(μ) is undominated. The crucial problematic piece of this ranking is PARSESYLL >> MAX-μ which will allow non-local shortening for maximal parsing.

(51) BAL, PARSESYLL >> MAX-μ, ALLFTL >> ALLFTR

The tableaux in (52) through (54) illustrate this hypothetical language. In this language, a word beginning with a heavy syllable followed by an even number of light syllables separating it from the word edge will not exhibit shortening of the heavy syllable, as shown in (52), because an (ˈH) is a licit balanced trochee and maximal parsing into feet while satisfying foot form is possible with (ˈH)(ˈLL).

(52) /HLL/ → (ˈH)(ˈLL); Maximal parsing possible without shortening

/HLL/	BAL	PARSESYLL	MAX-μ	ALLFTL
→ a. (ˈH)(ˈLL)				1
b. (ˈLL)L		**W** 1	**W** 1	**L**
c. (ˈHL)L	**W** 1	**W** 1		**L**

However, when a word begins with a heavy syllable and has an *odd* number of light syllables separating it from the word edge, it will show shortening of the heavy syllable, as shown in (53). In this case, maximal parsing is possible while obeying foot form only in candidate (a), in which shortening of the first heavy syllable has taken place to form light-syllable trochees (ˈLL)(ˈLL).

(53) /HLLL/ → (ˈLL)(ˈLL); Shortening for maximal parsing

/HLLL/	BAL	PARSESYLL	MAX-μ	ALLFTL
→ a. (ˈLL)(ˈLL)			1	2
b. (ˈHL)(ˈLL)	**W** 1		**L**	2
c. (ˈH)L(ˈLL)		**W** 1	**L**	2
d. (ˈH)(ˈLL)L		**W** 1	**L**	**L** 1

Meanwhile words in this hypothetical language show apparent left-to-right parsing of sequences of light syllables as shown in (54).

(54) /LLLLL/ → (ˈLL)(ˈLL)L; Apparent left-to-right parsing

/LLLLL/	BAL	PARSESYLL	MAX-μ	ALLFTL
→ a. (ˈLL)(ˈLL)L		1		2
c. L(ˈLL)(ˈLL)		1		**W** 4

These inputs show that a language with apparent left-to-right parsing must consider the parity of the syllables following an initial heavy syllable in order to know whether to parse it as a monosyllabic foot or to shorten and parse it as a disyllabic foot.

The prediction is in fact even slightly more complicated, as we must consider inputs with heavy syllables in places other than the initial syllable. When an input like /HLH/ in (55) is considered, for example, we see that it is not the initial heavy that shortens. Rather, it is the final one, as the winning candidate (a) shows. In this case there is competition between candidate (a), (ˈH)(ˈLL), and candidate (d), (ˈLL)(ˈH), which both feature perfect foot form, maximal parsing, and one violation of MAX-μ. Candidate (a) wins because it violates ALLFTL less. This behavior is familiar from the unattested language discussed in the previous section, which showed similar preferences due to alignment constraints (Crowhurst and Hewitt 1995a; further discussion in section 4).

(55) /HLH/ → (ˈH)(ˈLL); Shortening for maximal parsing[25]

/HLH/	BAL	PARSESYLL	MAX-μ	ALLFTL
→ a. (ˈH)(ˈLL)			1	1
b. (ˈHL)(ˈH)	**W** 1		**L**	**W** 2
c. (ˈH)(ˈLH)	**W** 1		**L**	1
d. (ˈLL)(ˈH)			1	**W** 2
e. (ˈH)L(ˈH)		**W** 1	**L**	**W** 2

Summarizing the statement of when to shorten heavy syllables in this language is not very straightforward, though the statement of when *not* to shorten is rather familiar from the previous section. The summary is provided algorithmically in (56).

(56) Algorithm for determining which heavy syllables to shorten.

- Parse a heavy syllable as a monosyllabic foot if it occurs in an odd numbered syllable of an odd-parity word and is the closest to the left edge of all heavy syllables satisfying this requirement.
- Shorten all other heavy syllables in the word (actually, in the language), and create (ˈLL) feet.

[25] This case would not be considered an example of *trochaic* shortening, because the potentially offending sequence would have been (ˈLH) rather than (ˈHL). This type of shortening is called *iambic* shortening, but it occurs exclusively in trochaic languages (e.g., Latin: Allen 1965, 1973; Prince 1990; Mester 1994). The constraint BALANCED militates indiscriminately against both types of unbalanced foot.

One obvious consequence of the statement in (56) is that even-parity words contain only light syllables, while odd-parity words may have at most one heavy syllable, necessarily occurring in an odd-numbered syllable. Thus, in this language, vowel length cannot be contrastive in even-parity words, even-numbered syllables of any word, nor in any syllable following a heavy syllable counting from left-to-right. The list in (57) shows some input-output pairs in this hypothetical language for illustration.[26]

(57) Inputs and outputs in hypothetical language

Odd-parity words	Even parity words
LLL → (ˈLL)L	LLLL → (ˈLL)(ˈLL)
HLL → (ˈH)(ˈLL)	HLLL → (ˈLL)(ˈLL)
HLH → (ˈH)(ˈLL)	LHLL → (ˈLL)(ˈLL)
LLH → (ˈLL)(ˈH)	LLHL → (ˈLL)(ˈLL)
LLHLL → (ˈLL)(ˈH)(ˈLL)	HLLHLL → (ˈLL)(ˈLL)(ˈLL)
HHHHH → (ˈH)(ˈLL)(ˈLL)	HHHH → (ˈLL)(ˈLL)

The generalization about stress and the distribution of vowel length (or syllable heaviness in general) in this hypothetical language is complicated and non-local, and yet parallel OT predicts this language using only the constraints from the previous section that were necessary for a standard account of Fijian. This language clearly demonstrates a non-local interaction between stress and shortening, because the number and position of light and heavy syllables throughout the word have to be taken into account in order to know whether to shorten a given heavy syllable.

This is something we would never expect to see in a natural language stress system, and HS/IFO correctly predicts it should not occur. This is because HS/IFO makes decisions about metrification one foot at a time, and when the dominant parsing mode is left to right (because ALLFTL >> ALLFTR) this means that the inputs /HLL/ and /HLLL/ will not look very different from the point of view of choosing optimal foot

[26] The formulation of the constraint BAL is non-trivial when it comes to spelling out these predictions. However, parallel OT will predict a language which makes non-local parsing decisions in the sense defined here regardless of how the (ˈHL)-dispreferring constraint is formulated, though its precise characteristics will vary depending on whether the constraint disfavors only (ˈHL) or both (ˈHL) and (ˈLH), as BAL does.

structure. Yet the hypothetical unattested language would require them to be treated differently, with /HLL/ being parsed as (ˈH)LL on the first iteration and /HLLL/ being first parsed as (ˈHL)LL, as in (58).

(58) a. HLL → (ˈH)LL → (ˈH)(ˈLL)
b. HLLL → (ˈHL)LL → (ˈLL)LL → (ˈLL)(ˈLL)[27]

Importantly, the reason these derivations cannot coexist in HS/IFO with these constraints is that the ranking required to produce (ˈH)LL at the first iteration would not also produce (ˈHL)LL. This is exactly as we saw in section 3.1.3, and the same illustration stands. Barring constraints that can indirectly simulate derivational look-ahead, which were argued not to exist in section 3.1.3, HS/IFO cannot analyze this unattested language.

The striking parallel between the hypothetical unattested language from section 3.1 and this hypothetical language should by this point be obvious. In the former, parallel OT's global evaluation chooses the best heavy syllable to parse as a monosyllabic foot; in an odd-parity word the leftmost heavy syllable will make the 'best' monosyllabic foot. In the latter, the global evaluation does the exact same thing, but it additionally simultaneously shortens all other heavy syllables in order to parse maximally while maintaining balanced feet. Thus, the problem that arises here is essentially the same as the one in the previous section with the addition of shortening. We can see then that a language may appear to go to great lengths to satisfy PARSESYLL in parallel OT, and that the results are often non-local in character.

3.2.3 Comparison to Hayes' (1995) analysis

Hayes (1995) presents a serial analysis of Fijian trochaic shortening that differs in crucial ways from the HS/IFO analysis presented above, despite the fact that both theories involve serial generation. In Hayes' framework foot building is not necessarily ordered with respect to other processes, but instead may apply 'persistently' (Hayes 1995:114). That is, in some languages, including Fijian, foot-building rules apply

[27] Or possibly HLLL → (ˈHL)LL → (ˈHL)(ˈLL) → (ˈLL)(ˈLL)

whenever they can, which effectively means both before and after shortening. The shortening rule is thus stated as in (59), resulting in the derivation in (60).

(59) Trochaic shortening rule (Hayes 1995: 146)[28]

$\sigma_{\mu\mu} \rightarrow \sigma_{\mu}$ / __ σ_i where σ_i is metrically stray

(60) Hayes (1995) Fijian shortening derivation

HL	
(ˈH)L	*footing*
(ˈL)L	*Rule (59); H syll shortens before stray syllable*
(ˈLL)	*persistent footing / incorporation*

It is not possible to recapitulate this derivation in HS/IFO, despite the fact that both theories employ sequential derivations. Persistent footing allows Hayes to first build a monosyllabic foot on the heavy syllable and then later incorporate the following light syllable, after the heavy syllable is shortened. In HS/IFO with the assumption of strict inheritance, (ˈH)L will not lead to (ˈLL), so the same path is not an option. In addition, HS/IFO requires that processes be motivated by plausibly universal markedness constraints rather than language-specific rules. Even if strict inheritance were not assumed, there would be no plausible motivation for shortening the heavy syllable of (ˈH)L to create (ˈL)L, which is a poor foot. As we have already seen, the derivation has no ‘look-ahead’ and would not tolerate shortening to (ˈL)L in order to eventually get (ˈLL). Thus, Hayes’ analysis cannot be repeated in HS/IFO, with or without strict inheritance.

In our analysis the motivation for shortening is an unbalanced trochee, which requires that the unbalanced trochee be built before shortening can occur. In contrast, Hayes’ theory cannot easily allow (ˈHL) feet to be created by foot-building rules because such feet are argued in the larger theory not to exist in moraic trochee languages. In HS/IFO, foot form constraints are violable, rather than absolute. Thus, there is no cost

[28] Hayes does not include reference to the right-edge word boundary in his rule because he wants a general rule for both medial and final HL-shortening.

associated with building a foot that is ‘marked’ in the Jakobsonian sense, unless the constraints and their ranking are poised to prevent the building of such feet. If all constraints that disprefer such feet are ranked below a constraint that is best satisfied by creating them, they will be created. This was also an implicit feature of the analysis of English trisyllabic shortening by Prince (1990).

An interesting feature of Hayes’ analysis, however, is that he proposes that the ultimate motivation for the shortening process in Fijian is a kind of metrical parsing maximization. By this account, vowel shortening permits the sequence …(ˈLL)#, which is more maximally parsed than …(ˈH)L#, the only other option in a theory with a fixed foot inventory because of the absolute ban on unbalanced trochees. This argument is rather parallel in character. To formalize the notion of shortening *in order to* parse maximally would require that the result of parsing and the result of shortening be evaluated at the same time. In Hayes’ framework this is not an option, and instead he states the shortening rule formally as was shown in (59). Although this rule does not reference parsing maximization directly, persistent footing ensures that the output of rule (59) is reparsed into an (ˈLL) foot, and the existence of the rule is argued, meta-theoretically, to follow from a cross-linguistic tendency to maximize the number of syllables parsed into feet.

If the process of trochaic shortening in Fijian were successfully argued to follow from parsing maximization, a parallel theory would seem to have an obvious advantage in capturing this, for the reasons just cited. HS/IFO is in no better position to capture the generalization behind Hayes’ intuition than Hayes’ own theory is, since the processes of shortening and metrification occur sequentially in HS/IFO. In contrast, parallel OT compares metrification and shortening (and everything else) in parallel, making obvious what the consequences of shortening would be for foot building. It might seem then that the fate of HS/IFO as a model of metrification rests on whether trochaic shortening is indeed motivated in order to maximize parsing, as Hayes suggests. But there are several reasons to believe that it is not (at least not directly), and to consequently favor the serial analysis of trochaic shortening over a parallel one.

First, the process may happen throughout the word, but it is obligatory in a word-final HL sequence. It seems unlikely that a constraint preferring parsing maximization would somehow reference the word edge. In our analysis we relied on the assumption that languages may require a foot to align with the word edge, formalized with ALIGNWDR/L, but this constraint is equally satisfied by a foot of any size. Second, for a process to be truly motivated by a maximal parsing preference, it would seem that PARSESYLL or some equivalent should be the motivating constraint. To implement Hayes' suggestion in parallel OT, we can thus assume that PARSESYLL should be ranked over MAX-μ so that an /…HL#/ sequence becomes …(ˈLL)#. However, the same ranking causes trochaic shortening to occur in medial sequences of /HL/ as well, and may in fact lead to the undesirable prediction discussed in the previous section if additional analytic tools (e.g., other constraints) are not employed. Indeed, this ranking produces *true* shortening-for-parsing-maximization, and we have seen that the results are not a language we want our theory to predict. The results suggest that the explanation provided by Hayes to account for the rule in (59) does not reflect a true generalization of natural language stress systems.

3.2.4 Summary of stress and shortening

This section has essentially replicated the results of the previous section. Admitting a standard parallel OT analysis of Fijian also, by factorial typology, admits patterns which are nominally similar yet unattested. Thus, although the notion of shortening to maximize parsing cannot be captured in a serial theory, it appears that it cannot be constrained in a parallel one. At the heart of these predictions are the same characteristics of parallel OT we found section 3.1 – parallel evaluation permits non-local interactions because of its ability to consider all possible metrical parses at one time. On the other hand, the account of trochaic shortening in HS/IFO was shown to be entirely satisfactory for analyzing Fijian because it supplies a derivation whose components evoke well-attested output preferences, and this theory does not predict these unattested non-local interactions.

3.3 Stress and edge restrictions

The previous two sections showed that standard stress constraints have unwanted predictions in a parallel theory; two problems with essentially the same etiology were explored – high-ranking PARSESYLL can enforce odd metrical parsing (section 3.1) and unattested patterns heavy syllable shortening (section 3.2). Iterative foot optimization cannot reproduce these predictions because they would require the derivation to be able to look ahead to influence early parsing decisions. In the present section we turn to another locus of non-local stress systems in parallel OT: the interaction of edge constraints with metrical rhythm and foot form.

It is not uncommon for languages to interrupt regular parsing at the right edge of a word. Extrametricality can target segments, syllables, or (arguably) feet, and has been typically treated in Optimality Theory with constraints from the NONFINALITY family, discussed by Prince and Smolensky (1993/2004). Languages that parse words into iambic feet are at a particular risk for stressing the final syllable, since an iambic foot that lines up with the right edge of a word would do just that. Such languages employ various strategies for avoiding a final stress, with 'rhythmic reversal' being a common choice. In this process, the final two syllables in a word are parsed as a trochee rather than the usual iamb in order to satisfy a constraint that militates against final stressed syllables. Prince and Smolensky (1993/2004) cite Choctaw, Munsee, Southern Paiute, Ulwa, and Axininca as languages that show this process, and Aguaruna has a similar restriction before syncope of unstressed vowels, according to McCarthy (to appear-a).

Importantly, the languages exhibiting this process show that the effects of this constraint are felt locally, in that the word-final foot is the only one that undergoes rhythmic reversal. In this section I will illustrate with data from the Apurucayali dialect of Axininca, showing that HS/IFO and parallel OT can both deal with local rhythmic reversal for edge restrictions. However, as with previous sections, parallel OT additionally predicts that such effects may permeate the word, resulting in non-local interactions not predicted by HS/IFO. Evidence for languages with such non-local responses to edge constraints is not convincing, as the only know case is shown to be problematic in section 3.3.3.

3.3.1 Local rhythmic reversal in HS/IFO

Stress in the Apurucayali dialect of Axininca (Payne, Payne, and Santos 1982; McCarthy and Prince 1993b), an Arawakan language of Peru, demonstrates a kind of local interaction in which although the language prefers iambic feet, it may parse the final two syllables of a word into a trochee if building an iamb would cause the foot's prominence to fall on a word-final syllable.[29] The data in (61) illustrate. In (a) are words with a left-to-right iambic parse that leave a stray syllable at the right edge. In (b) and (c) are words that would stress the final syllable if an iamb appeared at the right edge, but a trochee appears instead.

(61) Axininca stress (Payne, Payne, and Santos 1982: 188-9, 193)[30]

a. (ʧʰoˈri)na	'species of palm'
(iˈʧʰi)(kaˈki)na	'he has cut me'
b. (kiˈmi)(ˈtaka)	'perhaps'
(hoˈti)(ˈtana)	'he let me in'
c. (ˈsari)	'macaw'
(ˈkito)	'shrimp'

This pattern can be straightforwardly analyzed in HS/IFO with a derivation building left-aligned iambs except when doing so would stress the final syllable, in which case iambic foot form is sacrificed in order to avoid the final stress. The constraint we will employ to militate against final stress is NONFINALITY(STRESS) defined in (62). We will also assume PARSESYLL >> ALLFTL >> ALLFTR in order to enforce iterative left-to-right parsing, and IAMB >> TROCHEE in order to have a default preference for right-headed feet. Finally, we also assume that FTBIN(μ) is undominated, since feet surfacing in Axininca contain at least two moras.

(62) **NONFINALITY(STRESS)**: Assign one violation mark for a word whose final syllable is stressed.

[29] The pattern of trochaic reversal alternates with simply leaving the final two syllables unfooted; which of these two options wins appears to be based in part on the segmental make-up of the preceding stressed syllable and the target vowel, but may also be partly random in words in which the appropriate conditions are met (see Payne, Payne, and Santos 1982: 193 for discussion).

[30] Degrees of stress ignored; which syllable is the primary stress is based on a somewhat complex calculation between the last two feet based partly on a prominence scale. See Payne, et al. (1982), or Hayes (1995) for a discussion of prominence in the Pichis dialect.

A derivation of a word with an even number of light syllables in Axininca will proceed by building iambs non-finally until the only option for building an iamb involves stressing the last syllable, and it will then build a trochee on the two remaining syllables. An example is shown in (63). Candidate (b) wins in the first iteration because regular iambic parsing is assumed. Then candidate (f), which is the eventual winner, wins in the second iteration because high-ranked NonFin(Str) prevents regular iambicity at the right word edge, and the ranking of ParseSyll over Iamb means that it's more important to parse words into feet than it is for the feet to be iambic, so (d) and (g) are losers (candidate (g) is also ruled out by high-ranking FtBin(μ)). The third iteration shows convergence because no more parsing is possible.

(63) Axininca /LLLL/ → (LˈL)(ˈLL)

/hotitana/ 'he let me in' **1st iteration**	NonFin(Str)	ParseSyll	AllFtL	Iamb	Trochee
a. hotitana		**W** 4			**L**
b. (hoˈti)tana		2			1
c. (ˈhoti)tana		2		**W** 1	**L**
2nd iteration					
d. (hoˈti)tana		**W** 2	**L**	**L**	1
e. (hoˈti)(taˈna)	**W** 1		2	**L**	**W** 2
f. (hoˈti)(ˈtana)			2	1	1
g. (hoˈti)(ˈta)na		**W** 1	2	**L**	1
3rd iteration					
h. (hoˈti)(ˈtana)			2	1	1

Output: (hoˈti)(ˈtana)

The derivation of an odd-parity word will proceed as shown in (64). The derivation begins by building an iamb at the left edge, just as in the previous derivation. It continues building iambs until the third iteration when only one unfooted syllable remains at the right edge, and in this case the syllable is left unparsed because parsing it violates NonFin(Str) (and also high-ranking FtBin(μ)), and (e) emerges as the winner.

(64) Axininca /LLLLL/ → (LˈL)(LˈL)L

/(iˈʧʰi)(kaˈki)na/ ‘he has cut me’ **1st iteration**	NONFIN(STR)	PARSESYLL	ALLFTL	IAMB	TROCHEE
a. (iˈʧʰi)kakina		3			1
b. (ˈiʧʰi)kakina		3		**W** 1	**L**
2nd iteration					
c. (iˈʧʰi)kakina		**W** 3	**L**		**L** 1
d. (iˈʧʰi)(kaˈki)na		1	2		2
3rd iteration					
e. (iˈʧʰi)(kaˈki)na		1	2		2
f. (iˈʧʰi)(kaˈki)(ˈna)	**W** 1	**L**	**W** 6		2

Output: (iˈʧʰi)(kaˈki)na

As this illustration has indicated, it is rather easy to account for Axininca’s local rhythmic reversal in HS/IFO. The constraint against final stressed syllables becomes relevant when parsing the final syllables of a word, and it enforces its preference then. The next section shows that parallel OT can also produce this pattern as well, but ranking permutation predicts other patterns which are only marginally attested.

3.3.2 Stress and edge restrictions in parallel OT

Following Prince and Smolensky’s (1993/2004:65) analysis of a very similar process of rhythmic reversal in Southern Paiute, it is possible to analyze the same process in Axininca in parallel OT. The same constraints and ranking from the HS/IFO analysis will suffice. The tableau in (65) illustrates with an even-parity word. Candidate (c) emerges as the winner because it satisfies NONFIN(STR) by violating IAMB, and the same ranking will easily also account for odd-parity words (tableau omitted).

(65) Axininca local rhythmic reversal in parallel OT

/hotitana/ 'he let me in'	NONFIN(STR)	PARSESYLL	ALLFTL	IAMB	TROCHEE
a. hotitana		**W** 4	**L**	**L**	**L**
b. (hoˈti)(taˈna)	**W** 1		2	**L**	**W** 2
→ c. (hoˈti)(ˈtana)			2	1	1
d. (hoˈti)(ˈta)na		**W** 1	2	**L**	1
e. (ˈhoti)(ˈtana)			2	**W** 2	**L**

This analysis appears to be on a par with the one provided in the previous section for accounting for local reversal in Axininca. However, as with previous illustrations, parallel OT predicts a non-local counterpart to the kind of local rhythmic reversal illustrated here. The winning candidate in tableau (65) features a stress clash, a dispreferred sequence of two stresses. With a constraint such as *CLASH highly ranked, (c) would be out on these grounds. The ranking of the other constraints can conspire to create the prediction illustrated in (66), that despite the ranking of IAMB >> TROCHEE, the winning output is parsed into trochees. This effect is non-local, because a constraint referencing the edge of the domain has affected the parsing at the opposite end of the word, in opposition to the ostensible direction of parsing in the language.

(66) Pseudo-Axininca: Non-local rhythmic reversal

/hotitana/ 'he let me in'	*CLASH	NONFIN(STR)	PARSESYLL	IAMB	TROCHEE
a. hotitana			**W** 4	**L**	
b. (hoˈti)(taˈna)		**W** 1		**L**	**W** 2
c. (hoˈti)(ˈtana)	**W** 1			**L** 1	**W** 1
d. (hoˈti)(ˈta)na	**W** 1		**W** 1	**L**	**W** 1
→ e. (ˈhoti)(ˈtana)				2	
f. (hoˈti)tana			**W** 2		**W** 1

This same ranking produces standard iambs in a word ending with an odd number of light syllables, because in such words iambic rhythm can be preserved without the potential for stressing the final syllable. This is shown in (67).

(67) Pseudo-Axininca: Otherwise iambs

/itʃʰikakina/ 'he has cut me'	*CLASH	NONFIN(STR)	PARSESYLL	IAMB	TROCHEE
→ a. (iˈtʃʰi)(kaˈki)na			1		2
b. (iˈtʃʰi)(ˈkaki)na	**W** 1		1	**W** 1	**L** 1
c. (ˈitʃʰi)(ˈkaki)na			1	**W** 2	**L**

Thus, the language predicted under this ranking is one in which odd-parity words have stress on even-numbered syllables, while even-parity words have stress on odd-numbered syllables. This stress system would superficially look like right-to-left syllabic trochees, but we can assume that a language showing true iambic parsing as in (67) will show the standard correlates of iambic rhythm, such as iambic lengthening (Hayes 1995), in odd-parity sequences, though not in even-parity ones. To give another example, in the local rhythmic reversal of Axininca, a long vowel in a final syllable is shortened when it is unstressed because of rhythmic reversal or when it is the stray unfooted syllable at the edge of the word.[31] We might then expect such de-lengthening to occur throughout even-parity words in the hypothetical language in (66) and (67), since even-parity words will be parsed exclusively into trochees.

This prediction cannot be obtained in HS/IFO because the effects of a constraint like *CLASH cannot permeate the word in the *opposite* direction of footing, which is what this hypothetical language shows. Because footing decisions are assumed to be irrevocable, a candidate in which an input iamb becomes an output trochee are not considered. However, even if we were to allow such a candidate we could not get the unattested prediction, because getting rid of *CLASH violations by iteratively reversing foot form will not result in harmonic improvement in words of more than two feet. This is illustrated in (68) and (69) below. As (68) shows, if *CLASH >> PARSESYLL as in the parallel OT example above, the clash-inducing foot is not even built in HS/IFO. But if PARSESYLL >> *CLASH, as in (69), then the candidate in (f) wins at the third iteration. But permeating backwards through the word is not possible in words longer than two feet, even when strict inheritance is not assumed because no net benefit is realized by

[31] This process does not apply to diphthongs, which are always stressed, even word-finally. This might suggest that diphthongs are heavier than long vowels, attracting stress more aggressively. See Kager (1989: Chapter 3) for arguments that syllables with diphthongs in Dutch are heavier than those with long vowels.

doing so iteratively. As can be seen by comparing candidates (g) and (h) in (69), one *CLASH violation is traded for another at the expense of an additional IAMB violation, which causes the derivation to converge on candidate (g) as the winner, without global rhythmic reversal.[32]

(68) *CLASH >> PARSESYLL; Final foot not created

/σσσσσσ/ **1st iteration**	NONFIN(STR)	*CLASH	PARSESYLL	IAMB	TROCHEE
a. σσσσσσ			**W** 6		**L**
b. (σˈσ)σσσσ			4		1
2nd iteration					
c. (σˈσ)σσσσ			**W** 4		**L** 1
d. (σˈσ)(σˈσ)σσ			2		2
3rd iteration					
e. (σˈσ)(σˈσ)σσ			2		2
f. (σˈσ)(σˈσ)(ˈσσ)		**W** 1	**L**	**W** 1	2
g. (σˈσ)(σˈσ)(σˈσ)	**W** 1		**L**		**W** 3

Output: (σˈσ)(σˈσ)σσ

(69) PARSESYLL >> *CLASH; Final foot present, but no global reversal

/σσσσσσ/ **1st iteration**	NONFIN(STR)	PARSESYLL	*CLASH	IAMB	TROCHEE
a. σσσσσσ		**W** 6			**L**
b. (σˈσ)σσσσ		4			1
2nd iteration					
c. (σˈσ)σσσσ		**W** 4			**W** 1
d. (σˈσ)(σˈσ)σσ		2			2
3rd iteration					
e. (σˈσ)(σˈσ)σσ		**W** 2	**L**	**L**	2
f. (σˈσ)(σˈσ)(ˈσσ)			1	1	2
4th iteration					
g. (σˈσ)(σˈσ)(ˈσσ)			1	1	2
h. (σˈσ)(ˈσσ)(ˈσσ)			1	**W** 2	**L** 1

Output: (σˈσ)(σˈσ)(ˈσσ)

[32] The prediction that rhythmic reversal could occur in words of two feet (but not longer) is not something we would want to predict either. This can be seen as constituting an argument for strict inheritance, which prevents any reversal from occurring.

Thus, the ability of *CLASH to create non-local stress systems by interacting with non-finality is quite limited in HS/IFO. This is in clear contrast to parallel OT, which optimizes over all possible metrical parses and will thus choose winners that show the effects of edge constraints non-locally. In the next section I turn to a language which has been argued to exhibit a non-local interaction of this kind, and I show that the stress system of this language resists coherent analysis even in a parallel theory.

3.3.3 Yidiɲ

The Australian language Yidiɲ (Dixon 1977ab) is often thought of as the canonical example of a language whose metrical phonology seems to require a parallel analysis (e.g., McCarthy 2002: 149-152). Indeed, Hung (1993, 1994) analyzes Yidiɲ stress as conforming to a generalization essentially similar to the one provided in the previous section as a prediction of parallel OT – stress generally avoids the final syllable and always alternates by syllables, making it possible to analyze the language as having undergone non-local rhythmic reversal in order to satisfy non-finality and *CLASH or some equivalent, as in Hung's RHYTHM constraint, which requires that every stressed syllable be followed by an unstressed one.

Dixon's description of Yidiɲ stress derives all sources of vowel length before stress is assigned, making the stress statement entirely dependent on the location of long vowels in the word. This has led other theorists to state the generalization along the lines of 'parse a word into trochees unless a long vowel appears in an even-numbered syllable, then parse the word into iambs', which is also a non-local description. This is essentially how the generalization is couched in Hayes (1980: 201) and Hayes (1995: 260).[33] However, odd-parity words in Yidiɲ *always* meet the criterion for being parsed into iambs, which is not accounted for under this alternative except by first describing the distribution of long vowels. Due to a process of penultimate lengthening, the penult (an even-numbered syllable) in odd-parity words nearly always contains a long vowel; thus,

[33] The justification for the process is cited by Hayes (1995) as the Iambic/Trochaic Law. The presence of length (or weight) favors iambic parsing due to the law, because a long vowel in an even-numbered syllable allows for the construction of at least one canonical iamb, i.e. (LˈH). And the presence of trochees in the absence of length conforms to trochees' desire to group elements of equal weight, also according to the law. However, even-parity words with trochaic rhythm may show length contrasts in the stressed syllables in this language, contra the Iambic/Trochaic Law.

identifying the target of lengthening would also seem to require having metrical structure available.

In either case, the construal of Yidiɲ's stress system seems to be that it relies on a view of the entire metrical parse in order for stress to be properly assigned, since whether a word is parsed into iambs or into trochees depends on the parity of the entire word and the distribution of its heavy syllables. The generalization in (70) covers most words of Yidiɲ.

(70) Stress in Yidiɲ

- Even-parity words are parsed into disyllabic trochees unless an even-numbered syllable contains a long vowel, in which case the word is parsed into iambs.
- Odd-parity words are always parsed into iambs because the penultimate syllable (an even-numbered syllable) contains a long vowel almost without exception.

Despite the non-local character of the stress statement in this language, no extant phonological analyses dealing with all the facts of Yidiɲ's stress system can be found in parallel OT. A primary source of problems is that the stress pattern is opaque in many even-parity words. The main source of non-lexical vowel length in Yidiɲ is the process of penultimate lengthening in odd-parity words. Many stems that are analyzed as being underlyingly odd-parity undergo penultimate lengthening and then deletion of the final syllable, leaving a word with a form like *gindáːn* 'moon' in the absolutive (bare) form, which surfaces as the trisyllabic stem *gindanu-* in the remainder of the paradigm (Dixon 1977a: 13). The vowel length in such stems and the concomitant stress on the second syllable derives exclusively from the underlying parity of the stem, no longer visible on the surface. Parallel OT's difficulty with opacity is well-known, and none of the phonological analyses of Yidiɲ in parallel OT currently known to me attempt to deal with this problem.[34]

[34] An exception is Hung (1994a), though her use of containment faithfulness is what allows an analysis, and it cannot be restated within correspondence theory (McCarthy and Prince 1995).

Partly in response to this stubborn detail Hayes (1997/99) proposes a rather radical reanalysis of Yidiɲ's morphophonology which involves doing away with a single underlying form and taking the bare absolutive form as the base for paradigmatic alternations. He thus argues that what appears to be lengthening and deletion can rather be analyzed in the reverse, as shortening and epenthesis. He presents evidence that the third vowel in such stems is actually predictable, obviating the need to store it in an underlying form. Shortening is then motivated by non-phonotactic constraints, which require a type of anti-faithfulness among paradigm members. This solution has much to recommend it, though it is in large part morphological. It thus remains to be seen to what extent Yidiɲ's stress system might present a problem of non-locality once the relationship among paradigm members and constraints mediating this relationship are better understood.

Thus we have seen that a language which has been put forward as evidence of parallelism in metrical structure building is itself not straightforwardly analyzed in a parallel theory without significant changes. The stress system of Yidiɲ deserves more work, which will no doubt lead to additional insights, adding to those that Dixon (1977ab), Nash (1979), Hayes (1980, 1982, 1997/99), Halle and Vergnaud (1987), Kirchner (1992), Hagberg (1993), Crowhurst and Hewitt (1995b), Hung (1993, 1994), and others have provided us. But the preliminary evidence suggests that we are justified in not giving up on the potential of iterative foot optimization in response to the apparent non-local character of Yidiɲ's stress system.

3.3.4 Summary of stress and edge restrictions

In sum, with respect to local versus non-local interactions of stress with constraints like NONFINALITY, the evidence we have suggests that HS/IFO is correct in its predictions for the existence of local process, such as rhythmic reversal in Axininca, and the non-existence of non-local processes, which have not been unequivocally demonstrated.

3.4 Summary of locality in stress

We find in all of the cases explored in this section that commonplace interactions in stress systems are local, while non-local interactions are not clearly attested, matching HS/IFO's predictions. Parallel OT on the other hand predicts the existence of unattested stress systems with myriad non-local interactions, using the same mechanisms necessary for it to account for the attested local processes. The only example of a seemingly non-local stress system was shown to be questionable, and until clearer examples emerge, the evidence suggests that we are justified in preferring HS/IFO as a theory of metrical structure and its interaction with weight, quantitative adjustments, and edge restrictions.

4 Other consequences

The previous section argued for an iterative approach to stress on empirical grounds: harmonic serialism with iterative foot optimization makes strong predictions about what kinds of stress systems we would expect, and those predictions appear to hold up. This section explores some of the formal consequences of adopting this model. I first discuss the properties of PARSESYLL in HS/IFO and then turn to why alignment constraints are appropriate for determining foot placement in this model.

4.1 PARSESYLL in HS/IFO

In standard stress analyses in parallel OT, PARSESYLL is the constraint that compels foot-building, and the analyses in this paper have followed suit. Its definition is repeated here in (71) (Prince and Smolensky 1993/2004).

(71) **PARSESYLL**: Assign one violation mark for each syllable that is not a member of some foot.

Although the standard definition is unchanged in the present model, PARSESYLL has a somewhat different character in HS/IFO than in parallel OT. We saw in sections 3.1 and 3.2 that high ranking PARSESYLL was partly responsible for many of the unwanted typological predictions made by parallel OT and that its preferences could compel non-local interactions with parallel evaluation. The same predictions were not made by HS/IFO, because the architecture of serial optimization does not allow these interactions using this constraint. There are additional differences though. In particular, because feet are built one at a time in HS/IFO, PARSESYLL will always favor building the largest foot possible because a larger foot leaves fewer syllables unparsed. In parallel OT, PARSESYLL does not necessarily have this property; the hypothetical forms (σ)(σ)(σ)(σ) and (σσ)(σσ) both maximally satisfy PARSESYLL, though the first does not feature maximally large feet.

The idea of preferring the largest possible foot in a stress derivation is not new, though this preference is not typically thought of as arising from the constraint that compels foot building. Rule-based theories of metrical parsing typically include a clause that requires maximal expansion of the foot (Hayes 1995: 102-3; Halle and Vergnaud 1987:15; Prince 1980). The directional parsing algorithms of rule-based theories required a clause to this effect to ensure that the largest licit foot was built even when a smaller foot would satisfy parameterized requirements on foot form. This ‘maximality condition’ ordered the foot-building rule to look around to make sure it was not settling for a smaller foot. Here the ability of PARSESYLL to favor the maximal licit foot expansion plays a similar role as the maximality condition from rule-based metrical parsing. A difference however is that PARSESYLL is a violable constraint and its preferences may go unheeded if higher ranking constraints disagree, while the maximality condition was meant to be an inviolable principle of foot building. On this point we again find a similarity between HS/IFO and Prince (1990)’s Harmonic Parsing. Prince states that the maximality condition should not be retained because building the *best* foot should replace building the *biggest*. In the constraint-based HS/IFO we also choose the best foot given the constraint ranking, but if PARSESYLL is high enough ranked, the best foot will also be one of the biggest feet.

One might wonder whether PARSESYLL is duplicating the role of the constraints intended to enforce minimum foot size, FTBIN(μ) and FTBIN(σ), or vice versa, but it turns out that both kinds of constraints are needed. First, FTBIN constraints do not compel foot-building; they can only assess a particular foot for its size. In fact, under some rankings they may discourage foot-building if all the feet that meet their requirements are ruled out for other reasons. Second, the preferences of FTBIN(μ) and PARSESYLL may overlap when light syllables are at issue, in which case both constraints will prefer a disyllabic foot, but FTBIN(μ) is equally satisfied by a heavy syllable in its own foot, while PARSESYLL still prefers something larger, if possible. FTBIN(σ) and PARSESYLL overlap much more often in their preferences, both preferring disyllabic feet when possible,[35] but these constraints can also conflict, and the resolution of their conflict is important for modeling whether a language allows monosyllabic feet at all. In the case of Pintupi, the syllabic trochee language analyzed in section 2.3, the ranking FTBIN(σ) >> PARSESYLL ensures that the final syllable in odd-parity words is not parsed.

The ability of PARSESYLL to prefer a larger foot was already shown to be important in the analysis of the generalized trochee languages presented in section 3.1. Such languages prefer to build disyllabic feet but will allow a monosyllabic foot when a single heavy syllable is the only thing left in the word. This was captured in the HS analysis precisely with PARSESYLL's preference for larger feet. A representative example from the discussion of generalized trochees in section 3.1 is shown in (72). Although candidate (a) is not ruled out by FTBIN(μ) because the first syllable is heavy, candidate (b) is preferred by PARSESYLL because it gets rid of more unparsed syllables. Since PARSESYLL has to outrank FTBIN(σ) to allow monosyllabic feet at all, FTBIN(σ) is not the deciding factor.

(72) PARSESYLL prefers maximal expansion; Weṛgaia /HLL/ → (ˈHL)L

/delguna/ 'to cure' **1st iteration**	FTBIN(μ)	PARSESYLL	FTBIN(σ)
a. (ˈdel)gu.na		**W** 2	**W** 1
→ b. (ˈdel.gu)na		1	

This behavior of PARSESYLL is an interesting feature of present framework. It produces the right results for generalized trochee languages and reproduces a condition

[35] The situation would change a bit if GEN were allowed to build feet of more than two syllables. In this case the relationship between FTBIN(σ) and PARSESYLL would be more like that of FTBIN(μ) and PARSESYLL.

on metrical parsing that other theories state separately. And happily, it is prevented by the framework of HS/IFO from inducing problematic non-local interactions.[36]

4.2 Gradient Alignment in HS/IFO

The analyses in this paper have also employed alignment constraints – specifically ALLFTL and ALLFTR – and have assumed their standard definitions with 'gradient' assessment, repeated here in (73) and (74). These constraints come from the Generalized Alignment family (McCarthy and Prince 1993a).

(73) **ALLFTL**: For each foot in a word assign one violation mark for every syllable separating it from the left edge of the word.

(74) **ALLFTR**: For each foot in a word assign one violation mark for every syllable separating it from the right edge of the word.

In OT, directionality – whether stress is assigned left to right or right to left – does not exist as such, but instead emerges from the preferences of constraints on the placement of feet. Alignment constraints do this by preferring all feet to be as far to one edge of the word as possible. While this restriction is never fully satisfied in words with more than one foot (that is, when PARSESYLL >> ALIGNMENT), the result nonetheless resembles a directional parse, as shown in (75).

(75) a. Iterative parsing when ALLFTL >> ALLFTR: (σσ)(σσ)σ
b. Iterative parsing when ALLFTR >> ALLFTL: σ(σσ)(σσ)

The story is similar in HS/IFO, in that directionality is again not overtly specified, but is also emergent, as was discussed in section 2.4.1. As in parallel OT, alignment constraints in HS assign a number of violation marks to each foot based on how far away from the word edge it is, and the violation marks for the feet are added up to determine

[36] See McCarthy (to appear-a) for a discussion of a possible alternative constraint that compels foot building in HS. The proposed alternative shares with PARSESYLL this property of preferring larger feet when feet are built by iterative optimization.

the number of violation marks that are assigned to a candidate. The difference, of course, is that HS/IFO effectively looks at only one additional foot at a time.

Although a number of criticisms of gradient alignment in parallel OT have surfaced, and some proposals have been made for doing away with such constraints (McCarthy 2003; Kager 2001), in this section I will illustrate that alignment constraints are well-suited to determining foot placement in HS/IFO. Alignment constraints are shown to behave better in a serial model than a parallel one and to be the best among logical alternatives for determining foot placement in HS.

4.2.1 Alignment is better in HS than in parallel OT

Crowhurst and Hewitt (1995a) observe that in parallel OT the apparent parsing directionality that emerges in a language when monosyllabic feet are allowed requires a ranking of ALLFTL and ALLFTR opposite from what we would expect based on (75). This is illustrated in (76); when ALLFTL outranks ALLFTR for example, the candidate with apparent left-to-right parsing wins if monosyllabic feet are prohibited (candidate (76)a), but the candidate with apparent right-to-left parsing wins if monosyllabic feet are allowed (candidate (76)d). That is, whether the ranking ALLFTL >> ALLFTR results in a left to right parse depends on the ranking of PARSESYLL and FTBIN(σ), because their ranking indicates whether monosyllabic feet are allowed in the language. Similarly, the apparent left-to-right parse in (a), with no monosyllabic foot, requires the ranking ALLFTL >> ALLFTR, while the apparent left-to-right parse in (c), with a monosyllabic foot, would require the opposite ranking.

(76) Foot alignment and monosyllabic feet (Crowhurst and Hewitt 1995a)

			Apparent direction	ALLFTL marks	ALLFTR marks
Monosyllabic feet prohibited	a.	(σσ)(σσ)σ	L→R	0+2 = **2**	1+3 = **4**
	b.	σ(σσ)(σσ)	R→L	1+3 = **4**	0+2 = **2**
Monosyllabic feet allowed	c.	(σσ)(σσ)(σ)	L→R	0+2+4 = **6**	1+3 = **4**
	d.	(σ)(σσ)(σσ)	R→L	0+1+3 = **4**	2+4 = **6**

This reversal is unexpected from the point of view of rule-based directional parsing, which treats (σσ)(σσ)σ and (σσ)(σσ)(σ) as a class, as both are the result of starting with foot-building at the left and iterating rightward. The gradient foot alignment constraints instead treat (σσ)(σσ)σ and (σ)(σσ)(σσ) as a class, because both are derived from the same ranking, ALLFTL >> ALLFTR.

In HS/IFO, alignment constraints do not show this behavior. Instead, under the ranking ALLFTL >> ALLFTR, apparent left to right footing emerges, and vice versa for ALLFTR >> ALLFTL, regardless of whether monosyllabic feet are allowed. The derivation in (77) shows that ALLFTL >> ALLFTR produces (σσ)(σσ)σ, which has apparent left to right parsing, when monosyllabic feet are prohibited, while the derivation in (78) shows that the same ranking of the alignment constraints produces (σσ)(σσ)(σ), also with apparent left to right parsing, when monosyllabic feet are allowed. Thus, the effect of left to right foot parsing emerges in HS/IFO with the ranking ALLFTL >> ALLFTR regardless of whether monosyllabic feet are allowed, contrary to the outcome in parallel OT. (And vice versa for ALLFTR >> ALLFTL.)

(77) ALLFTL >> ALLFTR, monosyllabic feet prohibited

/σσσσσ/ **1st iteration**	FTBIN(σ)	PARSESYLL	ALLFTL	ALLFTR
a. σσσσσ		**W** 5		
→ b. (σσ)σσσ		3		3
c. σ(σσ)σσ		3	**W** 1	**L** 2
2nd iteration				
d. (σσ)σσσ		**W** 3	**L**	**L** 3
→ e. (σσ)(σσ)σ		1	2	4
f. (σσ)σ(σσ)		1	**W** 3	**L** 3
3rd iteration				
→ g. (σσ)(σσ)σ		1	2	4
h. (σσ)(σσ)(σ)	**W** 1	**L**	**W** 6	4

Output: (σσ)(σσ)σ

(78) ALLFTL >> ALLFTR, monosyllabic feet allowed

/σσσσσ/ **1st iteration**	PARSESYLL	FTBIN(σ)	ALLFTL	ALLFTR
a. σσσσσ	**W** 5			
b. (σσ)σσσ	3			3
c. σ(σσ)σσ	3		**W** 1	**L** 2
2nd iteration				
d. (σσ)σσσ	**W** 3		**L**	**L** 3
e. (σσ)(σσ)σ	1		2	4
f. (σσ)σ(σσ)	1		**W** 3	**L** 3
3rd iteration				
g. (σσ)(σσ)σ	**W** 1	**L**	**L** 2	4
h. (σσ)(σσ)(σ)		1	6	4

Output: (σσ)(σσ)(σ)

The reason for this difference is that monosyllabic feet in non-moraic-trochee languages are a kind of last resort. Such feet are never the only kind of foot allowed in the language, but instead arise under duress – usually at the end of a parse in an odd-parity word when the choice is either to create a monosyllabic foot or not parse the syllable. These languages prefer to build canonical disyllabic feet when possible, because these are more harmonic, other things being equal, but allow monosyllabic feet when necessary to fully parse the word. Under the ranking in tableau (78), HS/IFO accounts for this restriction by choosing a disyllabic foot as optimal at intermediate iterations when there are enough syllables to create one, because PARSESYLL prefers maximal feet as discussed in the previous section. Even though monosyllabic feet are allowed, they are not preferred over disyllabic feet in the absence of high-ranked constraints penalizing disyllabic feet,[37] and thus monosyllabic feet are not chosen unless a disyllabic foot is not a possible option. This is essentially the same reasoning we saw with generalized trochee languages in section 3.1, here extended to languages that would allow any kind of syllable (i.e., heavy or light) in a monosyllabic foot. Again the character of an HS derivation is that it has no foresight – it does not 'know' that building

[37] And importantly, if there were constraints penalizing disyllabic feet, they would apply equally in even- and odd-parity words. In the language(s) at issue here only odd-parity words show monosyllabic feet, while even-parity words are happily parsed into disyllabic feet.

a monosyllabic foot at the outset would allow ultimately better satisfaction of ALLFTL at a subsequent iteration.

So far this illustration has simply highlighted a difference with respect to the preferences of alignment in parallel OT and HS/IFO. Despite the fact that parallel OT patterns differently from both HS and rule-based directional parsing in what languages it groups together by parsing 'direction', it may not be immediately clear whether this property of alignment in parallel OT has unwanted consequences. If the status of monosyllabic feet is constant in a language (i.e., monosyllables are always licit feet) and the apparent directionality is constant, then the alignment constraints have no problem describing stress patterns with monosyllabic feet, albeit with a different ranking than we might have initially thought. This is essentially the conclusion of Crowhurst and Hewitt (1995a).

However, there are reasons to think that this consequence of foot alignment in parallel OT is indeed undesirable. In a language that does not uniformly allow monosyllables to be licit feet, we predict that apparent directionality in the language may vary among words. For example, if syllables must be heavy to form monosyllabic feet, we predict that although the ranking of ALLFTL and ALLFTR remains the same in the language, the apparent directionality of the metrical parsing for each word will vary depending on whether and where a heavy syllable occurs. The unattested GT-like language observed by Hyde (2007) and discussed in section 3.2.2 has exactly this property. In this unattested language predicted by the standard approach to stress in parallel OT, the directionality is non-uniform; some words show apparent left-to-right footing (e.g., example (32)), while some show apparent right-to-left parsing (e.g., examples (34) and (35)), and still some words show neither because the left-most heavy syllable occurs in the middle of the word (e.g., example (33)).

Another potential problem posed by this prediction of alignment in parallel OT is a language with optional monosyllabic feet. Optionality may be formalized in OT with constraints that have variable rankings with respect to one another (Anttila 1997, 2002). In a language with variation between, for example, (σσ)(σσ)σ and (σσ)(σσ)(σ), the ranking of PARSESYLL and FTBIN(σ) must be in variation so that sometimes a syllable goes unparsed if it would lead to a monosyllabic foot, and sometimes monosyllabic feet are tolerated in order to maximally parse syllables into feet. However, to get this pattern in parallel OT, the ranking of ALLFTL and ALLFTR must also be in variation, and must co-vary with the ranking of PARSESYLL and FTBIN(σ). That is, for (σσ)(σσ)σ to emerge, FTBIN(σ) must outrank PARSESYLL and ALLFTL must outrank ALLFTR. But when (σσ)(σσ)(σ) wins, *both* rankings are reversed, such that PARSESYLL outranks FTBIN(σ) and ALLFTR outranks ALLFTL.

To my knowledge, no theory of variation in OT has a mechanism to force pairs of constraints into co-varying rankings. Instead, standard theories of variation in OT predict that if both pairs of constraints are variably ranked, then the ranking chosen for one pair of constraints would vary orthogonally with the ranking chosen for the other pair. Thus, the variable ranking between PARSESYLL and FTBIN(σ) on the one hand and ALLFTL and ALLFTR on the other predicts four-way variation between (σσ)(σσ)σ, (σσ)(σσ)(σ), σ(σσ)(σσ), and (σ)(σσ)(σσ) within a single language, which is an undesirable and unattested prediction, as no language appears to stress words in this way. The tableau in (79) and the summary in (80) show this result.

(79) Rankings in variation

/σσσσσ/	PARSESYLL	FTBIN(σ)	ALLFTL	ALLFTR
(σσ)(σσ)σ	1		2	4
(σσ)(σσ)(σ)		1	6	4
σ(σσ)(σσ)	1		4	2
(σ)(σσ)(σσ)		1	4	6

(80) Possible outcomes of variation

a. (σ)(σσ)(σσ) PARSESYLL >> FTBIN(σ) and ALLFTL >> ALLFTR

b. (σσ)(σσ)(σ) PARSESYLL >> FTBIN(σ) and ALLFTR >> ALLFTL

c. (σσ)(σσ)σ FTBIN(σ) >> PARSESYLL and ALLFTL >> ALLFTR

d. σ(σσ)(σσ) FTBIN(σ) >> PARSESYLL and ALLFTR >> ALLFTL

These examples suggest that the behavior of alignment constraints in parallel OT discussed by Crowhurst and Hewitt (1995a) is indeed something we should be glad to shed in HS/IFO. The serial model adopted in this paper does not reproduce this prediction, despite adopting the same constraints and definitions. Instead the emergent directionality from the ranking of ALLFTL and ALLFTR is constant.

4.2.2 *Alternatives to alignment*

Although the previous section showed that alignment in HS is more well-behaved than in parallel OT, it is reasonable to ask whether moving to serial evaluation of restricted candidate sets would allow some alternative to alignment which crucially does not rely on gradience, since this property has been much-maligned, particularly on formal grounds (McCarthy 2003; Eisner 1999; Potts and Pullum 2002, a.o.). In this section I discuss two potential contenders mentioned by McCarthy (2003) and show that they cannot in fact guarantee an appropriate parse in HS/IFO, particularly for bidirectional stress systems, while gradient alignment can. This section concludes that gradient alignment constraints are better than 'categorical' alternatives for determining foot placement in HS with IFO.

4.2.2.1 'Categorical' alignment I

McCarthy (2003: 79) offers in passing two alternatives to gradient alignment which fall under the categorical constraint schema he defines. I will discuss these two alternatives in turn and show that both fail as methods of determining foot placement in HS with IFO.

The first offered alternative that I will discuss are the constraints defined in (81) and (82), which can be considered categorical equivalents of ALLFTL and ALLFTR, respectively.

(81) ***σ/__...FT**: Assign one violation mark for every syllable appearing to the left of some foot.

(82) ***σ/FT…__**: Assign one violation mark for every syllable appearing to the right of some foot.

In the standard definition of alignment, a syllable can incur more than one alignment violation, subject to the number of feet in the word. In this proposed alternative formulation, a syllable can incur at most one violation of the constraint, and it does so when it appears between some foot and the appropriate word edge. The table in (83) shows the violation marks assessed by ALLFTL and *σ/__...FT to candidates with one, two, and three left-aligned feet. In (a) and (b) the constraints match in their assessments, but in (c), gradient ALLFTL assigns more violation marks than *σ/__...FT; the former assigns marks for every syllable to the left of *each* foot, meaning some syllables incur more than one violation mark, while the latter assigns only one mark per syllable occurring to the left of *some* foot.

(83) Comparison of marks assessed by ALLFTL and *σ/__...FT

		ALLFTL	*σ/__...FT
a.	(σσ)σσσσσ	0	0
b.	(σσ)(σσ)σσσ	2	2
c.	(σσ)(σσ)(σσ)σ	6	4

Similarly, the table in (84) shows the same candidates and the violation marks they receive from ALLFTR and the categorical equivalent in (82). Each candidate shown receives five violation marks from *σ/FT…__, because in each case there are five syllables that appear to the right of a foot. Thus, the number of violation marks does not change with additional feet placed to the right of the foot in (a).

(84) Comparison of marks assessed by ALLFTR and *σ/FT…__

		ALLFTR	*σ/FT…__
a.	(σσ)σσσσσ	5	5
b.	(σσ)(σσ)σσσ	8	5
c.	(σσ)(σσ)(σσ)σ	9	5

This alternative works for contiguous foot building of the type shown in most of the examples in this paper. A representative derivation for a language with left-to-right disyllabic feet is shown in (85).

(85) Derivation with categorical foot placement constraints

/σσσσσ/ **1st iteration**	FTBIN(σ)	PARSESYLL	*σ/__...FT	*σ/FT…__
a. σσσσσ		**W** 5		
b. (σ)σσσσ	**W** 1	**W** 4		**W** 4
c. (σσ)σσσ		3		3
d. σ(σσ)σσ		3	**W** 1	**L** 2
2nd iteration				
e. (σσ)σσσ		**W** 3	**L**	3
f. (σσ)(σ)σσ	**W** 1	**W** 2	2	3
g. (σσ)(σσ)σ		1	2	3
h. (σσ)σ(σσ)		1	**W** 3	3
3rd iteration				
i. (σσ)(σσ)σ		1	2	3
j. (σσ)(σσ)(σ)	**W** 1	**L**	**W** 4	3

Output: (σσ)(σσ)σ

However, to see why these constraints cannot be what dictates foot placement in HS/IFO, we need to look at bidirectional stress systems. Examples include the languages Piro (Matteson 1965), Lenakel (verbs and adjectives only; Lynch 1974, 1977, 1978), and Garawa (Furby 1974). In both Piro and Lenakel a single foot is placed at the right edge of the word and other feet iterate from the left, as shown in (86) with data from Lenakel; Garawa shows the opposite pattern. I will first show how the standard alignment constraints are able to get this pattern in HS/IFO, and then I turn to why the categorical constraints cannot.

(86) Bidirectional stress in Lenakel[38]

Odd parity: ˈtɨnaˈkamarolˈkeykey ‘you (pl.) will be liking it’

Even parity: ˈnɨmaˈmarolˈkeykey ‘you (pl.) were liking it’

In order to derive a pattern like this in HS/IFO with the constraints we have been using, the derivation must proceed by first building the static right-edge foot and then iterating feet from the left. I will illustrate how this works and then explain why it must

[38] Data from Hayes (1995:168). Degrees of stress ignored (rightmost is primary).

be this way. Since it is clear that Lenakel must have a foot at the right edge of every word we will employ the constraint ALIGNWDR, familiar from our analysis of Fijian in section 3.2, which demands that a word have a foot aligned with its right edge. This should be ranked above our standard foot alignment constraint ALLFTL so that the foot indeed is placed at the right. ALLFTL must in turn dominate ALLFTR to account for why the non-rightmost stresses align to the left. And PARSESYLL must dominate ALLFTL to get multiple feet iterating from the left. The derivation of a seven-syllable word in HS/IFO with gradient alignment is shown in (87). The rankings just asserted can be verified by examining the tableau.

(87) Derivation of 7-syllable word in Lenakel w/ gradient alignment

/tɨnakamarolkeykey/ **1st iteration**	PARSESYLL	ALIGNWDR	ALLFTL	ALLFTR
a. tɨ.na.ka.ma.rol.key.key	**W** 7	**W** 1	**L**	
b. (ˈtɨ.na)ka.ma.rol.key.key	5	**W** 1	**L**	**W** 5
c. tɨ.na.ka.ma.rol(ˈkey.key)	5		5	
2nd iteration				
d. tɨ.na.ka.ma.rol(ˈkey.key)	**W** 5		5	**L**
e. tɨ.na.ka(ˈma.rol)(ˈkey.key)	3		**W** 8	**L** 2
f. (ˈtɨ.na)ka.ma.rol(ˈkey.key)	3		5	5
3rd iteration				
g. (ˈtɨ.na)ka.ma.rol(ˈkey.key)	**W** 3		**L** 5	**L** 5
h. (ˈtɨ.na)ka(ˈma.rol)(ˈkey.key)	1		**W** 8	**L** 7
i. (ˈtɨ.na)(ˈka.ma)rol(ˈkey.key)	1		7	8

Output: (ˈtɨ.na)(ˈka.ma)rol(ˈkey.key)

This is the only way that the bidirectional derivation can be ordered. It is not possible to account for this stress system by building feet from the left edge and then 'skipping' the penultimate syllable in odd-parity words to achieve the appearance of a bidirectional system. This is because of the fact that higher-ranked constraints effectively have their preferences satisfied first in HS, which is what determines ordering. If ALLFTL were to outrank ALIGNWDR, then it will be optimal to add contiguous feet from the left edge, and when faced with the choice of *(ˈtɨ.na)(ˈka.ma)(ˈrol.key)key* or *(ˈtɨ.na)(ˈka.ma)rol(ˈkey.key)* at the third iteration, the former will win (assuming both options equally satisfy foot form constraints, etc.). In order to have the latter option win, ALIGNWDR must outrank ALLFTL, but under this ranking, the candidate with a foot on the right edge, *tɨ.na.ka.ma.rol(ˈkey.key)*, will win over *(ˈtɨ.na)ka.ma.rol.key.key* at the first iteration because the candidate with the right-aligned foot better satisfies the higher ranked constraint. Thus, bidirectional derivations must begin by satisfying the top-ranked ALIGNWD constraint, then iterate feet from the opposite edge.

When we replace ALLFTL and ALLFTR with the categorical alternatives in (81) and (82), we do not achieve the desired result. The violation marks assigned by these constraints do not identify the correct winner at the second iteration. The following sub-derivation illustrates. The first iteration looks the same as the first iteration from the derivation with gradient alignment in (87), but in the second iteration, candidates (e) and (f) tie on PARSESYLL and *σ/__...FT, while *σ/FT…__ prefers candidate (e), *tɨ.na.ka(ˈma.rol)(ˈkey.key)*, because it incurs fewer additional violations of this constraint.

(88) Wrong winner with categorical alignment

		/tɨnakamarolkeykey/ **1st iteration**	PARSESYLL	ALIGNWDR	*σ/__...FT	*σ/FT…__
	a.	tɨ.na.ka.ma.rol.key.key	**W** 7	**W** 1	**L**	
	b.	(ˈtɨ.na)ka.ma.rol.key.key	5	**W** 1	**L**	**W** 5
•	c.	tɨ.na.ka.ma.rol(ˈkey.key)	5		5	
➤		**2nd iteration**				
	d.	tɨ.na.ka.ma.rol(ˈkey.key)	**W** 5		5	**L**
☠→	e.	tɨ.na.ka(ˈma.rol)(ˈkey.key)	3		5	2
	f.	(ˈtɨ.na)ka.ma.rol(ˈkey.key)	3		5	**W** 5

It is possible to rescue the second iteration by employing ALIGNWDL, which assigns a violation mark for a word that does not have a foot at its *left* edge, and ranking it above *σ/FT…__ to force (f) to win in (88). However, even if this strategy were employed to rescue the second iteration, an even bigger problem would arise in the third iteration. As the sub-derivation in (89) shows, the categorical alignment constraints do not distinguish the candidates in (b) and (c) in (89) precisely because they both already have a foot at the left and right edge, meaning additional feet do not add additional violations. In fact, the categorical alignment constraints are not even able to be ranked relative to each other on the basis of this derivation. Indeterminacy of this sort is not easily fixed.

(89) Indeterminacy in third iteration with categorical alignment

〈(ˈtɨ.na)ka.ma.rol(ˈkey.key)〉 **3rd iteration**	PARSE-SYLL	ALIGN WDR	ALIGN WDL	*σ/__... FT	*σ/FT… __
a. (ˈtɨ.na)ka.ma.rol(ˈkey.key)	**W** 3			5	5
→ b. (ˈtɨ.na)ka(ˈma.rol)(ˈkey.key)	1			5	5
→ c. (ˈtɨ.na)(ˈka.ma)rol(ˈkey.key)	1			5	5

Thus, for bidirectional systems, the categorical foot placement constraints defined in (81) and (82) will not suffice to determine the placement of non-peripheral feet. McCarthy (2003: 79) observes the these constraints behave this way, but he argues this to be a virtue. Here we see that the indeterminacy leads to failure of HS/IFO to analyze Lenakel. Because peripheral feet are built first in a bidirectional parse, we must face the choice between (89)b and (89)c at this point in the derivation, but these alternative constraints cannot decide between them.

4.2.2.2 'Categorical' alignment II

The other possible replacement for gradient alignment constraints that is mentioned by McCarthy (2003; also Pater 2007) is to formulate local constraints that directly prefer contiguous footing by penalizing feet with an adjacent unfooted syllable, as in (90) and (91).

(90) ***FT/σ__**: Assign one violation mark for every foot with an adjacent unfooted syllable to the left.

(91) ***FT/__σ**: Assign one violation mark for every foot with an adjacent unfooted syllable to the right.

However, this proposal also fails to determine an optimal parse in bidirectional stress systems, for very similar reasons. The tableau in (92) illustrates an attempt at the same seven-syllable derivation with these constraints. The constraint ALIGNWDL is again needed to get the correct winner in the second iteration, and the third iteration again fails to return an optimal candidate.

(92) Derivation of 7-syll word, w/ new constraints

/σσσσσσσ/ **1st iteration**	PARSE-SYLL	ALIGN WDR	*FT/σ__	ALIGN WDL	*FT/__σ
a. tɨ.na.ka.ma.rol.key.key	**W** 7	**W** 1	**L**	1	
b. (ˈtɨ.na)ka.ma.rol.key.key	5	**W** 1	**L**	**L**	**W** 1
c. tɨ.na.ka.ma.rol(ˈkey.key)	5		1	1	
2nd iteration					
d. tɨ.na.ka.ma.rol(ˈkey.key)	**W** 5		1	**W** 1	**L**
e. tɨ.na.ka(ˈma.rol)(ˈkey.key)	3		1	**W** 1	**L**
f. (ˈtɨ.na)ka.ma.rol(ˈkey.key)	3		1		1
3rd iteration					
g. (ˈtɨ.na)ka.ma.rol(ˈkey.key)	**W** 3		1		1
→ h. (ˈtɨ.na)ka(ˈma.rol)(ˈkey.key)	1		1		1
→ i. (ˈtɨ.na)(ˈka.ma)rol(ˈkey.key)	1		1		1

In common with the previous alternatives, this variation was also observed by McCarthy (2003) to have the property of not distinguishing candidates like (h) and (i). In HS/IFO, we would not be able to account for bidirectional stress systems at all if this were the only mechanism we had for determining foot placement. We need gradient alignment to analyze the stress system of Lenakel, which wants the outcome in (i), and to distinguish it from a language like Garawa, in which (h) would emerge as optimal.

It would seem then that these non-gradient alternatives to generalized alignment constraints encounter problems because they are not sufficiently deterministic. Both of these categorical alternatives to alignment are subject to the same problem and we are thus in the position of preferring a gradient alignment analysis on the grounds that it allows us to derive bidirectional stress systems. Standard alignment constraints can handle bidirectional systems because the constraints that prefer regular iteration from one edge are gradiently defined and thus ensure feet will be as far to the dominant edge as they can be, modulo the satisfaction of other high-ranked constraints on metrical structure and location of other feet; this is something the proposed alternatives cannot do.

4.3 Summary of other consequences

This section has shown that HS/IFO has several formal consequences that further suggest its advantage over parallel OT. The constraint PARSESYLL was shown to have desirable effects with iterative foot building. HS/IFO restricts this constraint from compelling non-local interactions, and it also simulates a violable maximality condition on foot size, which was useful in our analysis of generalized trochees in section 3.1. Meanwhile, gradient alignment constraints, which have received much negative attention in parallel OT, were shown to be well-behaved in HS/IFO. Some potential alternatives to gradient alignment were shown to be insufficient for analyzing bidirectional stress systems, which leads us to prefer the gradient versions. The results of this section suggest we are justified in retaining some of the standard stress constraints in HS/IFO.

5 Conclusion

In this paper I have attempted to demonstrate the potential of Harmonic Serialism with iterative foot optimization to model rhythmic stress and its interactions with syllable weight, vowel shortening, and edge restrictions. HS/IFO predicts a certain amount of derivationally-defined locality in stress systems, and attested stress systems match these predictions. Parallel OT on the other hand was shown to be insufficiently restrictive because of its ability to analyze both local and non-local interactions. Attestation of truly non-local stress systems is extremely weak, and thus we should prefer HS/IFO because it more accurately reflects stress typology. Combined with the results of section 4, which showed that familiar constraints behave predictably in HS/IFO, this model holds much promise as a novel way to analyze and account for stress systems in natural language.

6 References

Alber, B. (1997). Quantity sensitivity as a result of constraint interaction. In G. Booij & J. van de Weijer (Eds.), *Phonology in progress - progress in phonology. HIL Phonology Papers* (pp. 1-45). The Hague: Holland Academic Press.

Alber, B. (2005). Clash, lapse, and directionality. *Natural Language & Linguistic Theory, 23*, 485-542.

Allen, W. S. (1965). *Vox Latina: a guide to the pronunciation of classical Latin.* Cambridge: Cambridge University Press.

Allen, W. S. (1973). *Accent and rhythm. Prosodic features of Latin and Greek: a study in theory and reconstruction*. Cambridge: Cambridge University Press.

Anttila, A. (1997). Deriving variation from grammar. In F. Hinskens, R. van Hout & L. Wetzels (Eds.), *Variation, Change, and Phonological Theory* (pp. 35-68). Amsterdam: John Benjamins. ROA-63.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory, 20*, 1-42.

Crowhurst, M., & Hewitt, M. (1995a). Directional footing, degeneracy, and alignment, *Proceedings of the Northeast Linguistic Society* (Vol. 25, pp. 47-61).

Crowhurst, M., & Hewitt, M. (1995b). Prosodic overlay and headless feet in Yidin^y. *Phonology, 12*, 39-84.

Dixon, R. M. W. (1977a). Some phonological rules in Yidin^y. *Linguistic Inquiry, 8*, 1-34.

Dixon, R. M. W. (1977b). *A grammar of Yidin^y*. Cambridge: Cambridge University Press.

Dixon, R. M. W. (1988). *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

Eisner, J. M. (1999). Doing OT in a straitjacket (Handout). Available at http://www.cs.jhu.edu/~jason/#ucla99.

Elfner, E. (2008). Prosodic circumscription in harmonic serialism (handout). HUMDRUM at Rutgers 2008.

Furby, C. (1974). *Garawa phonology*. Canberra: Australian National University.

Hagberg, L. (1993). *An autosegmental theory of stress*. Doctoral dissertation: University of Arizona, Tucson. (Published as SIL e-books 3. Available at http://www.sil.org/silepubs/index.asp?series=941).

Halle, M., & Vergnaud, J.-R. (1987). *An essay on stress*. Cambridge, MA: MIT Press.

Hansen, K. C., & Hansen, L. E. (1969). Pintupi phonology. *Oceanic Linguistics, 8*, 153-170.

Hansen, K. C., & Hansen, L. E. (1978). *The core of Pintupi grammar*. Alice Springs, Northern Territory, Australia: Institute for Aboriginal Development.

Hanson, K., & Kiparsky, P. (1996). A parametric theory of poetic meter. *Language, 72*, 287-335.

Hayes, B. (1980). *A metrical theory of stress rules*. Doctoral dissertation: MIT.

Hayes, B. (1982). Metrical structure as the organizing principle of Yidin[y] phonology. In H. van der Hulst & N. Smith (Eds.), *The Structure of Phonological Representations, Part I*. Dordrecht, the Netherlands: Foris Publications.

Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. In M. Niepokuj, M. VanClay, V. Nikiforidou & D. Jeder (Eds.), *Proceedings of the Berkeley Linguistics Society: Parasession on Poetics, Metrics, and Prosody* (Vol. 11, pp. 429-446). Berkeley, CA.

Hayes, B. (1987). A revised parametric metrical theory. In J. McDonough & B. Plunkett (Eds.), *Proceedings of the North East Linguistic Society* (Vol. 17, pp. 274-289). Amherst: GLSA.

Hayes, B. (1991). Metrical stress theory: principles and case studies. Ms. UCLA. (Early version of Hayes 1995).

Hayes, B. (1995). *Metrical stress theory: principals and case studies*. Chicago: University of Chicago Press.

Hayes, B. (1997). Anticorrespondence in Yidiɲ. Ms. UCLA. (Available at http://www.linguistics.ucla.edu/people/hayes/yidiny/anticorresp.pdf).

Hayes, B. (1999). Phonological restructuring in Yidin[y] and its theoretical consequences. In B. Hermans & v. Oostendorp (Eds.), *The Derivational Residue in Phonological Optimality Theory* (pp. 175-205). Amsterdam: John Benjamins.

Hercus, L. (1986). *Victorian languages: a late survey*. Canberra: Australian National University.

Hewitt, M. (1994). Deconstructing foot binarity in Koniag Alutiiq. Ms. University of British Columbia. ROA-12.

Hint, M. (1973). *Eesti keele sõnafonologia* (Vol. 1: Rõhusüsteemi fonoloogia ja morfofonoloogia põhiprobleemid). Talinn: Eeste NSV Teaduste Akadeemia.

Hung, H. J. (1993). Iambicity, rhythm, and non-parsing. Ms. University of Ottawa. ROA-9.

Hung, H. J. (1994). *The Rhythmic and Prosodic Organization of Edge Constituents*. Doctoral dissertation: Brandeis.

Hyde, B. (2007). Non-finality and weight-sensitivity. *Phonology, 24,* 287-334.

Jacobs, H. (2008). Quantity-insensitive stress and the need for OT with Candidate Chains. Ms. Radboud University Nijmegen.

Jesney, K. (2008). Positional Faithfulness, non-locality, and the Harmonic Serialism solution. Poster presented at NELS 39.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris Publications.

Kager, R. (1992a). Are there any truly quantity-insensitive systems?, *Proceedings of the Berkeley Linguistics Society* (Vol. 18, pp. 123-132). Berkeley, CA.

Kager, R. (1992b). Shapes of the generalized trochee, *Proceedings of the West Coast Conference on Formal Linguistics* (Vol. 11, pp. 298-312).

Kager, R. (1994). Ternary rhythm and alignment theory. Ms. Utrecht University. ROA-35.

Kager, R. (2001). Rhythmic directionality by positional licensing (Handout). Fifth HIL Phonology Conference: University of Potsdam. ROA-514.

Kimper, W. (2008). Local optionality and harmonic serialism. Ms. University of Massachusetts Amherst. ROA-988.

Kirchner, R. (1992). *Harmonic phonology within one language: an analysis of Yidin^y*. Master's thesis: University of Maryland.

Lynch, J. D. (1974). *Lenakel phonology*. Doctoral dissertation: University of Hawai'i.

Lynch, J. D. (1977). *Lenakel dictionary, Pacific Linguistics C55*. Canberra: Australian National University.

Lynch, J. D. (1978). *A grammar of Lenakel, Pacific Linguistics B55*. Canberra: Australian National University.

Matteson, E. (1965). *The Piro (Arawakan) language*. Berkeley and LA: University of California Press.

McCarthy, J. J. (2000). Harmonic serialism and parallelism. In M. Hirotani (Ed.), *Proceedings of the North East Linguistics Society* (Vol. 30, pp. 501-524). Amherst: GLSA.

McCarthy, J. J. (2002). *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.

McCarthy, J. J. (2003). OT constraints are categorical. *Phonology, 20*, 75-138.

McCarthy, J. J. (2006). Restraint of analysis. In E. Bakovic, J. Ito & J. J. McCarthy (Eds.), *Wondering at the natural fecundity of things: essays in honor of Alan Prince* (pp. 213-239). Santa Cruz, CA: Linguistics Research Center.

McCarthy, J. J. (2007a). *Hidden generalizations : phonological opacity in optimality theory*. Oakville, CT: Equinox.

McCarthy, J. J. (2007b). Slouching towards optimality: Coda reduction in OT-CC. *Phonological Studies (Journal of the Phonological Society of Japan), 7*.

McCarthy, J. J. (to appear-a). The serial interaction of stress and syncope. *Natural Language & Linguistic Theory*.

McCarthy, J. J. (to appear-b). The gradual path to cluster simplification. *Phonology*.

McCarthy, J. J., & Prince, A. (1986/1996). *Prosodic morphology*: Rutgers Technical Repot TR-32. Available at http://ruccs.rutgers.edu/ruccs/publications.php.

McCarthy, J. J., & Prince, A. (1993a). Generalized alignment. *Yearbook of Morphology*, 79-153.

McCarthy, J. J., & Prince, A. (1993b). Prosodic morphology: constraint interaction and satisfaction. University of Massachusetts and Rutgers University. ROA-482.

McCarthy, J. J., & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey & S. Urbanczyk (Eds.), *Papers in Optimality Theory*. Amherst: GLSA.

Myers, S. (1987). Vowel shortening in English. *Natural Language & Linguistic Theory, 5*, 485-518.

Nash, D. (1979). Yidin^{y} stress: a metrical account, *Proceedings of the Northeast Linguistic Society* (Vol. 9). *CUNYForum*, 7/8, 112–130.

Pater, J. (2007). Harmonic grammar with harmonic serialism. Handout from UMass Amherst colloquium series talk, November 2007.

Pater, J. (2008a). Serial harmonic grammar. Invited talk, UNC-Chapel Hill Spring Colloquium, April 2008.

Pater, J. (2008b). Optimization and linguistic typology. Ms. University of Massachusetts Amherst. ROA-982.

Payne, D., Payne, J., & Santos, J. (1982). *Morfologia, fonologia, y fonetica del Asheninca del Apurucayali (Campa - Arawak Preandino)*. Yarinacocha, Peru: Summer Institute of Linguistics.

Potts, C., & Pullum, G. K. (2002). Model theory and the content of OT constraints. *Phonology, 19,* 361-393.

Prince, A. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry, 11*, 511-562.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry, 14*, 19-100.

Prince, A. (1985). Improving tree theory, *Proceedings of the Berkeley Linguistics Society* (Vol. 11, pp. 471-490). Berkeley, CA.

Prince, A. (1990). Quantitative consequences of rhythmic organization, *Proceedings of the Chicago Linguistic Society* (Vol. 26, pp. 355-398).

Prince, A. (2002). Arguing optimality. In A. Carpenter, A. Coetzee & P. de Lacy (Eds.), *University of Massachusetts Occasional Papers in Linguistics 26: Papers in Optimality Theory II* (pp. 269-304). Amherst: GLSA. ROA-562.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: constraint interaction in generative grammar*. Malden, MA: Blackwell.

Schütz, A. J. (1985). *The Fijian language*. Honolulu: University of Hawaii Press.

Selkirk, E. (1984). *Phonology and Syntax*. Cambridge, MA: MIT Press.

Steriade, D. (1988). Greek accent: a case for preserving structure. *Linguistic Inquiry, 19*, 271-314.

Wolf, M. (2008). *Optimal Interleaving*. Doctoral dissertation: University of Massachusetts Amherst.