PART III

Issues and Answers in Optimality Theory

9. Inventory Theory and the Lexicon

All grammatical constraints are violable, in principle. A constraint such as ONS, 'syllables have onsets', in and of itself and prior to its interaction with other constraints, does not assert that syllables lacking onsets are impossible, but rather that they are simply less harmonic than competitors possessing onsets. Its function is to sort a candidate set by measuring adherence to (equivalently: divergence from) a formal criterion. Constraints therefore define relative rather than absolute conditions of ill-formedness, and it may not be immediately obvious how the theory can account for the absolute impossibility of certain structures, either within a given language or universally. Yet in the course of the preceding analyses we have seen many examples of how Optimality Theory explains language-particular and universal limits to the possible. In this section, we identify the general explanatory stategy that these examples instantiate, and briefly illustrate how this strategy can be applied to explaining segmental inventories. We then consider implications for the lexicon, proposing a general induction principle which entails that the structure of the constraints in a language's grammar is strongly reflected in the content of its lexicon. This principle, Lexicon Optimization, asserts that when a learner must choose among candidate underlying forms which are equivalent in that they all produce the same phonetic output and in that they all subserve the morphophonemic relations of the language equally well, the underlying form chosen is the one whose output parse is most harmonic.

9.1 Language-Particular Inventories

We begin by examining a simple argument which illustrates the central challenge of accounting for absolute ill-formedness in a theory of relative well-formedness:

"For Optimality Theory, syllables without onsets are not absolutely ill-formed, but only relatively. The syllable .VC. (for example) is *more* ill-formed than the syllable .CV., but .VC. is not *absolutely* ill-formed. How can Optimality Theory bar .VC. from any language's syllable inventory?

"What Optimality Theory would need in order to outlaw such syllables is some additional mechanism, like a *threshold* on ill-formedness, so that when the graded ill-formedness of syllables passes this threshold, the degree of ill-formedness becomes absolutely unacceptable."

The fallacy buried in this argument has two facets: a failure to distinguish the inputs from the outputs of the grammar, coupled with an inappropriate model of grammar in which the ill-formed are those *inputs* which are rejected by the grammar. In Optimality Theory, the job of the grammar is not to accept or reject *inputs*, but rather to assign the best possible structure to every input. The place to look for a definition of ill-formedness is in the set of *outputs* of the grammar. These outputs are, by definition, well-formed; so what is ill-formed — absolutely ill-formed — is any structure which is never found among the outputs of the grammar. To say that .VC. syllables are not part of the inventory of a given language is not to say that the grammar rejects /VC/ and the like as input, but rather that no ***output*** of the grammar ever contains .VC. syllables.

We record this observation in the following remark:

(277) **Absolute ill-formedness**. A structure φ is (absolutely) ill-formed with respect to a given grammar iff there is no input which when given to the grammar leads to an output that contains φ.

Note further that in a demonstration that .VC. syllables are ill-formed according to a given grammar, the input /VC/ has no *a priori* distinguished status. We need to consider every possible input in order to see whether its output parse contains a syllable .VC. Of course, /VC/ is a promising place to start the search for an input which would lead to such a parse, but, before concluding that .VC. syllables are barred by the grammar, we must consider all other inputs as well. Perhaps the optimal parse of /C/ will turn out to be .□́C., providing the elusive .VC. syllable. It may well be possible to show that if *any* input leads to .VC. syllables, then /VC/ will — but in the end such an argument needs to be made.

If indeed .VC. syllables are ill-formed according to a given grammar, then the input /VC/ must receive a parse other than the perfectly faithful one: .VC. At least one of the faithfulness constraints PARSE and FILL must be violated in the optimal parse. We can therefore generally distinguish two paths that the grammar can follow in order to parse such problematic inputs: violation of PARSE, or violation of FILL. The former we have called 'underparsing' the input, and in some other accounts would correspond to a 'deletion repair strategy'; the latter, overparsing, corresponds to an 'epenthesis repair strategy'. (In §10.3 we explicitly compare Optimality Theory to some repair theories.) These two means by which a grammar may deal with problematic inputs were explicitly explored in the Basic CV Syllable Structure Theory of §6. There we found that .VC. syllables were barred by either

[i] requiring onsets: ranking either PARSE or FILL^{Ons} lower than ONS; or
[ii] forbidding codas: ranking either PARSE or FILL^{Nuc} lower than −COD.

One particularly aggressive instantiation of the underparsing strategy occurs when the optimal structure assigned by a grammar to an input is the *null* structure: no structure at all. This input is then grammatically completely unrealizable, as discussed in §4.3.1. There is some subtlety to be reckoned with here, which turns on what kinds of structure are asserted to be absent in the null output. In one sense, the *null* means 'lacking in realized phonological content', with maximal violation of PARSE, a possibility that can hardly be avoided in the candidate set if underparsing is admitted at all. In another sense, the null form will fail to provide the morphological structure required for syntactic and semantic interpretation, violating M-PARSE. To achieve full explicitness, the second move requires further development of the morphological apparatus; the first requires analogous care in formulating the phonetic interpretation function, which will be undefined in the face of completely unparsed phonological material. In this discussion, we will gloss over such matters, focusing on the broader architectural issues.

It would be a conceptual misstep to characterize null parsing as *rejection of the input* and to appeal to such rejection as the basis of a theory of absolute ill-formedness. For example, it would be wrong to assert that a given grammar prohibits .VC. syllables because the input /VC/ is assigned the null structure; this is a good hint that the grammar may bar .VC. syllables, but what needs to be demonstrated is that *no* input leads to such syllables. In addition, a grammar which assigns some non-null structure to /VC/, for example .□V.〈C〉, might nonetheless prohibit .VC. syllables.

Subject to these caveats, it is clear that assigning null structure to an input is one means a grammar may use to prevent certain structures from appearing in the output. The Null Parse is a possible candidate which must always be considered and which may well be optimal for certain particularly problematic inputs. We have already seen two types of examples where null structures can be optimal. The first example emerged in the analysis of Latin minimal word phenomenon in §4.3.1, where, given a certain interpretation of the data, under the pressure of FTBIN and LX≈PR, the optimal parse of the monomoraic input is null (but see Mester 1992:19-23). The second was in the CV Syllable Structure Theory of §6, where it was shown that the structure assigned to /V/ is null in any language requiring onsets and enforcing ONS by underparsing: that is, where PARSE is the least significant violation, with {ONS, FILL^{Ons}} $\gg$ PARSE; as in ex. (134), p.101.

9.1.1 Harmonic Bounding and Nucleus, Syllable, and Word Inventories

Absolute ill-formedness, explicated in (277), is an emergent property of the interactions in a grammar. Showing that a structure φ is ill-formed in a given language requires examination of the system. One useful strategy of proof is to proceed as follows. First we let A denote an arbitrary candidate parse which contains the (undesirable) structure φ. Then we show how to modify *any such analysis* A to produce a particular (better) competing candidate parse B of the same input, where B does not contain φ and where B is provably more harmonic than A. This is sufficient to establish that no structure containing φ can ever be optimal. The structure φ can never occur in any output of the grammar, and is thus absolutely ill-formed. We have called this method of proof "Harmonic Bounding" — it establishes that every parse containing the structure φ is bettered by, bounded above by, one that lacks φ.

The strategy of Harmonic Bounding was implicitly involved, for example, in the analysis of the minimal word phenomenon (§4.3.1). In this case, the impossible structure is $\varphi = [\mu]_{\text{PrWd}}$. We examined the most important type of input, a monomoraic one like /*re*/, and showed that the analysis containing φ, A = $[[$*ré*$]_{\text{F}}]_{\text{PrWd}}$, is less harmonic than a competitor B = 〈*re*〉, the Null Parse, which lacks φ. The method of constructing B from A is simply to replace structure with no structure.

To complete the demonstration that the Latin constraint hierarchy allows no monomoraic words in the output, we must consider every input that could give rise to a monomoraic word. We need to examine inputs with less than one mora, showing that they do not get overparsed as a single empty mora: $[[\acute{\square}_{\mu}]_{\text{F}}]_{\text{PrWd}}$. We also must consider inputs of *more than one* mora, showing that these do not get underparsed, with only one mora being parsed into the PrWd: $[[\mu]_{\text{F}}]_{\text{PrWd}}\langle\mu\cdots\rangle$. Both of these are also harmonically bounded by the Null Parse of the relevant inputs. On top of whatever violation marks are earned by complete structuring of monomoraic input — marks that are already sufficient to establish the superiority of the Null Parse — these moraic over- and under-parses incur *FILL and *PARSE marks as well, and it is even clearer that a monomoraic parse cannot be optimal.

Similarly, in the analysis of Lardil in §7, we provided the core of the explanation for why no words in its inventory can be monomoraic. The result is the same as in Latin, but enforcement of LX≈PR and FTBIN for monomoraic inputs is now by *over*parsing rather than by *under*parsing, due to differences in the constraint ranking. The structure we wish to exclude is again $\varphi = [\mu]_{\text{PrWd}}$, and, as in Latin, we examined monomoraic inputs such as /maɽ/ (182), p.130, to see if their parses contained φ. In each case, the optimal parse is a bisyllabic competitor B with an unfilled second mora. We also examined vowel-final *bimoraic* inputs (184), p.134, because, for longer inputs, a final vowel is optimally *un*parsed, a pattern which would lead to monomoraicity if universally applied. However, both moras in bimoraic inputs must be parsed, so again we fail to produce a monomoraic output. Inputs with three or more moras leave a final vowel unparsed, but parse all the others (183), p.132. Thus, there are no inputs, long or short, which produce monomoraic outputs.

It is worth emphasizing that, even though the lack of monomoraic words in the Latin and Lardil inventories is a result of the high ranking of LX≈PR and FTBIN in the domination hierarchy, it would be distinctly incorrect to summarize the Optimality Theory explanation as follows: "LX≈PR and FTBIN are superordinate therefore unviolated, so any monomoraic input is thereby rendered absolutely ill-formed." An accurate summary is: "LX≈PR and FTBIN dominate a FAITHFULNESS constraint (PARSE in Latin; FILL in Lardil), so for *any input at all* — including segmentally monomoraic strings as a special case — monomoraic parses are always less harmonic than available alternative analyses (Null Parse for Latin, bisyllable for Lardil); therefore outputs are never monomoraic."

Successful use of the Harmonic Bounding argument does not require having the optimal candidate in hand; to establish *φ in the absolute sense, it is sufficient to show that there is always a B-without-φ that is better than any A-with-φ. Whether any such B is optimal is another question entirely. This can be seen clearly in the kind of argument pursued repeatedly above in the development of the Basic Segmental Syllable Theory in §8. For example, as part of the process of deriving the typology of segmental inventories licensed by various syllable positions, we showed that the inventory of possible nuclei could not include a segment α in any language in which *P/α ≫ {FILL^{Nuc}, *M/α}).[70] These are languages in which it is

[i] more important to keep α out of the Nucleus (P = 'peak') than to fill the Nucleus, and

[ii] more important to keep α out of the Nucleus than to keep it out of the syllable margins.

The φ we want to see eliminated is the substructure Nuc/α, in which the segment α is dominated by the node Nucleus. Let A denote an arbitrary parse containing Nuc/α = ά, so that a segment α appearing in the input string is parsed as a nucleus: A = ~ά~. The bounding competitor B is identical to A except that the structure in question, Nuc/α, has been replaced by the string in which α is an onset sandwiched between two empty nuclei; B = ~□́.α□́~. In terms of the slash-for-domination notation, the crucial replacement pattern relating A to B can be shown as

A = … Nuc/**α** … B = … Nuc/□. Ons/**α** Nuc/□. … .

We have then the following argument:

(278) **Harmonic Bounding Argument** showing α is an impossible nucleus

a. Assumed constraint ranking

*P/α ≫ {FILL^{Nuc}, *M/α}

[70] Here, *M/α and *P/α are the constraints against parsing α as a Margin (Onset, Coda) and as a Peak (Nucleus), respectively; this is the contrapositive of the Possible Peak Condition (230).

b. Structures

i. $\varphi = \acute{\alpha}$ (segment α qua nucleus)

ii. $A = \sim \acute{\alpha} \sim$ (any parse taking α to be a nucleus)

iii. $B = \sim \acute{\square}.\alpha\acute{\square} \sim$ (analysis A modified in a specific way to make α nonnuclear)

c. Argument: show that B bests A.

It should be clear that B is always more harmonic than A in the given languages. The mark *P/α incurred by nucleizing α in A is worse than both the marks *M/α (for marginalizing α) and *FILL^{Nuc} (for positing empty nuclei) that are incurred by B. Hence, in such a grammar the optimal parse can never include φ = Nuc/α, no matter what the input. The conclusion is that α is not in the inventory of possible nuclei for these languages. However, we cannot conclude that every occurence of α is in onset position, as in the bounding analysis B, or indeed, without further argument, that *any* occurrence of α is in onset position. There may be other analyses that are even more harmonic than B in specific cases; but we are assured that α will never be a nucleus in any of these. (In fact, under certain rankings consistent with (278a) α will be banned from the surface altogether, barred from the onset as well as the nucleus, as an 'untenable association', (225), p. 156.)

The Harmonic Bounding strategy is explicitly carried out for syllable inventories in the CV theory in the appendix, and is implicitly involved in a number of other results derived above. Samek-Lodovici (1992ab) makes independent use of the same method of proof (taking B to be a kind of Null Parse) to establish the validity of his Optimality theoretic analysis of morphological gemination processes.

9.1.2 Segmental Inventories

Having illustrated the way prosodic inventories are delimited, from the structural level of the syllable position (e.g. Nuc) up through the syllable itself to the word, we can readily show how the technique extends downward to the level of the segment. Now we take as inputs not strings of already formed segments, but rather strings of feature sets. These must be optimally parsed into segments by the grammar, just as (and at the same time as) these segments must be parsed into higher levels of phonological structure. The segmental inventory of a language is the set of segments found among the optimal *output parses* for all possible inputs.

We now illustrate this idea by analyzing one particular facet of the segmental inventory of Yidin^{y} (Kirchner 1992b). Our scope will be limited: the interested reader should examine the more comprehensive analysis of the Yidin^{y} inventory developed in Kirchner's work, which adopts the general Optimality Theory approach to inventories, but pursues different analytic strategies from the ones explored here.

The consonant inventory of Yidin^y looks like this:

Labial	*Coronal*	*Retroflex Coronal*	*Palatalized Coronal*	*Velar*
b	d		d^y	g
m	n		n^y	ŋ
	l			
	r	ŗ		

Here [r] is a "trilled apical rhotic" and [ŗ] an "apical postalveolar (retroflex) rhotic continuant," according to Dixon (1977:32).

Complex articulations are found only at coronal place of articulation; this is the generalization we wish to derive. The complexities include palatalization in [d^y, n^y] and the retroflexion in [ŗ]. (A similar but more articulated system is found in Lardil; see (148), §7.1, p. 109.) We propose to analyze the normal and palatalized coronals as follows, along lines developed in Clements 1976, 1991 and Hume 1992:

(279) **Representation of Coronals**

a. Normal b. Palatalized

In line with the findings of Gnanadesikan 1992 and Goodman in prep, we hold that retroflexion is dorsalization rather than coronalization (as it is in Kirchner 1992b). To focus the discussion, we will deal only with the coronalized coronals. As a compact representation of these structures, we will use bracketing to denote the structure of the Place node, according to the following scheme:

(280) **Bracketting Notation for Place Geometry**

a. [α] 'feature α occupies C-Place, there is no V-Place' node.

b. [α β] 'feature α occupies C-Place and feature β occupies V-Place'

With this notation, structure (279.a) is denoted by [Cor] and structure (279.b), by [Cor Cor].

In this representational system, the palatalized coronals are literally *complex*, with two places of articulation, while the other, unmarked coronals are literally simple. The generalization is now clear: of all the possible structurally complex places, only one is admitted into the Yidin^y lexicon: the one in which the primary and secondary places are both Cor — generally held to be the unmarked place of articulation (Avery & Rice 1989, and see especially the papers in Paradis & Prunet 1991, reviewed in McCarthy & Taub 1993).

Informally speaking, two generalizations are involved:

(281) **Coronal unmarkedness (observation).** "Don't have a place of articulation other than Coronal."

(282) **Noncomplexity (observation)**. "Don't have structurally complex places of articulation."

Our goal is to analyze the interaction between coronal unmarkedness and complexity markedness. This is of particular interest because it exemplifies a common pattern of interaction: each constraint is individually violated, but no form is admitted which violates both of them at once. There are consonants with single Lab or Dors specifications, violating coronal unmarkedness, and there are consonants with two place specifications, violating noncomplexity. But no consonant with any noncoronal place feature has a complex specification. We dub this generalization pattern *banning the worst of the worst*.

The worst-of-the-worst interaction is absent in the Basic CV Syllable Structure Theory. The two dimensions of well-formedness there — Onset well-formedness (more harmonic when present) and Coda well-formedness (more harmonic when absent) — operate independently. Requiring Onset, prohibiting Coda will generate the entire Jakobson Typology; the *worst-of-the-worst languages do not appear. Such a language would allows onsets to be absent, and codas to be present, but not in the same syllable; its inventory would include CV, V, CVC but exclude VC. This inventory is not possible according to the Basic CV Syllable Structure Theory, and we know of no reason to believe that this is anything but a desirable result.

The techniques already developed enable a direct account of the interaction between coronality and structural complexity. We assume that the input to the grammar is a string of root nodes each with a set of (unassociated) features. The output is an optimal parse in which these features are associated to root nodes (with the root nodes associated to syllable-position nodes, and so on up the prosodic hierarchy). To minimize distractions, let's assume a universally superordinate constraint requiring root notes to have a child PL (Place) node. (This parallels the assumption made in §6 that the syllable node always has a child Nuc, due to universal superordinance (123) of the relevant constraint NUC (119), p. 96.) For the present analysis of consonant inventories, we similarly assume a universally superordinate constraint, or restriction on Gen, to the effect that in consonants the presence of V-Place entails the presence of C-Place. (This head/dependent type of relationship is conveniently encoded in the bracketing notation of (280), because the configuration [α is always interpreted as 'α is C-Pl'.)

Our focus will be on which of the place features in an input feature set gets associated to the PL node. As always, unparsed input material is phonetically unrealized; underparsing is therefore a principal means of barring certain feature combinations from the inventory. If certain infelicitous combinations of features should appear in an input feature set, the grammar may simply leave some of them unparsed; the feature combinations which surface phonetically define a segmental inventory from which certain ill-formed feature combinations have been absolutely banned.

In Yidin^y, the feature set {Cor, Cor} gets completely parsed. Both Cor features are associated to the PL node in the optimal parse, and the segment surfaces as d^y or n^y, depending on which other features are in the set. On the other hand, the set {Lab, Lab} does *not* get completely parsed: the inventory does not include complex labials. In contrast, the unit set {Lab} does get completely parsed; the language has simple labials.

To minimize notation we will deal only with Cor and Lab; any other non-coronal place features receive the same analysis for present purposes as Lab.

Coronal unmarkedness can be formally stated as the following universal Harmony scale:

(283) **Coronal Unmarkedness, Harmony Scale:** PL/Cor $\succ$ PL/Lab

The notation 'PL/Cor' refers to a structural configuration in which PL dominates Cor, understood to be through some intermediate node — either C-Pl or V-Pl. The simplest theory, which we develop here, treats the two intermediate nodes alike for purposes of Harmony evaluation.

Following the same analytic strategy as for Universal Syllable Position/Segmental Sonority Prominence Alignment (213), §8, p.150, we convert this Harmony scale to a domination ranking of constraints on associations:

(284) **Coronal Unmarkedness, Domination Hierarchy:** *PL/Lab $\gg$ *PL/Cor

Following the general 'Push/Pull' approach to grammatical parsing summarized in §8.1 (186), the idea here is that all associations are banned, some more than others. The constraint hierarchy (284) literally says that it is a more serious violation to parse labial than to parse coronal. Coronal unmarkedness in general means that to specify PL as coronal is the least offensive violation. The constraint *PL/Lab is violated whenever Lab is associated to a PL node; this constraint universally dominates the corresponding constraint *PL/Cor because Lab is a less well-formed place than Cor. In addition to these two associational constraints we have the usual FAITHFULNESS constraints PARSE and FILL. They are parametrized by the structural elements they pertain to; in the present context, they take the form:

(285) **PARSE$^{\text{Feat}}$:** An input feature must be parsed into a root node.

(286) **FILL$^{\text{PL}}$:** A PL node must not be empty (unassociated to any features).

Just as with the segmental syllable theory, we have a set of deeply conflicting universal constraints: association constraints (*PL/Lab, *PL/Cor), which favor no associations, and FAITHFULNESS constraints which favor associations (PARSE$^{\text{Feat}}$ from the bottom up, FILL$^{\text{PL}}$ from the top down). This conflict is resolved differently in different languages by virtue of different domination hierarchies. The four constraints can be ranked in 4! = 24 ways overall; Universal Grammar, in the guise of Coronal Unmarkedness (283), rules out the half of these in which *PL/Lab is ranked below *PL/Cor, leaving 12 possible orderings, of which 8 are distinct. These induce a typology of segment inventories which includes, as we will shortly see, the Yidin^y case.

In languages with a wider variety of complex segments than Yidin^y, we need to distinguish an input which will be parsed as [Cor Vel] – a velarized coronal like [t^ɣ]– from an input which will be parsed as [Vel Cor] – a palatalized velar like [k^y]. (Both these segments occur, for example, in Irish and Russian). For this purpose we assume that the feature set in the first input is {Cor, Vel′} and in the second, {Cor′, Vel}; the notation f′ means that the feature f is designated in the feature set as secondary, one which is most harmonically parsed in the secondary place position. That is, we have the constraint:

(287) *** [f′**. f′ is not parsed as the primary place of articulation (not associated to C-Pl).

Since f and f′ designate the same place of articulation, parsing either of them incurs the same mark *PL/f; there are no separate marks *PL/f′ because *PL/f refers only to the place of articulation f.

Now we are ready to analyze the interaction between coronal unmarkedness and complexity in Yidin^y. The analysis is laid out for inspection in table (288):

(288) **Segmental Inventory**

Input POA's	*Candidates*	FILL^{PL}	*PL/Lab	PARSE^{Feat}	*PL/Cor	*[f′
Coronalized Coronal	a. ☞ [Cor Cor′]				* *	
{PL, Cor, Cor′}	b. [Cor′ Cor]				* *	* !
	c. [Cor] ⟨Cor′⟩			* !	*	
$t^y \mapsto t^y$	d. [Cor′] ⟨Cor⟩			* !	*	*
	e. [ ] ⟨Cor, Cor′⟩	* !		* *		
Labialized Labial	f. [Lab Lab′]		* * !			
{PL, Lab, Lab′}	g. ☞ [Lab] ⟨Lab′⟩		*	*		
$p^w \mapsto p$	h. [ ] ⟨Lab, Lab′⟩	* !		* *		
Coronalized Labial	i. [Lab Cor′]		* !		*	
{PL, Lab, Cor′}	j. [Lab] ⟨Cor′⟩		* !	*		
	k. ☞ [Cor′] ⟨Lab⟩			*	*	*
$p^y \mapsto t$	l. [ ] ⟨Lab, Cor′⟩	* !		* *		
Labialized Coronal	m. [Cor Lab′]		* !		*	
	n. ☞ [Cor] ⟨Lab′⟩			*	*	
{PL, Cor, Lab′}	o. [Lab′] ⟨Cor⟩		* !	*		*
$t^w \mapsto t$	p. [ ] ⟨Cor, Lab′⟩	* !		* *		
Simple Coronal {PL, Cor}	q. ☞ [Cor]				*	
$t \mapsto t$	r. [ ] ⟨Cor⟩	* !		*		
Simple Labial {PL, Lab}	s. ☞ [Lab]		*			
$p \mapsto p$	t. [ ] ⟨Lab⟩	* !		*		

The size of the table gives a misleading impression of intricacy. The idea behind this analysis is quite simple. Association must be forced, since the anti-association constraints *PL/α militate against it. The location of PARSE amid the anti-association constraints marks a kind of cut-off point: those *PL/α below PARSE are overruled and association of their α is compelled; those above PARSE, by constrast, are under no bottom-up pressure to associate. Only the top-down pressure of FILL will compel association – but since violations must be minimal, only minimal association can be forced. Glancing across the top of the tableau, one can see that all Cor's will be forced into association by PARSE, but Lab-association, driven only by FILL, will be minimal.

Here we give the details of the argument just outlined. Since *PL/Lab ≫ $\text{PARSE}^{\text{Feat}}$, it is more harmonic to leave Lab features unparsed (incurring *$\text{PARSE}^{\text{Feat}}$) than to associate them to PL (incurring *PL/Lab). Thus, *ceteris paribus*, Lab features remain unparsed.

The only reason that Lab nodes are ever parsed at all is to satisfy FILL^{PL}, which dominates *PL/Lab. FILL is exactly the *ceteris* that is not *paribus*. If the only features available in the set are Lab features, then failing to parse all of them would leave PL unfilled, earning a worse mark *FILL^{PL} than is incurred by parsing one of the Lab nodes.

On the other hand, only *one* Lab feature need be parsed to satisfy FILL^{PL}. When two are available, as in (f−h), parsing both would only increase the degree of violation of *PL/Lab. Since violations are minimal, the least necessary concession is made to FILL^{PL}. If two Labs are available in the set, one of them satisfies its intrinsic tendency to remain unparsed, while the other sacrifices this for the higher goal of ensuring that PL is not completely empty.

The situation is reversed for Cor, however; it is more harmonic to parse these features than to leave them unparsed, because $\text{PARSE}^{\text{Feat}}$ ≫ *PL/Cor.

As we see from the tableau, the Yidin^{y} inventory includes simple labials, as in rows (g,s), simple coronals, as in rows (k,n,q), and complex coronals as in row (a) but no other complex Places.[71] The grammar foils the attempt to create a complex labial from the input {PL,Lab,Lab′} in rows (f−h) by underparsing this set: a simple labial is output, as in (g), with one of the Lab features unparsed. The input {PL,Lab,Cor′}in rows (i−l) also fails to generate a complex segment, because the grammar parses only the Cor feature, outputting a simple coronal, row (k). The same output results from the input {PL,Cor,Lab′} of rows (m−p). This then is an instance of what we called 'Stampean Occultation' in §4.3.1; potential complex places involving Lab cannot surface, because the grammar always interprets them as something else, behind which they are effectively hidden. In the simplest case, the learner would never bother to posit them (see §9.3 for discussion).

[71] In the tableau, a label like 'Labialized Labial' for the input {PL,Lab,Lab′}is keyed to what *would result* from a faithful parse. The actual grammar underparses this input, and the output is a simple labial. Such labels are intended to aid the reader in identifying the input collocation and do not describe the output.

9.2 Universal Inventories

In addition to language-particular inventories, any theory must make possible an account of universal inventories. We have already seen a number of examples of universal inventory construction, and the preceding analysis of segmental inventories provides yet another, which we will now explore. The general issue of universal inventories has two aspects which we will exemplify; the following statements are intended to fix the terms of the discourse.

(289) **Absolute Universal Inventory Characterizations.**

a. **Absence**. A structure φ is absent from the universal inventory if, for every possible grammar and every possible input, the optimal output parse of that input for that grammar lacks φ.

b. **Presence**. A structure φ is universally present in language inventories if, for any possible grammar, there is some input whose optimal parse in that grammar contains φ.

(290) **Relative Universal Inventory Characterizations:**

An *implicational universal* of the form 'ψ in an inventory implies φ in the inventory' holds if, for every possible grammar in which there is some input whose optimal parse includes ψ, there is an input whose optimal parse in that same grammar includes φ.

The phrase 'possible grammar' refers to the well-formedness constraints provided by Universal Grammar, interacting via a particular domination hierarchy consistent with the domination conditions imposed by Universal Grammar.

9.2.1 Segmental Inventories

The segmental inventory of Yidin^y, barring only the worst-of-the-worst (complex, with at least one noncoronal Place), is but one of the inventories in the universal typology generated by the 12 possible domination hierarchies which can be constructed from the four constraints *PL/Cor, *PL/Lab, FILL^{PL}, PARSE^{Feat}, consistent with the universal domination condition (283) that yields Coronal Unmarkedness. This typology includes, for example, inventories which exclude all segments with complex places, and inventories which exclude all labials. The basic sense of the typology emerges from a couple of fundamental results, demonstrated below; these results correspond directly to the informal observations of Noncomplexity (282) and Coronal Unmarkedness (281), taken as implicational universals:

(291) **Complex ⇒ Simple.** $[\pi\ \psi'] \Rightarrow [\pi], [\psi]$

If the segment inventory of a language includes a complex segment with primary place π and secondary place ψ, it has a simple segment with place π and a simple segment with place ψ.

(292) **Lab ⇒ Cor** [···Lab···] ⇒ [···Cor···]

If the segment inventory of a language admits labials, it admits coronals.

a. **Harmonic Completeness w.r.t. Simple Segments:** [Lab] ⇒ [Cor]
If a language has simple labials, then it has simple coronals.

b. **Harmonic Completeness w.r.t. Primary Place:** [Lab ψ'] ⇒ [Cor ψ']
If a language has a complex segment with primary place Lab and secondary place ψ, then it has a complex segment with primary place Cor and secondary place ψ.

c. **Harmonic Completeness w.r.t. Secondary Place:** [π Lab′] ⇒ [π Cor′]
If a language has a complex segment with secondary place Lab and primary place π, then it has a complex segment with secondary place Cor and primary place π.

Recall that we are using 'Lab' to denote any non-coronal place of articulation. All noncoronals satisfy these implicational universals, because like Lab they all satisfy the Coronal Unmarkedness constraint domination condition (284). Both 'Lab' and 'Cor' should be taken here as no more than concrete place-holders for 'more marked entity' and 'less marked entity'.

Harmonic completeness means that when a language admits forms that are marked along some dimension, it will also admit all the forms that are less marked along that dimension. More specifically, if some structure is admitted into a language's inventory, and if a subpart of that structure is swapped for something more harmonic, then the result is also admitted into that language's inventory. Like the syllable structure results — for example, (215) of §8.2.1, p.152 — the implications **Complex ⇒ Simple** and **Lab ⇒ Cor** ensure harmonic completeness in exactly this sense.

These results entail that *only* harmonically complete languages are admitted by the constraint system, no matter what rankings are imposed. In other words, harmonic completeness in POA is a necessary condition for the admissibility of a language under the constraint system at hand. This result is not as strong as we would like: it leaves open the possibility that there are nevertheless some harmonically complete languages that the system does not admit. For example, if the factorial typology turned out to generate only those languages where the distinctions among the coronals were *exactly the same* as those among the labials, the theorems **Complex ⇒ Simple** and **Lab ⇒ Cor** would still hold true, for such languages are harmonically complete. (In fact, we know by construction that this is not the case: the Yidin^y hierarchy allows secondary articulations among the coronals but nowhere else.) What we want, then, is that harmonic completeness be also a sufficient condition for admissibility, so that *all* harmonically complete languages are admitted. Let us single out and name this important property:

(293) **Strong Harmonic Completeness (SHARC) Property.**

If a typology admits *all and only* the harmonically complete languages, then we say that it has Strong Harmonic Completeness (SHARC) .

If a typology has the SHARC, then it manifests what has been referred to in the literature as 'licensing asymmetry'. For place of articulation, in the circumscribed realm we have been examining, this comes out as follows:

(294) **POA Licensing Asymmetry**. In any language, if the primary place Lab licenses a given secondary place, then so does Cor; but there are languages in which the secondary places licensed by Cor are a strict superset of those licensed by Lab.

In the common metaphor, Cor is a 'stronger' licenser of secondary places than Lab. With the SHARC, there is the broader guarantee that every asymmetric system is possible. We know that the system of constraints examined here has the POA licensing asymmetry property, because harmonic completeness is a necessary property of admitted languages, and because we have produced at least one (Yidin^y) where the secondary articulations among the coronals are a strict superset of those permitted with labials. The factorial typology of the constraint system presented here does not in fact have the SHARC, as the reader may determine, but is a step in that direction.

It is worth noting that the SHARC is undoubtedly not true of POA systems in languages, and therefore not true of the entire UG set of constraints pertaining to POA. Indeed, it is unlikely that harmonic completeness is even a necessary condition on POA systems, as John McCarthy has reminded us. With respect to labialization, for instance, many systems have k^w or g^w with no sign of t^w or d^w. With respect to **Simple ⇒ Complex**, one recalls that Irish has velarized labials and palatalized labials, but no plain labials. McCarthy points to the parallel case of Abaza, which has pharyngealized voiceless uvulars but not plain ones. We do not see this as cause for dismay, however. Virtually any theory which aims to derive implicational universals must include subcomponents which, in isolation, predict the necessity of harmonic completeness and even its sufficiency as well. The constraints discussed here are a very proper subset of those relevant to POA. In particular, the key domination hierarchy is concerned only with context-free comparison of single features, and contains no information about effects of combination (labial+velar, round+back, ATR+high, etc.), which greatly alter the ultimate predictions of the system (Chomsky & Halle 1968: ch. 9, Cairns 1969, Kean 1974, Stevens & Keyser 1989, Archangeli & Pulleyblank 1992). Optimality Theory, by its very nature, does not demand that individual constraints or constraint groups must be *true* in any simple a-systematic sense. What this means is that an established subsystem or module can be enriched by the introduction of new constraints, without necessarily revising the original impoverished module at all. (We have already seen this in the transition from the basic syllable structure theory to the analysis of Lardil.) This fact should increase one's Galilean confidence that finding a subtheory with the right properties is a significant advance.

The POA subtheory examined here derives the relative diversity of coronals in inventory from the single fact of their unmarkedness. These two characteristics are so commonly cited together that it can easily be forgotten that underspecification theory cannot relate them. This important point comes from McCarthy & Taub 1993:

> Equally important as evidence for the unmarked nature of coronals is the fact that they are extremely common in phonemic inventories, where they occur with great richness

> of contrast… .[The] phonetic diversity of coronals is represented phonologically by setting up a variety of distinctive features that are dependent on the feature coronal....
>
> As explanations for different aspects of coronal unmarkedness, underspecification and dependent features are distinct or even mutually incompatible. By the logic of dependency, a segment that is specified for a dependent feature … must also be specified for the corresponding head feature … For example, even if the English plain alveolars *t, d, l, r* and *n* are underspecified for [coronal] the dentals *θ/ð* and palato-alveolars *č/ǰ/š/ž* must be fully specified to support the dependent features [distributed] and [anterior]. As a consequence, the dentals and palato-alveolars should not participate in the syndrome of properties attributed to coronal underspecification, and conversely, the plain alveolars should not function as a natural class with the other coronals until application of the [coronal] default rule.

It seems clear that the only way out is to abandon underspecification in favor of markedness theory (cf. Mohanan 1991). This is an ill-advised maneuver if it means embracing nothing more substantial than an elusive hope. The present theory shows that solid formal sense can be made of the notion of markedness, and, more significantly, that results about subtleties of inventory structure — permitted featural combinations — can be deduced from hypotheses about the relative markedness of individual atomic features. The coronal diversity result parallels the result in §8.3.2 that onsets are stronger licensers of segments than codas. In the syllable structure case, it is the structural markedness of the Cod node relative to the Ons node which impairs its ability to license segments. Here, licensing is diminished by the markedness of Lab as a place relative to Cor. Formally, the relationship of licenser to licensed is quite different in the two cases, but in both cases the markedness of the licenser governs its ability to license. We have, then, a very general mode of subtheory construction within Optimality Theory which allows us to argue from the markedness of atomic components to limitations on the structure of systems.

We now turn to the demonstrations of (291) and (292), with the goal of identifying a general technique for establishing such implicational universals.[72]

The argument establishing (291) runs as follows:

(295) **Proof of Complex ⇒ Simple:** For the case of the secondary place, i.e, proof that if a language has [π ψ′] it has [ψ]:

a. By definition of admission into the inventory, the output [π ψ′] must appear in an optimal parse of some input; the only possible such input is {PL,π,ψ′}.

---

[72] Another related technique, used in §8 and to an extended degree in Legendre, Raymond & Smolensky 1993, can be effectively used here as well; the results are more general but the technique is a bit more abstract. This other technique, which might be called the Technique of Necessary and Sufficient Conditions, goes as follows. Step 1: Determine necessary and sufficient conditions on the ranking of constraints in a hierarchy in order that each of the relevant structures be admitted into the inventory by that constraint ranking. Step 2: Examine the logical entailments that hold among these conditions: arguments of the form: in order to admit structure ϕ it is necessary that the constraints be ranked in such-and-such a way, and this entails that the constraint ranking meets the sufficient conditions to admit structure ψ. To carry out Step 1, to determine the necessary and sufficient conditions for a structure ϕ to be admitted, one takes a general parse containing ϕ and compares it to all alternative parses of the same input, and asks, how do the constraints have to be ranked to ensure that ϕ is more harmonic than all the competitors? And this in turn is done by applying the Cancellation/Domination Lemma, (192) of §8.2.6, p.142: for each mark *m incurred by ϕ, and for each competitor C, if *m is not cancelled by an identical mark incurred by C then it must be dominated by at least one mark of C.

In the present context, this technique gives the following results (Step 1):

(א) In order that [χ] be admitted into an inventory it is necessary and sufficient that:
either PARSE$^{\mathrm{Feat}}$ or FILL$^{\mathrm{PL}}$ ≫ *PL/χ

(ב) In order that [π ψ] be admitted into an inventory it is necessary and sufficient that:
a. PARSE$^{\mathrm{Feat}}$ ≫ *PL/ψ, and
b. either PARSE$^{\mathrm{Feat}}$ or *[f′ ≫ *PL/π, and
c. either PARSE$^{\mathrm{Feat}}$ or *FILL$^{\mathrm{PL}}$ ≫ *PL/π

From here, Step 2 is fairly straightforward. The result Complex ⇒ Simple (291) for the secondary place ψ follows immediately, since (ב.a) ⇒ (א) for χ = ψ. The result Complex ⇒ Simple for the primary place π follows similarly since (ב.c) ⇒ (א) for χ = π.

For the Harmonic Completeness results (292), we use the Coronal Unmarkedness domination condition (284)

*PL/Lab ≫ *PL/Cor

This means that whenever any of the domination conditions in (א) or (ב) hold of the feature Lab, it must also hold of the feature Cor; for in that case, each asserts that some constraint must dominate *PL/Lab, which means the same constraint must also dominate *PL/Cor since *PL/Lab ≫ *PL/Cor. Spelling this observation out in all the cases a–c of (292) proves the result Lab ⇒ Cor.

b. This means that [π ψ′] (incurring two marks *PL/π, *PL/ψ) must be more harmonic than all competing parses of the input {PL,π,ψ′}, including [π]⟨ψ′⟩ (incurring the marks *PL/π, *PARSE^{Feat}).

c. This entails that PARSE^{Feat} must dominate *PL/ψ.

d. This in turn implies that with the input {PL,ψ}, the parse [ψ] (incurring *PL/ψ) is more harmonic than its only competitor, [ ]⟨ψ⟩ (incurring *PARSE^{Feat} [as well as *FILL^{PL}]), hence [ψ] is the optimal parse.

e. Which means that the simple segment [ψ] is admitted into the segmental inventory.

Broadly put, the argument runs like this. Association must be compelled, over the resistance of the anti-association constraints. Either PARSE or FILL can be responsible. The existence of [π ψ′] in an optimal output guarantees that association of ψ is in fact compelled by the grammar and indeed compelled by PARSE, since FILL would be satisified by merely parsing π. Therefore, the association [ψ] must also occur, driven by PARSE. A similar but slightly more complex argument also establishes that [π] must be admitted.

The parallel argument establishing (292) is just a little more complicated:

(296) **Proof of Lab ⇒ Cor:** For the case of simple segments, (292.a):

a. If a grammar admits simple labials, then the feature Lab in some input feature set must get associated to PL: [Lab] must appear in the optimal parse of this input.

b. In order for this to happen, the association [Lab incurring *PL/Lab, must be more harmonic than leaving Lab unparsed (incurring *PARSE^{Feat}, and also possibly *FILL^{PL} if there are no other features in the set to fill PL).

c. This means the language's domination hierarchy must meet certain conditions: either

[i] PARSE^{Feat} ≫ *PL/Lab

or

[ii] FILL^{PL} ≫ *PL/Lab.

d. These conditions [i−ii] on the ranking of *PL/Lab entail that the same conditions must hold when *PL/Lab is replaced by the universally lower-ranked constraint *PL/Cor: since *PL/Lab ≫ *PL/Cor, by Coronal Unmarkedness (283), if [i], then:

[i′] PARSE^{Feat} ≫ *PL/Lab ≫ *PL/Cor;

if [ii], then:

[ii′] FILL^{PL} ≫ *PL/Lab ≫ *PL/Cor.

e. This in turn entails that parsing Cor must be better than leaving it unparsed: the input {PL,Cor} must be parsed as [Cor] (incurring *PL/Cor), since the alternative [ ] ⟨Cor⟩ would incur both *FILL^{PL} and *PARSE^{Feat}, at least one of which must be a worse mark than *PL/Cor by d.

f. This means that coronals are admitted into the inventory.

Again, the argument can be put in rough-and-ready form. Association must be compelled, either bottom-up (by PARSE) or top-down (by FILL). The appearance of [Lab — primary labial place — in an optimal output of the grammar guarantees that labial association has in fact been compelled one way or the other. Either a dominant PARSE or a dominant FILL forces violation of *PL/Lab 'don't have a labial place'. The universal condition that labial association is worse than coronal association immediately entails that the less drastic, lower-ranked offense of coronal association is also compelled, by transitivity of domination.

The two proofs, (295) and (296), illustrate a general strategy:

(297) **General Strategy for Establishing Implicational Universals $\psi \Rightarrow \varphi$**

a. If a configuration ψ is in the inventory of a grammar G, then there must be some input I_ψ such that ψ appears in the corresponding output, which, being the optimal parse, must be more harmonic than all competitors.

b. Consideration of some competitors shows that this can only happen if the constraint hierarchy defining the grammar G meets certain domination conditions.

c. These conditions entail — typically by dint of universal domination conditions — that an output parse containing φ (for some input I_φ) is also optimal.

9.2.2 Syllabic Inventories

The general strategy (297) was deployed in §8 for deriving a number of implicational universals as part of developing the Basic Segmental Syllable Theory. One example is the Harmonic Completeness of the inventories of Possible Onsets and Nuclei (215), which states that if τ is in the onset inventory, then so is any segment less sonorous than τ, and if α is in the nucleus inventory, then so is any segment more sonorous than α. A second example is (254), which asserts that if τ is in the inventory of possible codas, then τ is also in the inventory of possible onsets. That the converse is *not* an implicational universal is the content of Onset/Coda Licensing Asymmetry (258).

So far, our illustrations of universal inventory characterizations have been of the implicational or relative type (290). Examples of the absolute type (289) may be found in the Basic CV Syllable Structure Theory of §6. A positive example is the result (128), p. 98, that every syllable inventory contains CV, the universally optimal syllable. A negative example is the result (144), p.

105, which states that, in syllabic theory (which does not include constraints like LX≈PR), two adjacent empty syllable positions (phonetically realized as two adjacent epenthetic segments) are universally impossible: the universal word inventory, under the Basic Theory, includes no words with two adjacent epenthetic segments.

9.3 Optimality in the Lexicon

The preceding discussions have been independent of the issue of what inputs are made available for parsing in the actual lexicon of a language. Under the thesis that might be dubbed *Richness of the Base*, which holds that *all* inputs are possible in all languages, distributional and inventory regularities follow from the way the universal input set is mapped onto an output set by the grammar, a language-particular ranking of the constraints. This stance makes maximal use of theoretical resources already required, avoiding the loss of generalization entailed by adding further language-particular apparatus devoted to input selection. (In this we pursue ideas implicit in Stampe 1969, 1973/79, and deal with Kisseberth's grammar/lexicon 'duplication problem' by having no duplication.) We now venture beyond the Richness of the Base to take up, briefly, the issue of the lexicon, showing how the specific principles of Optimality Theory naturally project the structure of a language's grammar into its lexicon.

Consider first the task of the abstract learner of grammars. Under exposure to phonetically interpreted grammatical outputs, the underlying inputs must be inferred. Among the difficulties is one of particular interest to us: the many-to-one nature of the grammatical input-to-output mapping, arising from the violability of FAITHFULNESS. To take the example of the Yidin^y segmental inventory illustrated above in the tableau (288), two different inputs surface as a simple labial: the input {PL,Lab} which earns the faithful parse [Lab], and the input {PL,Lab,Lab′} which is parsed [Lab]⟨Lab′⟩. These outputs are phonetically identical: which underlying form is the learner to infer is part of the underlying segmental inventory? Assuming that there is no morphophonemic evidence bearing on the choice, the obvious answer — posit the first of these, the faithfully parsable contender— is a consequence of the obvious principle:

(298) **Lexicon Optimization**[73]. Suppose that several different inputs $I_1, I_2, \ldots, I_n$ when parsed by a grammar G lead to corresponding outputs $O_1, O_2, \ldots, O_n$, all of which are realized as the same phonetic form Φ — these inputs are all *phonetically equivalent* with respect to G. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled O_k. Then the learner should choose, as the underlying form for Φ, the input I_k.

---

[73] The term 'lexicon' here is really overly restrictive, since this is actually a principle for inducing underlying forms in general, not just those of lexical entries. For example, it can apply in syntax as well. The rules of the syntactic base might well generate structures such as $[[[[[\text{he}]_{DP}]_{DP}]_{DP}]_{DP}]_{DP}$ as well as simple $[\text{he}]_{DP}$. But, as we shall see, the principle (298) will imply that the simpler alternative will be selected as the underlying form.

This is the first time that parses of *different inputs* have been compared as to their relative Harmony. In all previous discussions, we have been concerned with determining the output that a given input gives rise to; to this task, only the relative Harmony of competing parses of the same input is relevant. Now it is crucial that the theory is equally capable of determining which of a set of parses is most harmonic, even when the inputs parsed are all different.

Morphophonemic relations can support the positing of input-output disparities, overriding the Lexicon Optimization principle and thereby introducing further complexities into lexical analysis. But for now let us bring out some of its attractive consequences. First, it clearly works as desired for the Yidin^y consonant inventory. Lexicon Optimization entails that the analysis of the Yidin^y constraint hierarchy (288) simultaneously accomplishes two goals: it produces the right *outputs* to provide the Yidin^y inventory, and it leads the learner to choose (what we hypothesize to be) the right *inputs* for the underlying forms. The items in the Yidin^y lexicon will not be filled with detritus like feature sets {PL,Cor,Lab′} or {PL,Lab,Lab′}. Since the former surfaces just like {PL,Cor} and the latter just like {PL,Lab}, and since the parses associated with these simpler inputs avoid the marks *PARSE^{Feat} incurred by their more complex counterparts, the needlessly complex inputs will never be chosen for underlying forms by the Yidin^y learner.[74]

Lexicon Optimization also has the same kind of result – presumed correct under usual views of lexical contents – for many of the other examples we have discussed. In the Basic CV Syllable Structure Theory, for example, Lexicon Optimization entails that the constraints on surface syllable structure will be echoed in the lexicon as well. In the typological language family $\sum^{CV}{}_{del,del}$, for example, the syllable inventory consists solely of CV. For any input string of Cs and Vs, the output will consist entirely of CV syllables; mandatory onsets and forbidden codas are enforced by underparsing (phonetic nonrealization). Some inputs that surface as [CV] are given here:

(299) **Sources of CV in $\sum^{CV}{}_{del,del}$**

a /CVV/ → .CV.⟨V⟩
b /CCVV/ → ⟨C⟩.CV.⟨V⟩
c /CCVVV/ → ⟨C⟩⟨C⟩.CV.⟨V⟩⟨V⟩

The list can be extended indefinitely. Clearly, of this infinite set of phonetically equivalent inputs, /CV/ is the one whose parse is most harmonic (having no marks at all); so *ceteris paribus* the $\sum^{CV}{}_{del,del}$ learner will not fill the lexicon with supererogatory garbage like /CCVVV/ but will rather choose /CV/. Ignoring morphological combination (which functions forcefully as *ceteris imparibus*) for the moment, we see that CV-language learners will never insert into the lexicon any underlying forms that violate the (surface) syllable structure constraints of their language; that is, they will always choose lexical forms that can receive faithful parses given their language's syllable inventory.

Morphological analysis obviously enlivens what would otherwise be a most boringly optimal language, with no deep/surface disparities at all. So let's add to our CV language some stem+affix morphology. Away from the stem/affix boundary, the lack of deep/surface disparities will clearly

---

[74] The Yidin^y system follows the pattern called 'Stampean Occultation' in §4.3.1 above. The principle of Lexical Optimization thus makes explicit the content of the Occultation idea.

remain, but at this boundary we can see a bit of interesting behavior beginning. As an example, we consider the CV idealization (with slight simplification) of the justly celebrated case of deep/surface disparity in Maori passives discussed in Hale 1973. The language is in the typological class we've called $\sum^{(C)V}$: onsets optional, codas forbidden; the paradigm of interest is illustrated in (300):

(300) **CV inflectional paradigm** (phonetic surface forms cited)

	uninflected	*inflected*
a.	CVCV	CVCV***V***
b.	CVCV	CVCV***CV***

The inflected form is composed exactly of the uninflected form followed by additional material, which in case (300a) consists only of V, and in case (300b) consists of CV. At issue is how to analyze the suffix. One analysis is shown in (301):

(301) **Phonological Analysis:**

	stem	+affix	uninflected	inflected
a.	CVCV	+***V***	.CV.CV.	.CV.CV.***V***.
b.	CVCV*C*	+***V***	.CV.CV.⟨*C*⟩	.CV.CV.*C**V***.

This analysis runs as follows: the inflection is a suffix +V; there are two classes of stems: V-final (a), and C-final (b); the means used to enforce the prohibition on codas is underparsing, so that in the uninflected form, the final C is not parsed. (In terms of the CV Syllable Structure Theory of §6, the language is $\sum^{(C)V}{}_{del}$.)

In his discussion of Maori, Hale calls this analysis the *phonological* analysis; in derivational terms, it calls on a phonological rule of final C deletion. This is to be contrasted with what Hale calls the *conjugation* analysis, shown in (302):

(302) **Conjugation Analysis:**

	stem	+affix	uninflected	inflected
I.	CVCV	+**V**	.CV.CV.	.CV.CV.**V**.
II.	CVCV	+**CV**	.CV.CV.	.CV.CV.**CV**.

Here class-I and class-II stems are distinguished in the lexicon with a diacritic whose force is to select different forms for the inflectional affix: +V in the first case and +CV in the second.

The question now is this: how do these two analyses fare with respect to the Lexicon Optimization principle (298)? In the conjugation analysis, the parses are completely faithful, incurring no marks at all. By contrast, the phonological analysis requires underparsing for uninflected C-final (class-II) stems: this incurs a *PARSE mark. Thus, as stated in (298), the Lexicon Optimization principle would appear to favor the conjugation analysis. Indeed, in general it favors analyses that minimize deep/surface disparities, and that maximize faithful parsing, thereby avoiding *PARSE and *FILL marks. Yet it is clear that in many circumstances, phonological analyses like (301) are to be preferred to conjugational analyses like (302). The deficiency in the formulation (298) of Lexicon Optimization is that it attempts a form-by-form optimization, without taking into consideration, for example, the optimization (minimization) of the number of allomorphs associated with an affix.

In general, morphological analysis entails that morphemes will appear in multiple combinations with other morphemes; the underlying form for a morpheme which is optimal (in the sense of Lexicon Optimization) when it appears in some combinations will not be optimal when it appears with others. The conjugational analysis avoids this by limiting the possible combinations (the class-I form of the affix, +V, can only co-occur with class-I stems), at the obvious cost of not minimizing the number of allomorphs for the affix. It seems clear that Lexicon Optimization must be reformulated so that, instead of form-by-form optimization, a more global optimization of the lexicon is achieved, in which more deep/surface disparities are accepted in order to minimize the constraints on allowed morphological combination which are part and parcel of conjugational analyses.

One simple way of formulating such a global lexicon optimization would be in terms of minimizing the totality of underlying material contained in the lexicon (precisely the kind of solution proposed in Chomsky & Halle 1968: ch. 8). Applied to the problem of deciding between the phonological and conjugational analyses illustrated by our CV example, such a Minimal Lexical Information approach would go something like this. The conjugation and phonological analyses share a common core consisting of a set of uninflected stems and the affix +V. In addition to this core, the conjugational analysis requires an additional allomorph for the affix, +CV, and a diacritic for each stem indicating which allomorph it takes. (In an actual language, an additional allomorph instantiating +CV would be needed for each possible consonant that can instantiate C.) The phonological analysis requires, in addition to the shared core, additional final Cs on the class-II stems. (In an SPE-style segmental derivational theory, the phonological analysis also requires an extra final C deletion rule, but in any modern syllabic theory this comes for free, courtesy of the same grammatical structure — for us, constraint hierarchy — that determines the syllable inventory in the first place.) Specification of all the stem-final Cs in the phonological analysis, and specification of all the diacritics distinguishing the conjugational classes in the conjugational analysis, require basically the same amount of lexical information — depending on details of accounting and of distribution of possible final Cs which we set aside here. What is left to differentiate the quantity of lexical information required by the two analyses is simply the additional allomorphic material +CV in the conjugational analysis. Thus, if the omitted details are properly handled, the principle of Minimal Lexical Information would appear to favor the phonological analysis — if only by the barest of margins.

It is quite possible that this accounting grossly underassesses the costs of multiple allomorphs. The true cost of unnecessary allomorphs may not be that of *having* them — as assessed by the additional underlying material they contain — but rather in the increased difficulty of *learning* them; more precisely, of learning to identify a morpheme which has multiple exponents, each with its own idiosyncratic limitation on the other allomorphs or stems with which it can combine. The problem of detecting the combinatorial structure underlying stem+affix may well be *much* easier when the affix has a unique exponent, even when compared to the case of just two allomorphs. Evidence bearing indirectly on this claim comes from a series of learning experiments carried out by Brousse & Smolensky (1989) and Brousse (1991) using connectionist ('neural') network learning techniques. Networks were trained to identify inputs which possessed the structure *stem+affix*, where *stem* was any member of a set S and *affix* any member of another set A; the network had to learn the classes S and A as well as the means of morphologically decomposing the inputs into *stem+affix*. This task was very robustly learnable; from a tiny proportion of exclusively positive examples, the networks acquired a competence (under a somewhat subtle definition) extending far beyond the training examples. Next, limitations on legal combination were imposed: networks had to learn to identify inputs with either of the legal forms $stem_I$+$affix_I$ or $stem_{II}$+$affix_{II}$, distinguishing them from other inputs, such as those with the illegal forms $stem_I$+$affix_{II}$ or $stem_{II}$ + $affix_I$. (Here $stem_I$ is any member of a set S_I of 'class-I stems', $stem_{II}$ any member of a set S_{II} of 'class-II stems', $affix_I$ any member of a set A_I of 'class-I affixes', and $affix_{II}$ any member of a set A_{II} of 'class-II affixes'.) This task was completely unlearnable by the same networks that had no trouble at all in learning the first task, in which stem/affix combination is not constrained in the way it must be in conjugational analyses. Thus there may be very strong learnability pressure to minimize combinatorial constraints, i.e., to minimize conjugational classes and the number of exponents of each morpheme.

While properly reformulating Lexicon Optimization from a form-by-form optimization to a global lexicon optimization is a difficult problem, one that has remained open throughout the history of generative phonology, a significant step towards bringing the Minimal Lexical Information principle under the scope of Lexicon Optimization as formulated in (298) is suggested by a slight reformulation, the Minimal Redundancy principle: to the maximal extent possible, information should be excluded from the lexicon which is predictable from grammatical constraints. Such considerations figure prominently, e.g., in discussions of underspecification (e.g. Kiparsky's Free Ride). An example of the consequences of this principle, if taken to the limit, is this: in a language in which *t* is the epenthetic consonant, a *t* interior to a stem which happens to fall in an environment where it would be inserted by epenthesis if absent in underlying form should for this very reason *be* absent in the underlying form of that stem. A rather striking example of this can be provided by the CV Theory. Consider a $\sum^{(C)V}{}_{ep}$ language (onsets not required; codas forbidden, enforced by overparsing — 'epenthesis'). The Minimal Lexical Redundancy principle would entail that a stem that surfaces as .CV.CV.CV. must be represented underlyingly as /CCC/, since this is overparsed as .C□́.C□́.C□́., which is phonetically identical to .CV.CV.CV.: it is redundant to put the V's in the lexicon of such a language. Given the constraints considered thus far, Lexicon Optimization as stated in (298) selects /CVCVCV/ and not /CCC/ in this case; again, avoiding deep/surface disparities whenever possible. But this is at odds with the principle that the lexicon should not contain information which can be predicted from the grammar. The approach to parsing we have developed suggests an interesting direction for pursuing this issue. As stated in (186), the Push/Pull Parsing approach views parsing as a struggle between constraints which *prohibit* structure and constraints which *require* structure. As noted in §3.1, the most general form of the structure-prohibiting constraint is *STRUC which penalizes any and all structure. There is a specialization of it which would be invisible during parsing but which can play an important role in learning:

(303) ***SPEC**: Underlying material must be absent.

Each underlying feature in an input constitutes a violation of this constraint[75]. But these violations cannot influence parsing since the underlying form is fixed by the input, and no choice of alternative output parses can affect these violations of *SPEC. But Lexicon Optimization is an inverse of parsing: it involves a fixed phonetic *output*, and varying underlying *inputs*; thus, among phonetically equivalent inputs, *SPEC favors those with fewest featural and segmental specifications.

Now an interesting change occurs if *SPEC outranks FAITHFULNESS: Lexicon Optimization (298) selects /CCC/ over /CVCVCV/ in the CV theory example — since minimizing FAITHFULNESS violations (and thereby deep/surface disparities) is now less important than minimizing underlying material. If on the other hand, FAITHFULNESS dominates *SPEC, we are back to /CVCVCV/ as the optimal underlying form.

Clearly a great deal of work needs to be done in seriously pursuing this idea. Still, it is remarkable how the addition of *SPEC to the constraint hierarchy can allow Lexicon Optimization — in its original straightforward formulation (298) — to capture an important aspect of the Minimal Lexical Information and Minimal Redundancy principles. It remains to be seen whether a constraint like *SPEC can supplant other possible constraints aimed specifically at limiting allomorphy, demanding (for example) a 1:1 relation between a grammatical category and its morphemic exponent. It is important to note that the addition of *SPEC makes no change whatever to any of the analyses we have considered previously. This raises the intriguing question of whether there are other constraints which are invisible to parsing — the operation of the grammar — but which play indispensable roles in grammar acquisition.

---

[75] The constraint is thus identical to the featural measure of lexical complexity in Chomsky & Halle 1968:381.

10. Foundational Issues and Theory-Comparisons

"If this is the best of all possible worlds, what are the others like?"
— *Candide, ou l'optimisme*, Ch. VI.

10.1 Thinking about Optimality

10.1.1 Fear of Optimization

We distinguish three species of qualm that have dissuaded people from thinking about optimality-based theories of linguistic form.

Q1. *Computation*. "Optimization is computationally intractable. Even simple optimization problems typically turn out to be inordinately expensive in terms of computational time and space. Many problems based on satisfaction of well-formedness conditions (much less *relative* well-formedness conditions) are even undecidable."

Q2. *Loss of Restrictiveness*. "In order to handle optimality, you must use numbers and use counting. The numerical functions required belong to a vast class which cannot be constrained in a reasonable way. Arbitrary quantization will be required, both in the weighting of degrees of concordance with (and violation of) individual constraints and in the weighting of the importance of disparate constraints with respect to each other. The result will be a system of complicated trade-offs (e.g. '*one* serious violation of $\mathbb{A}$ can be overcome when *three* moderate agreements with $\mathbb{B}$ co-occur with *two* excellent instances of $\mathbb{C}$'), giving tremendous descriptive flexibility and no hope of principled explanation. Therefore, the main goal of generative grammatical investigation is irredeemably undermined."

Q3. *Loss of Content*. "Appeal to scalar constraints — degrees of well-formedness — leads inevitably to a functionalizing narrative mush of the 'better for this/better for that' sort. By means of such push-pull, any imaginable state-of-affairs can be comfortably (if hazily) placed in a best of all possible worlds. Vagueness of formulation is reinstated as the favored mode of discourse, and Pullum's worst fears are realized."

10.1.2 The Reassurance

Q1. *Computation*. This qualm arises from a misapprehension about the kind of thing that grammars are. It is not incumbent upon a grammar to compute, as Chomsky has emphasized repeatedly over the years. A grammar is a function that assigns structural descriptions to sentences; what matters formally is that the function is well-defined. The requirements of explanatory adequacy (on theories of grammar) and descriptive adequacy (on grammars) constrain and evaluate the space of hypotheses. Grammatical theorists are free to contemplate any kind of formal device in pursuit of these goals; indeed, they *must* allow themselves to range freely if there is to be any hope of discovering decent theories. Concomitantly, one is not free to impose arbitrary additional meta-constraints (*e.g.* 'computational plausibility') which could conflict with the well-defined basic goals of the enterprise.

In practice, computationalists have always proved resourceful. All available complexity results for known theories are stunningly distant from human processing capacities (which appear to be easily linear or sublinear), yet all manner of grammatical theories have nonetheless been successfully implemented in parsers, to some degree or another, with comparable efficiency (see e.g. Barton, Berwick, & Ristad 1987; Berwick, Abney, and Tenny 1991.) Furthermore, it is pointless to speak of relative degrees of failure: as a failed image of psychology, it hardly matters whether a device takes twice as long to parse 5 words as it takes to parse 4 words, or a thousand times as long. Finally, real-world efficiency is strongly tied to architecture and to specific algorithms, so that estimates of what can be efficiently handled have changed radically as new discoveries have been made, and will continue to do so. Consequently, there are neither grounds of principle nor grounds of practicality for assuming that computational complexity considerations, applied directly to grammatical formalisms, will be informative.

Q2. *Loss of Restrictiveness through Arithmetic.* Concern is well-founded here. As we have shown, however, recourse to the full-blown power of numerical optimization is not required. *Order*, not *quantity* (or *counting*), is the key in Harmony-based theories. In Optimality Theory, constraints are ranked, not weighted; harmonic evaluation involves the abstract algebra of order relations rather than numerical adjudication between quantities.

Q3. *Loss of Content through Recourse to the Scalar and Gradient*. Here again there is a real issue. Recourse to functional explanations, couched in gradient terms, is often accompanied by severe loss of precision, so that one cannot tell how the purported explanation is supposed to play out over specific cases. A kind of informal terminological distinction is sometimes observed in the literature: a 'law' is some sort of functional principle, hard to evaluate specifically, which grammars should generally accord with, in some way or other, to some degree or other; a 'rule' is a precise formulation whose extension we understand completely. Thus, a 'law' might hold that 'syllables *should* have onsets,' where a 'rule' would be: 'adjoin C to V.' 'Laws' typically distinguish better from worse, marked from unmarked; while 'rules' construct or deform.

Linguistic theory cannot be built on 'laws' of this sort, because they are too slippery, because they contend obscurely with partly contradictory counter-'laws', because the consequences of violating them cannot be assessed with any degree of precision. With this in mind, one might feel compelled to view a grammar as a more-or-less arbitrary assortment of formal rules, where the principles that the rules subserve (the 'laws') are placed entirely outside grammar, beyond the purview of formal or theoretical analysis, inert but admired. It is not unheard of to conduct phonology in this fashion.

We urge a re-assessment of this essentially formalist position. If phonology is separated from the principles of well-formedness (the 'laws') that drive it, the resulting loss of constraint and theoretical depth will mark a major defeat for the enterprise. The danger, therefore, lies in the other direction: clinging to a conception of Universal Grammar as little more than a loose organizing framework for grammars. A much stronger stance, in close accord with the thrust of recent work, is available. When the scalar and the gradient are recognized and brought within the purview of theory, Universal Grammar can supply the very substance from which grammars are built: a set of highly general constraints which, through ranking, interact to produce the elaborate particularity of individual languages.

10.2 The Connectionism Connection, and other Computation-based Comparisons

10.2.1 Why Optimality Theory has nothing to do with connectionism

The term 'Harmony' in Optimality Theory derives from the concept of the same name proposed in 'Harmony Theory', part of the theory of connectionist (or abstract neural) networks (Smolensky 1983, 1984ab, 1986). It is sometimes therefore supposed that Optimality Theory should be classified with the other connectionist approaches to language found in the literature (McClelland & Kawamoto 1986; Rumelhart & McClelland 1986; Lakoff 1988, 1989; McMillan & Smolensky 1988; Stolcke 1989; Touretzky 1989, 1991; Elman 1990, 1991, 1992; Goldsmith & Larson 1990; Larson 1990, 1992; Legendre, Miyata, & Smolensky 1990abc, 1991ab; Hare 1990; Rager & Berg 1990; St. John & McClelland 1990; Berg 1991; Jain 1991; Miikkulainen & Dyer 1991; Touretzky & Wheeler 1991; Goldsmith 1992; Wheeler & Touretzky 1993 is a small sample of this now vast literature; critiques include Lachter & Bever 1988, Pinker & Prince 1988). Despite their great variety, almost all of these connectionist approaches to language fall fairly near one or the other of two poles, which can be characterized as follows:

(304) **Eliminativist connectionist models.** Representing the mainstream connectionist approach to language, the primary goals of these models are to show:

a. that basic analytic concepts of generative theory 'can be eliminated' in some sense;
b. that numerical computation can eliminate computing with symbolically structured representations;
c. that knowledge of language can be empirically acquired through statistical induction from training data.

(305) **Implementationalist connectionist models.** At the other pole, these models aim to contribute to the theory of language by studying whether (and how) more-or-less standard versions of concepts from generative grammar or symbolic natural language processing can be computationally implemented with connectionist networks. As with symbolic computational approaches, the claim is that limitations on what can be (efficiently) computed bear importantly on issues of language theory.

The conspicuous absence of connectionist models in this work shows how far Optimality Theory is from either of these poles; both eliminativist and implementationalist connectionism depend crucially on the study of specific connectionist networks. All three of the prototypical objectives in eliminativist research (304a–c) are completely antithetical to Optimality Theory. And as for the implementationalist approach, rather than arguing for the contribution of Optimality Theory based on issues of connectionist implementation, we have not even entertained the question.[76]

10.2.2 Why Optimality Theory is deeply connected to connectionism

That Optimality Theory has nothing to do with eliminativist or implementationalist connectionism is related to the fact that, fundamentally, Harmony Theory itself has little to do with eliminativist or implementationalist connectionism. Harmony Theory develops mathematical techniques for the theory of connectionist computation which make it possible to *abstract away* from the details of connectionist networks. These techniques show how a class of connectionist networks can be analyzed as algorithms for maximizing Harmony, and, having done so, how Harmony maximization itself, rather than the low-level network algorithms used to implement it, can be isolated as one of the central characteristics of connectionist computation. Optimality Theory constitutes a test of the hypothesis that *this* characterization of connectionist computation is one which can enrich — rather than eliminate or implement — generative grammar: by bringing into the spotlight optimization as a grammatical mechanism.

Optimality Theory has abstract conceptual affinities with Harmonic Grammar (Legendre, Miyata, & Smolensky, 1990a), a grammar formalism which is mathematically derivable from Harmony Theory and general principles concerning the realization of symbolic structure within distributed connectionist networks (Dolan 1989, Dolan & Dyer, 1987, Legendre, Miyata & Smolensky 1991, 1992, Smolensky 1987, 1990; relatedly, Pollack 1988, 1990). Harmonic Grammar is more intimately connected than Optimality Theory to a connectionist substrate: the relative strengths of constraints are encoded numerically, rather than through non-numerical domination hierarchies. Harmonic Grammar has been applied to the formulation of a detailed account of the complex interaction of syntactic and semantic constraints in unaccusativity phenomena in French (Legendre, Miyata, & Smolensky 1990bc, 1991; Smolensky, Legendre & Miyata, 1992).

It is instructive to ask what happens to Optimality Theory analyses when they are recast numerically in the manner of Harmonic Grammar. Suppose we assess numerical penalties for violating constraints; the optimal form is the one with the smallest total penalty, summing over the whole constraint set. A relation of the form $\mathbb{C}_1 \gg \mathbb{C}_2$ means considerably more than that the penalty for violating $\mathbb{C}_1$ is greater than that for violating $\mathbb{C}_2$. The force of strict domination is that *no* number of $\mathbb{C}_2$ violations is worth a single $\mathbb{C}_1$ violation; that is, you can't compensate for violating $\mathbb{C}_1$ by pointing to success on $\mathbb{C}_2$, no matter how many $\mathbb{C}_2$ violations are thereby avoided. In many real-world situations, there will be a limit to the number of times $\mathbb{C}_2$ can be violated by any given input; say, 10. Then if p_k is the penalty for violating $\mathbb{C}_k$, it must be that p_1 is greater than $10 \times p_2$.

[76] It is clear that Optimality Theory can be readily implemented using non-connectionist computation, and study of implementational issues in both connectionist and non-connectionist systems is a large open area for research.

The same reasoning applies on down the constraint hierarchy $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \cdots \gg \mathbb{C}_n \gg \cdots$. If $\mathbb{C}_3$ admits a maximum of 10 violations, then $p_2 > 10\times p_3$, and $p_1 > 10\times 10\times p_3$. For p_n, we'll have $p_1 > 10^{n-1}\times p_n$, if we cling artificially to 10 as the standard number of possible violations per constraint. The result is that, in order to represent the domination relation, the penalties must grow exponentially. Optimality Theory, on this practical construal, represents a very specialized kind of Harmonic Grammar, with exponential weighting of the constraints.

When we remove the artifice of limiting the number of violations per constraint, it becomes clear that the real essence of the domination idea is that the penalty for violating $\mathbb{C}_1$ is *infinitely* greater than the penalty for violating $\mathbb{C}_2$. The notion of Harmony in Optimality Theory, then, cannot be faithfully mapped into any system using standard arithmetic. Nevertheless, Optimality Theory is recognizable as a regimentation and pushing to extremes of the basic notion of Harmonic Grammar. The interested reader is referred to Smolensky, Legendre, & Miyata 1992, in which the relations between Harmonic Grammar, Optimality Theory, and principles of connectionist computation are subjected to detailed scrutiny.

10.2.3 Harmony Maximization and Symbolic Cognition

The relation of Optimality Theory to connectionism can be elaborated as follows (for extended discussion, see Smolensky, 1988, in press). In seeking an alternative to eliminativist and implementationalist connectionism, a natural first question to ask is whether, and how, connectionist principles might be capable of informing generative grammar. Suppose connectionism is viewed as a computational hypothesis concerning the structure of cognition at a level lower than that assumed in standard symbolic cognitive theory — a level closer to, but not as low as, the neural level. The question then becomes, how can connectionist computational principles governing this lower level of description constructively interact with principles operating at the higher level of description where grammar has traditionally been carried out? As a first step toward a reconciliation of the kinds of processes and representations assumed by connectionism to operate at the lower level with those assumed to operate in grammar, it seems necessary to find ways of introducing symbolic principles into connectionist theory and means of importing connectionist principles into symbolic theory. The former can take the form of new principles of structuring connectionist networks so that their representational states can be formally characterized at a higher level of analysis as symbolically structured representations.

In the reverse direction, principles of connectionist computation need to be introduced into symbolic grammatical theory. What principles might these be? Perhaps the most obvious are the principles of mutual numerical activation and inhibition which operate between connectionist units (or 'abstract neurons'). Work along these lines includes Goldsmith and Larson's *Dynamic Linear Model* (DLM), in which the levels of activity of mutually exciting and inhibiting units are taken to represent, typically, levels of prominence of adjacent phonological elements, e.g., derived sonority, stress (Goldsmith 1992, Goldsmith & Larson 1990, Larson 1990, 1992). As Goldsmith and Larson have shown, linguistically interesting behaviors are observed in these models; a variety of results to this effect have been proven in Prince 1993, which provides detailed formal analysis of the models, including explicit mathematical expressions characterizing their behavior. (It must also be noted that this analysis reveals a number of non-linguistic behaviors as well.)

Principles of mutual activation and inhibition are the lowest-level principles operating in connectionist networks. Rather than attempting to import such low-level principles to as high-level a theoretical enterprise as the theory of grammar, an alternative strategy is to identify the highest level principles to emerge from connectionist theory, and attempt to import these instead. Such high level principles are presently in very short supply. One of the few available is Harmony maximization.

Stripping away the mathematical technicalities, the principle of Harmony maximization can be couched in the following quite general terms:[77]

(306) **Connectionist Harmony Maximization.** In a certain class of connectionist network, the network's knowledge consists in a set of conflicting, violable constraints which operate in parallel to define the numerical Harmonies of alternative representations. When the network receives an input, it constructs an output which includes the input representation, the one which best simultaneously satisfies the constraint set — *i.e.*, which has maximal Harmony.

That Harmony maximization could be imported into phonological theory as a leading idea was suggested by Goldsmith (1990), working within the context of concerns about the role of well-formedness constraints in influencing derivations; it plays a central role in the model called Harmonic Phonology (Goldsmith 1990, 1993; Bosch 1991, Wiltshire 1992). (Immediately below, we examine some features of the class of models to which this belongs.) In line with our assessment of fundamentally different modes of interaction between connectionist and symbolic theories, it is important to recognize that the Dynamic Linear Model is conceptually and technically quite distinct from the family of linguistic models employing notions of Harmony, and in understanding its special character it is necessary to be aware of the differences. The DLM is a discrete approximation to a forced, heavily-to-critically damped harmonic oscillator. The networks of the DLM do not conform in general to the formal conditions on activation spread which guarantee Harmony maximization (either equally-weighted connections going in both directions between all units, or no feedback— see Prince & Smolensky 1991 for discussion). To the best of our knowledge, no Harmony function exists for these networks. Further, while Harmonic Phonology is based on symbol structures, the representations in the DLM are crucially numerical and non-structural. Thus, even if a Harmony function existed for the networks, it is unlikely that the activation passing in them can be construed as Harmonic Rule Application.

Optimality Theory, by contrast, seeks to strengthen the higher-level theory of grammatical form. It can be viewed as abstracting the core idea of the principle of Harmony Maximization and making it work formally and empirically in a purely symbolic theory of grammar. We see this as opening the way to a deeper understanding of the relation between the cognitive realm of symbol systems and the subcognitive realm of activation and inhibition modeled in connectionist networks. The property of strict domination is a new element, one quite unexpected and currently unexplainable from the connectionist perspective,[78] and one which is crucial to the success of the enterprise.

---

[77] This principle, and an appreciation of its generality and importance, is the result of the work of a number of people, including Hopfield 1982, 1984; Cohen & Grossberg 1983; Hinton & Sejnowski 1983, 1986; Smolensky 1983, 1986; Golden 1986, 1988; and Rumelhart, Smolensky, Hinton, & McClelland 1986. There are almost as many names for the Harmony function as investigators: it (or its negative) also goes by the names Lyapunov-, energy-, potential-, or goodness-function.

[78] For a possible line of explanation, see Smolensky, Legendre & Miyata 1992 (§3.3).

10.3 Analysis of 'Phonotactics+Repair' Theories

As discussed in §1, Optimality Theory is part of a line of research in generative syntax and phonology developing the explanatory power of output constraints. Most other research in this line has been derivational and in phonology has tended to use constraints only for surface- (or level-)[79] unviolated conditions: *phonotactics*. The fact that these phonotactics are surface-true arises in these derivational theories from a variety of factors, including the blocking of phonological processes which would lead to violation of a phonotactic, and the triggering of repair rules which take representations violating a phonotactic and modify them in one way or another so that the phonotactic holds of the result.

In these *Phonotactics+Repair* theories, interactions between phonotactics and repair rules can be handled in a variety of ways, across and within particular theories. An individual repair may be associated with individual phonotactics or not; they may be ordered with respect to other phonological rules or not; a phonotactic may block a rule or not. As we have observed throughout this work, all these patterns of interaction between phonotactics and repair rules have effects which are obtained in Optimality Theory from the single device of constraint domination. Domination yields not only the effects of phonotactic/repair interactions, but also accomplishes all the other work of the grammar, including the prosodic parse and the effects of what in derivational theories are general phonological processes. This constitutes a pervasive unification of what is expressed in other theories through a fragmented diversity of incommensurable mechanisms.

In this section we explicitly compare Optimality Theory to two representatives of the family of Phonotactics+Repair theories: the persistent rule theory (Myers 1991), and the Theory of Constraints and Repair Strategies (Paradis 1988ab).We will focus on one issue of central importance: comparing the notions of conflict which operate in Optimality Theory on the one hand and in Phonotactics+Repair theories on the other. We will see how a special case of Optimality Theoretic resolution of constraint conflict by ranking directly yields results which Phonotactics+Repair theories achieve by regulating phonotactic/rule interaction. The configuration at issue is one in which certain phonological rules are *not blocked* by a phonotactic, leading to intermediate derivational states in which the phonotactic is violated, at which point one of a set of possible repair rules is selected to restore compliance with the phonotactic. The cases we examine are treated within Optimality Theory via a straightforward interaction pattern which is quite familiar from descriptive work in the theory.

While we will not consider Harmonic Phonology explicitly here, the general comparative analysis we will provide is relevant to it as well, as a member of the Phonotactics+Repair (henceforth, 'P+R') family of theories. The interaction of phonotactics, repair rules, and the general rules implementing phonological processes is structured in Harmonic Phonology in the following way (Goldsmith 1990, 1993, Bosch 1991, Wiltshire 1992). The overall derivation involves several levels of representation. At each level certain specified phonotactics apply. At each level there is a set of rules, essentially repair rules, which apply freely within that level, governed by the principle of Harmonic Rule Application: "phonological rules apply … just in case their output is better than their input with

[79] Throughout this section, properties predicated of the 'surface' will refer to properties which hold either level-finally or derivation-finally.

respect to some criteria specified by a phonotactic (of the relevant level)" (Goldsmith 1993: 252). Rules apply one at a time within a level, but the derivational step to another level involves a parallel application of a specified set of rules which are not subject to Harmonic Rule Application. As with the P+R theories we treat specifically below, rules are used to achieve results which in Optimality Theory arise through the interaction of violable constraints. Optimality Theory differs crucially in explicitly assessing Harmony using constraints many of which are surface- (or level-) violated. A specific comparison on this point was provided in the Lardil analysis, §7.4; all effects in the Optimality Theoretic account are the result of harmonic evaluation, while the Harmonic Phonology perspective requires that crucial parts of the analysis be attributed to cross-level rules to which harmonic principles do not apply. The main reason for this difference is that crucial well-formedness conditions in the Optimality Theoretic analysis are *not* surface-unviolated phonotactics. The constraints, unlike phonotactics, come from Universal Grammar — they cannot be gleaned from inspection of surface forms. Indeed, there is no hope of constructing UG in this way if its constraints must be inviolable, and conversely, no hope of constructing individual grammars from inviolable constraints if they must be universal. In this situation, we argue, it is necessary to go for a strong UG rather than cling to the notion that constraints are *a priori* inviolable.

The central idea, then, which distinguishes Optimality Theory from other related proposals in the generative literature, notably Phonotactics+Repair theories, is this: constraints which are violated on the surface do crucial work in the grammar.[80] In alternative approaches, the work done in Optimality Theory by surface-violable constraints is generally performed by derivational rules. The issue of *conflict* is central here, since in Optimality Theory, surface violations of constraints arise only when they are forced by conflicts with more dominant constraints. Such conflicts between constraints in Optimality Theory are related in non-obvious ways to conflicts which arise in other theories between surface-unviolated constraints and derivational rules. Clarifying the relation between constraint/constraint conflict in Optimality Theory and rule/phonotactic conflict in P+R theories is a main goal of this section.

Examining the relation between these two kinds of conflict will allow us to compare Optimality Theory to a few specific P+R proposals in the literature. Our goal is to explicitly relate an important class of accounts based on a combination of surface-true constraints and derivational rules to Optimality Theoretic accounts based exclusively on constraints, both surface-true and surface-violated, and in the process to relate rule/phonotactic conflict in other theories to constraint/constraint conflict in Optimality Theory. For this purpose we will sketch particular Optimality Theoretic accounts of phonological interactions which are kept as close as possible to selected P+R accounts; and we will flesh out these Optimality Theoretic accounts just sufficiently to allow us to concretely illustrate the following general observations, for an interesting class of cases:

---

[80] We are grateful to Robert Kirchner and John McCarthy for clarificatory discussion. See Kirchner (1992bc).

(307) **The Rule/Constraint Divide**

a. The work done in P+R theories by specific repair rules is included under Optimality Theory in the consequences of general, violable constraints which function generally within the grammar to ensure correct parsing. In many cases, they are kinds of FAITHFULNESS constraints, versions of PARSE and FILL.

b. P+R theories distinguish sharply between phonotactics and repair rules, which must be treated entirely differently. Optimality Theory makes no such distinction, exploiting the single theoretical construct of the violable constraint: higher-ranked constraints end up surface-unviolated; lower-ranked ones, surface-violated. Avoiding the ontological phonotactic/repair-rule distinction considerably strengthens Universal Grammar, because the same constraint which is surface-violated in one language (correlating with a repair rule) is surface-unviolated in another (corresponding to a phonotactic). Universal Grammar provides violable constraints which individual grammars rank; whether a constraint appears as surface-violated ('repair-like') or surface-unviolated ('phonotactic-like') in a given language is a consequence of the constraint's ranking in the grammar.

c. Under Optimality Theory, the universally fixed function Gen supplies all structures; there are no special structure-building or structure-mutating processes that recapitulate the capacities of Gen in special circumstances. Because of Gen, the correct form is somewhere out there in the universe of candidate analyses; the constraint hierarchy exists to identify it. In a nut-shell: all constraint theories, in syntax as well as phonology, seek to eliminate the *Structural Description* term of rules; Optimality Theory also eliminates the *Structural Change*.

(308) **Conflict and Violation**

a. Conflict between a phonotactic $\mathbb{C}$ and a phonological rule R does *not* correspond in Optimality Theory to conflict between $\mathbb{C}$ and the constraint $\mathbb{C}_R$ which does the main work of the rule R; both $\mathbb{C}_R$ and $\mathbb{C}$ are surface-unviolated, hence the two constraints cannot be in conflict.

b. Instead, the conflict in Optimality Theory is between the pair $\{\mathbb{C}, \mathbb{C}_R\}$ on the one hand, and a third constraint $\mathbb{C}_{\text{rstr}}$ on the other: this third constraint is one which is violated by the repair rule which is used in the P+R theory to enforce the phonotactic $\mathbb{C}$.

c. One consequence of the P+R approach is the conclusion that constraints which are unviolated on the surface must nonetheless be violated at intermediate stages of derivations. In Optimality Theory, surface-violable constraints which do some of the work of repair rules eliminate the need for temporary violation of surface-unviolated constraints.

It is clear that the very idea of *repair strategies* demands that surface-inviolable constraints be violated in the course of derivations: a repair strategy is a derivational process which takes a representation in which a constraint is violated and mutates it into a representation in which the constraint is satisfied. Such a process cannot take place unless representations violating the constraint are present in the derivation.

A derivational process is by definition a sequential procedure for converting an input into an output; a sharp ontological distinction between input and output is not possible since there is a host of intermediate representations bridging them. Constraints need to be evaluated over all such representations, and part of the reason theoretical complexities arise is that constraints which are unviolated on the surface may be violated in underlying forms (especially after morphological combination) and then at least for some time during the derivation; or constraints may be initially vacuously satisfied because they refer to structure that is not yet constructed, but as soon as the relevant structure is built, violation occurs. In the non-serial version of Optimality Theory, however, there is a sharp ontological difference between inputs and outputs: Markedness constraints are evaluated only with respect to the output; and Faithfulness constraints, which value realization of the input, must also look to the output to make their assessments. Any surface-unviolated constraint is therefore literally entirely unviolated in the language, as it would be a basic category error to say that the constraint is violated by underlying forms.

While it is obvious that a repair strategy approach requires that surface-unviolated constraints be violated at least temporarily during derivations, it is much less obvious, of course, that the work of repair strategies can in fact be done by violable constraints, the violations of which are governed by domination hierarchies. Evidence for this conclusion is implicit in most of the preceding sections, since many of the Optimality Theoretic analyses we have presented use the freedom of Gen and constraint violation to do work which is performed in other accounts by repair rules. In this section, we focus on explicit comparison between Optimality Theoretic and a few P+R accounts from the literature. It turns out that several of these comparisons have a structure which is already implicit in one of our basic analyses, the CV syllable structure typology of §6. So before turning to our comparisons, we set up their general structure by examining this simplest case.

10.3.1 CV Syllable Structure and Repair

Consider syllabification in the typological class $\sum^{\mathrm{CV(C)}}{}_{\mathrm{ep}}$: onsets are mandatory, enforced by overparsing. In §6.2.2.1, we examined the input /V/, and saw that it was parsed as .□V., surfacing as a CV syllable with an epenthetic onset. A simple analysis of this situation using a surface-unviolated phonotactic and a repair rule runs as follows. The language has a syllable structure constraint SYLLSTRUC given by the template CV(C). (In the terminology of §6, SYLLSTRUC encapsulates the constraint package {ONS, NUC, *COMPLEX}.) A syllabification rule R_{syll} erects a σ node and associates V to it (via a Nuc or μ node). This onsetless σ violates SYLLSTRUC and this violation triggers a repair rule R_{ep}: a C-epenthesis rule.

Even in this utterly simple situation, the most basic of the problematic issues which loom large in richer contexts are already present. Why doesn't the constraint SYLLSTRUC *block* the syllabification rule in the first place? (Indeed, this is just what might be assumed in a comparable analysis of the *under*parsing case, $\sum^{\mathrm{CV(C)}}{}_{\mathbf{del}}$.) What general principle licenses temporary violation of SYLLSTRUC?

These questions disappear when derivational rules are replaced by Optimality Theory's violable constraints. The work of the syllabification rule R_{syll} is done by the constraint PARSE, which happens to be surface-unviolated in this language. The work of the repair rule R_{ep} is performed by the constraint FILL (more precisely, FILL$^{\mathrm{Ons}}$) which happens to be surface-violated in this language.

For theory comparison, the cross-theoretic relation between the derivational rule R_{syll} and the Optimality Theory constraint PARSE is an important one, for which it is convenient to have a name; we will say that PARSE is a *postcondition* of R_{syll}: the constraint PARSE gives a condition ('underlying material must be parsed into syllable structure') which is satisfied *after* the operation of the rule R_{syll}, which parses underlying material into syllable structure. There is not a unique postcondition associated with a rule; and not just any one will result in an Optimality Theoretic account that works. The theory comparison enterprise on which we now embark is certainly not a mechanical one. The postcondition relation is a strictly *cross*-theoretic notion; since it is not a notion internal to Optimality Theory, the fact that it is not uniquely defined in no way undermines the well-definition of Optimality Theory.

The relation between the repair rule R_{ep} and the constraint FILL is also cross-theoretically important, and different from that relating R_{syll} and PARSE. The repair rule R_{ep} does *not* apply *unless* it is necessary to save the SYLLSTRUC constraint; it is precisely the avoidance of FILL violations which prevents overparsing (epenthesis), except when necessary to meet SYLLSTRUC. A postcondition of the rule R_{ep} is thus SYLLSTRUC. FILL, on the other hand, is a *restraint* on the repair rule: it is *violated* when the rule fires. (Like postcondition, restraint is a cross-theoretic relation which is not uniquely defined.)

The Optimality Theory treatment employs the basic domination relation:

(309) SYLLSTRUC $\gg$ FILL

This is central to achieving the same result as that obtained in the P+R account by stating that the repair rule applies *when necessary* to meet SYLLSTRUC (but otherwise not, 'in order to avoid violating FILL'). The domination (309) illustrates the general case: when a repair rule R_{rep} has as postcondition a (surface-unviolated) constraint (corresponding to a phonotactic) $\mathbb{C}_{\text{tac}}$, and when this rule is restrained by a (surface-violated) constraint $\mathbb{C}_{\text{rstr}}$, we must have:

(310) $\mathbb{C}_{\text{tac}} \gg \mathbb{C}_{\text{rstr}}$

The restraining constraint must be subordinate to the constraint corresponding to the phonotactic.

This situation is summarized in the following table:

(311) $\sum^{\text{CV(C)}}{}_{\text{ep}}$:

	Rule	**Constraint**
Repair/Phonotactic	R_{rep} = C-Epenthesis	$\mathbb{C}_{\text{tac}}$ = SYLLSTRUC
Process/Postcondition	R_{proc} = Syllabification	$\mathbb{C}_{\text{proc}}$ = PARSE
Restraining Constraint	*violated by* R_{rep}*:*	$\mathbb{C}_{\text{rstr}}$ = FILL

Ranking: $\{\mathbb{C}_{\text{tac}}, \mathbb{C}_{\text{proc}}\} \gg \mathbb{C}_{\text{rstr}}$

i.e. {SYLLSTRUC, PARSE} $\gg$ FILL

This table, it turns out, captures the general structure of several more complex theory-comparisons which will be spelled out in analogous tables below (314, 316, 318). It is important to note, however, that the labeled columns do not partition the theories, and a point-for-point occamite match-up is not on offer here. Along with the cited rules, the P+R theory includes an exact image of the Markedness constraint(s) SYLLSTRUC; and in addition includes a condition (functioning like PARSE) that causes the Syllabification rule to fire in the presence of unsyllabified material, and another condition, analogous to Faithfulness in general and here functioning like FILL, which restrains C-epenthesis, restricting repair to minimal modification. On the OT side, there is Gen, which supplies candidates corresponding to those produced by C-Epenthesis and Syllabification (as well as many more).

In the P+R account, the parsing of /V/ involves a conflict between the constraint SYLLSTRUC and the rule R_{syll}: this is a case of rule/phonotactic conflict. The locus of conflict in the Optimality Theoretic account is elsewhere, however. The clearest Optimality Theoretic counterpart of the rule R_{syll} is its postcondition PARSE, and there is no conflict between SYLLSTRUC and PARSE; the conflict is between the pair of constraints {SYLLSTRUC, PARSE} on the one hand and FILL on the other: and it is FILL which gets violated. The rule/phonotactic conflict between SYLLSTRUC and R_{syll} arises in the P+R account from the fact that R_{syll} chooses to procedurally implement PARSE with a construction that implicitly also tries to implement FILL: for R_{syll} constructs a syllable with no unfilled nodes. To see the consequences of this, let's trace the status of the three constraints SYLLSTRUC, PARSE, and FILL during the simple derivation of /V/ → .□V.

(312) Constraint Violation History of a Simple Derivation

Step	Form	Rule	SYLLSTRUC	PARSE	FILL
0	/V/			*	
1	.V.	R_{syll}	*		
2	.□V.	R_{ep}			*

The rule R_{syll} eliminates the PARSE violation by parsing V into a new σ, but chooses to construct this σ in such a way as to avoid violations of FILL; that is, it constructs a σ with no onset. The problem is that this then creates a violation of SYLLSTRUC which next needs repair. The P+R analysis requires a stage of derivation, step 1 (shaded), in which a phonological rule has been allowed to produce a violation of a surface-unviolated rule. And, of course, such a violation is necessary to trigger R_{ep}. The Optimality Theory account, however, involves no violation of the surface-unviolated constraint SYLLSTRUC: it involves a violation of the surface-violated constraint FILL.

10.3.2 General Structure of the Comparisons: Repair Analysis

The simple example of syllabification of /V/ in $\sum^{CV(C)}{}_{ep}$ illustrates a very general situation. In the P+R account, the story goes as follows. At one stage of a derivation, the conditions of a phonological process (*e.g.* Syllabification) are met; the process applies, creating a structure which violates a phonotactic; the conditions of a repair rule now being met, the repair applies; and then the phonotactic is satisfied.

The Optimality Theory view of this P+R account goes like this. The surface-unviolated phonotactic is a high-ranking constraint $\mathbb{C}_{tac}$. The phonological process achieves some postcondition, another constraint $\mathbb{C}_{proc}$. $\mathbb{C}_{proc}$ is not 'blocked' in any sense by $\mathbb{C}_{tac}$ because in fact the two *do not conflict*: there is a way of satisfying them both, at the expense of a third constraint $\mathbb{C}_{rstr}$ which is lower-ranked than both $\mathbb{C}_{tac}$ and $\mathbb{C}_{proc}$.

There is no constraint/constraint conflict between $\mathbb{C}_{proc}$ and $\mathbb{C}_{tac}$ even though there is rule/phonotactic conflict between R_{proc} and $\mathbb{C}_{tac}$. This is because the rule R_{proc} enforcing $\mathbb{C}_{proc}$ introduces a stage of derivation in which $\mathbb{C}_{tac}$ is violated in order to meet $\mathbb{C}_{proc}$. But the subsequent repair produces an ultimate structure which meets both $\mathbb{C}_{tac}$ and $\mathbb{C}_{proc}$, which is possible only because there is no constraint/constraint conflict between these two constraints. The constraint/constraint conflict is actually between the pair $\{\mathbb{C}_{tac}, \mathbb{C}_{proc}\}$ on the one hand, and $\mathbb{C}_{rstr}$ on the other. In this conflict, $\mathbb{C}_{rstr}$ loses: it is the constraint violated by the repair rule.

The Optimality Theory account of the same situation is simply this (from (311)):

(313) $\{\mathbb{C}_{tac}, \mathbb{C}_{proc}\} \gg \mathbb{C}_{rstr}$

In an unproblematic input (e.g., /CV/ above), $\mathbb{C}_{tac}$, $\mathbb{C}_{proc}$ and $\mathbb{C}_{rstr}$ are all satisfied. In a problematic input (e.g., /V/ above), $\mathbb{C}_{tac}$ and $\mathbb{C}_{proc}$ together force the violation of $\mathbb{C}_{rstr}$.

This general comparative analysis can be applied to relate a variety of P+R accounts to Optimality Theoretic accounts of the same phonemena, as we now see. It is useful to have a name for this strategy: we call it *Repair Analysis*.

We reiterate the importance of distinguishing theory-comparative and Optimality Theory-internal notions in this discussion. Within Optimality Theory, all constraints have exactly the same status. The theory does not recognize, for example, a difference between 'violable' and 'inviolable' constraints. All constraints are *potentially* violable, and which ones happen to emerge as violated on the surface is a logical *consequence* of the domination hierarchy, the set of inputs, and the content of the constraints (which determines which of them conflict on the inputs). Similarly, although Repair Analysis distinguishes constraints as $\mathbb{C}_{tac}$, or $\mathbb{C}_{proc}$, or $\mathbb{C}_{rstr}$, this distinction is entirely theory-comparative: from the Optimality Theory-internal perspective, they are all simply violable constraints, interacting in the only way sanctioned by the theory: strict domination. The distinction between $\mathbb{C}_{tac}$, $\mathbb{C}_{proc}$, and $\mathbb{C}_{rstr}$ only arises in comparing an Optimality Theoretic account to a P+R account; they are constraints which relate to elements (e.g., phonotactics and repair rules) which have markedly different theoretical status in the P+R account — the constraints have identical theoretical status in Optimality Theory.

Two major features of subsequent Repair Analyses are also simply illustrated in the example of syllabification in $\sum^{CV(C)}{}_{ep}$. The first feature is generality: the constraints involved in the Optimality Theoretic account are extremely general ones, which function pervasively in the grammar to define well-formedness. The effects of a specific repair rule (epenthesis in a specific kind of environment) are derived consequences of the interaction of general well-formedness constraints.

The constraints in the Optimality Theoretic account are general in another sense, beyond their general applicability within the given language: the constraints in question are the same ones which operate in other languages exemplifying typologically different syllabification classes — this was exactly the point of the CV Syllable Structure Theory developed in §6. Thus the second important feature illustrated in this example is universality. From the perspective of comparison to P+R theories, the point is this: a constraint $\mathbb{C}$ may happen to be surface-unviolated in language L_1 and formalized as a phonotactic in a P+R theory, and the same constraint may well be operative but surface-violated in another language L_2 — and therefore not treatable as a phonotactic. $\mathbb{C}$ may play the role $\mathbb{C}_{tac}$ in L_1, but may be demoted to the role of a subordinate constraint $\mathbb{C}_{rstr}$ in L_2. In the Optimality Theoretic treatment, $\mathbb{C}$ may have exactly the same form in both languages; but in Phonotactics+Repair theory, this is completely impossible.

This situation has been exemplified in a number of cases discussed in previous sections, and is quite clear in the syllable structure example. In the language $L_1 = \sum^{CV(C)}{}_{ep}$ discussed in §10.3.1, the surface-unviolated constraint is $\mathbb{C}_{tac}$ = SYLLSTRUC = {NUC, *COMPLEX, ONS} while the surface-violated constraint is $\mathbb{C}_{rstr}$ = FILL. However for a language L_2 in the family $\sum^{(C)V(C)}$, for example, the roles of ONS and FILL are interchanged: now ONS is surface-violated while FILL is surface-unviolated. The constraint ONS is part of the phonotactic $\mathbb{C}_{tac}$ in L_1; and similarly −COD is also part of the corresponding phonotactic for a language L_1' in the family $\sum^{CV}$. Yet at least one of ONS and −COD must be active in $L_2 = \sum^{(C)V(C)}$, even though both are surface-violated (since /CVCV/ must be syllabified .CV.CV. rather than .CVC.V., a pair of legal syllables in L_2). To see how ONS and −COD are demoted from the status of $\mathbb{C}_{tac}$ in L_1 and L_1' to the status of $\mathbb{C}_{rstr}$ in L_2, consider the following P+R account of L_2. A Core Syllabification Rule builds core syllables, and if any unsyllabified segments remain after core syllabification, these defects are repaired by rules of Coda Attachment (for free Cs) and Onsetless Open Syllable Construction (for free Vs). The first repair rule is restrained by the constraint −COD and the second by ONS. So in a Repair Analysis of this P+R account of L_2, ONS and −COD fill the role of $\mathbb{C}_{rstr}$.

The fact that Optimality Theory has no equivalent of the phonotactic/rule dichotomy, but rather a single category of potentially violable constraint, makes it possible for Universal Grammar to simply specify a set of general constraints: the distinction between surface-violated and surface-unviolated, then, is a derived language-particular consequence of constraint ranking. These universal constraints capture generalizations which, in P+R terms, link what appear as phonotactics and postconditions in some languages to what are effectively restraining constraints on repair rules in others.

10.3.3 Persistent Rule Theory

In the preceding discussion we have been contrasting Phonotactics+Repair approaches, which use constraints for surface-unviolated phonotactics only, and Optimality Theory, which uses constraints much more widely. One recent analysis of the role of constraints in generative phonology is Myers 1991, which argues that constraints must be used *less* widely: only for a *subset* of phonotactics. Phonotactics are argued to divide into two classes which need to be theoretically treated in two different ways: one, as constraints which block phonological rules, the other, via *persistent rules* which do not block other phonological rules (but may in some cases undo their effects). The conclusion that the second class of phonotactics should not be treated as constraints but rather as the result of derivational rules is one which we now attempt to reconcile with Optimality Theory, in which such rules are eschewed in favor of constraints (surface-violated as well as surface-unviolated).

The repair rules of persistent rule theory (henceforth *PRT*) are ‘persistent’ in the sense that they are not ordered with respect to other rules, but rather apply whenever their conditions are met. The persistence of these rules does not, however, bear on the applicability of Repair Analysis: what matters is only that these rules are repair rules.

The arguments for PRT consist centrally in showing that a subset of phonotactics do not block phonological rules; that these rules apply, generating intermediate representations which violate the phonotactic, representations which are then repaired by a persistent rule. We will consider two such cases, and apply Repair Analysis to show how the necessary interactions fall out of very simple constraint interactions within Optimality Theory. We reiterate that our objective here is not at all to give full alternative treatments of these phenomena, but rather to illustrate the application of Repair Analysis to some relevant examples from the literature.

10.3.3.1 English Closed Syllable Shortening

A simple application of Repair Analysis is to Myers’ analysis (his §2.4) of English vowel length alternations like *keep/kept*, *deep/depth*, *resume/resumption*. The PRT analysis assumes a phonotactic which bars CVVC syllables; we can take this to be a constraint *μμμ barring trimoraic syllables. In the P+R account, this phonotactic does not block the process of Syllabification of, e.g., the final two consonants in *kept*, although the result of such syllabification is an illicit CVVC syllable, which then triggers a repair rule of Closed σ Shortening. The resulting derivation involves first associating the underlying long vowel of the stem *keep* to two syllabified moras, then associating *p* to a mora in the same syllable, then delinking the second mora for the vowel.

An Optimality Theoretic account of these interactions gives a straightforward application of Repair Analysis. The following table shows the relevant rules and constraints, in exact correspondence with the table for CV syllabification (311):

(314) **English Closed Syllable Shortening**

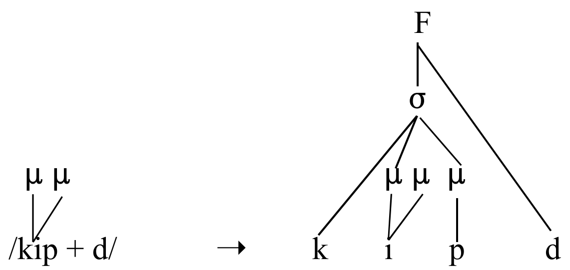
	Rule	**Constraint**
Repair/Phonotactic	R_{rep} = Closed σ Shortening	$\mathbb{C}_{\text{tac}}$ = *μμμ
Process/Postcondition	R_{proc} = Syllabification	$\mathbb{C}_{\text{proc}}$ = PARSE$^{\text{Seg}}$
Restraining Constraint	*violated by* R_{rep}:	$\mathbb{C}_{\text{rstr}}$ = PARSE$^{\mu}$

Ranking: $\{\mathbb{C}_{\text{tac}}, \mathbb{C}_{\text{proc}}\} \gg \mathbb{C}_{\text{rstr}}$ (313)

i.e.: {*μμμ, PARSE$^{\text{Seg}}$} $\gg$ PARSE$^{\mu}$

The phonotactic is $\mathbb{C}_{tac}$ = $*\mu\mu\mu$. The phonological process not blocked by the phonotactic is R_{proc} = Syllabification; the postcondition associated with this process is PARSE^{Seg}. The repair rule is R_{rep} = Closed σ Shortening. This is restrained by the constraint $\mathbb{C}_{rstr}$ = PARSE$^{\mu}$ which says that moras must be parsed (see §4.5): this is the constraint which must be violated in order to perform shortening. Here we assume the following analysis of *keep/kept*, designed to be minimally different from the PRT analysis:

(315)



Like the PRT account, ours does not treat the segmental alternations *i/*ε, *d/t*. For comparative purposes, we retain Myers' assumption of a (superordinate) constraint entailing that the final consonant is extrasyllabic (at least at this level of representation); we assume that segments attached to prosodic structure at any hierarchical level (e.g., the foot F) are phonetically realized, and that the failure to parse the second μ of *i* means the vowel is phonetically realized as short. We use failure to parse the second μ of *i* into σ here, rather than the PRT delinking of *i* to the second μ, in conformity with the general Optimality Theoretic principle that a parse of an input must always include the entire input representation (here, including two μs and their associations to *i*). The constraint ranking in (314) ensures that the parse in (315), which incurs the mark *PARSE$^{\mu}$, is the optimal one.[81]

We observe that the Optimality Theoretic treatment involves no violation of $*\mu\mu\mu$ whatsoever: there is no need for an intermediate stage of derivation in which the long vowel is fully parsed into syllable structure in order to provide the conditions for a Shortening rule; the second μ is simply never parsed. By using the (surface-violated, subordinate) constraint PARSE$^{\mu}$ instead of the derivational repair rule, the (surface-unviolated, superordinate) constraint $*\mu\mu\mu$ is spared even temporary violation.

In this case, the PRT account involves a conflict between the phonotactic $*\mu\mu\mu$ and the phonological process of Syllabification. But in the Optimality Theory account, without the Syllabification rule but with instead its postcondition PARSE^{Seg}, no conflict arises between this constraint and the constraint $*\mu\mu\mu$; rather, the conflict is between the combination {$*\mu\mu\mu$, PARSE^{Seg}} on the one hand, and the subordinate constraint PARSE$^{\mu}$ on the other. The Syllabification rule creates a conflict with $*\mu\mu\mu$ by doing syllabification in such a way as to satisfy PARSE$^{\mu}$ (in addition to PARSE^{Seg}): all μs are parsed into syllable structure, and this then needs to be undone by the repair rule.

[81] An absolutely direct assault is available if we recognize the 'canceled link' as a representational entity whose distribution is governed by a generalized version of PARSE.

In the Optimality Theoretic account, the correct interactions fall out directly from the simple constraint domination in (314), exactly the same domination pattern as in (311): the pattern characteristic of Repair Analysis. Furthermore, whereas repair rules (like Closed σ Shortening) are specialized rules which perform marked operations in order to overcome specific phonotactic violations created by phonological processes, the constraints which do the work in the Optimality Theoretic account are extremely general ones which are responsible for doing the central work in the grammar. Here, the repair rule of Closed σ Shortening performs the marked, anti-grammatical operation of Shortening or Delinking, in the very specifically circumscribed context of a Closed σ. By contrast, PARSE$^{\mu}$ and PARSE^{seg} are extremely general constraints which do the main grammatical work of ensuring the underlying material gets parsed … except when doing so would violate more highly-ranked constraints. The final 'except' clause is automatically furnished by the fundamental operation of the theory, and there is therefore no need to build the specific cases where this exception is realized into a specialized repair rule which undoes parsing exactly when it should never have occurred in the first place.

10.3.3.2 Shona Tone Spreading

Our second example of an argument from Myers 1991 against formalizing phonotactics as constraints is more complex: Shona Tone Spreading. Here the phonotactic is a prohibition on contour tones. The phonological process which is not blocked by this phonotactic are Association of syllables to tones, and Rightward Spread of High Tone. The repair rule is Simplification, which delinks the right tonal association of any syllable which has two tonal associations.

H tones are present in underlying forms, and L tones arise from a rule of Default. The basic spreading facts are that underlying H tones always dock and then spread right, unboundedly through a stem but only to the first σ after crossing a morphological boundary (which we'll denote '|'). The input /ku|mú|verengera/ — which has a single underlying H tone (denoted ´) on /mú/ — surfaces as *kumúvérengera*, with the underlying H tone spreading right one σ to yield *vé*, the remaining σs receiving default L tone.

The derivation of this output displays the now-familiar pattern in which the phonotactic is violated temporarily. On the first cycle, L tone is associated by default to the stem; next, with the innermost prefix /mú/ added, the underlying H tone spreads right to the first syllable of the stem, *ve*, which syllable is now doubly-linked to both H and L, in violation of the phonotactic; this double-association then triggers Simplification, which delinks the right (L) association of this syllable, satisfying the phonotactic.

The situation is more complex, but, with appropriate flexing, Repair Analysis applies here as well. The table corresponding to tables (311) and (314) is:

(316) **Shona Tone Spread**

	Rule	Constraint
Repair/Phonotactic	R_{rep} = Simplification (delink right T)	$\mathbb{C}_{tac}$ = *ǒ
Process/Postcondition	R_{proc}^{1} = Association	$\mathbb{C}_{proc}^{1}$ = PARSE^{T}
	R_{proc}^{2} = Spread H→	$\mathbb{C}_{proc}^{2}$ = *ő σ
Subordinate Constraints	*potentially violated by* R_{proc}^{2}:	$\mathbb{C}_{rstr}^{1}$ = *T†σ
	R_{sub}^{2} = Spread H→	$\mathbb{C}_{proc}^{2'}$ = *όσ
	R_{sub}^{3} = Default T is L	$\mathbb{C}_{rstr}^{2}$ = *ό

Ranking: $\{\mathbb{C}_{tac}, \mathbb{C}_{proc}\} \gg \mathbb{C}_{rstr}$ (313)

i.e. $\{\mathbb{C}_{tac}, \mathbb{C}_{proc}^{1}, \mathbb{C}_{proc}^{2}\} \gg \mathbb{C}_{rstr}^{1} \gg \mathbb{C}_{proc}^{2'} \gg \mathbb{C}_{rstr}^{2}$

The Optimality-theoretic constraint corresponding to the phonotactic is written *σ̌, where σ̌ denotes a doubly-associated σ. The first phonological process not blocked by this phonotactic, Association, has a postcondition PARSE^{T}. The second such process, Spread H→, requires a bit more discussion.

For present purposes, the following treatment of spreading will suffice. In Optimality Theory the phonological and morphological content of the input is assumed to be an identifiable part of every candidate output. In an output, let σ́ denote a syllable associated to H, and σ̋ a syllable associated to a tautomorphemic H. Then we can capture Shona's rightward spreading as a pair of constraints, the first of which is *σ́σ: this is violated when an L-bearing σ follows an H-bearing σ. The second, higher-ranked constraint is *σ̋σ, violated when an L-bearing σ follows a σ bearing a tautomorphemic H tone.[82]

We assume without further comment that a familiar set of constraints are imposed on the candidate sets generated by Gen: *e.g.*, all σs must be associated to tones, association lines do not cross, the OCP.

The Default rule of L-association is motivated by a (lowest-ranked) constraint barring H-bearing syllables: *σ́. The only remaining constraint is one which assesses a mark for each tonal association line which crosses a morphological boundary; we iconically name this constraint *T†σ.

The Optimality Theory analysis follows the general form of the previous cases: compare the constraint ranking in (316) to those in (311) and (314). Now, however, the lower-ranking constraints corresponding to $\mathbb{C}_{rstr}$ are crucially ranked with respect to one another. Because *T†σ is ranked lower than *σ̋σ and higher than *σ́σ, it follows that H tones will spread rightward, exactly one syllable across a morphological boundary: the rightmost H-bearing σ is then σ́ rather than σ̋; so failing to continue to spread violates *σ́σ, but continuing to spread violates higher-ranked *T†σ. Heteromorphemic H-tone association (violating *T†σ) can only be forced by *σ̋σ, which is satisfied

[82] A more complete analysis would attempt to capture the connection between *T†σ, which penalizes hetero-morphemic association, and the subhierarchy *σ̋σ ≫ *σ́σ, which asserts that hetero-morphemic associations are weaker licensers of "spreading".

by spreading a single syllable across the morpheme boundary. Because *σ́ is lowest-ranked, it will be violated in optimal forms in which H tones spread: but it ensures that any syllable not receiving H tone by spread from an underlying H tone will bear a L tone.

As in the preceding examples, in the Optimality Theory account, the constraint $\mathbb{C}_{tac}$ is not (even temporarily) violated. In /ku|mú|verengera/ → .ku|mú|vérengera., the syllable *ve* does not have the tonal association history ∅ → L → HL → H which it undergoes in the Phonotactics+Rules account; it never receives L-tone in the first place, and there is thus no need to revoke it. And like the previous example, the work of a specialized repair rule is done by extremely general constraints which are also responsible for doing the primary grammatical work of correctly assigning parses to inputs.

10.3.3.3 Summary

The analysis in this section is obviously preliminary, and the conclusion therefore a tentative one. According to Persistent Rule Theory, the role of constraints in grammar is restricted to that subset of surface-unviolated phonotactics which block phonological rules; other phonotactics arise as the consequence of persistent rules. We have seen that the failure of a phonotactic to block a phonological process is an inevitable outcome within a constraint-based theory in which there is no conflict between the constraint $\mathbb{C}_{tac}$ corresponding to the phonotactic and the constraint $\mathbb{C}_{proc}$ which is a postcondition of the phonological process; these two constraints can be simultaneously met, at the expense of a third subordinate constraint $\mathbb{C}_{rstr}$ which is surface-violated; this additional constraint is what is violated by the operation of the persistent repair rules in PRT. Optimality Theory handles such cases straightforwardly via the simple domination condition (313):

$$\{\mathbb{C}_{tac}, \mathbb{C}_{proc}\} \gg \mathbb{C}_{rstr}$$

The constraints involved in the Optimality Theoretic account are highly general ones which do the primary work of the grammar: but because the complete set of constraints $\mathbb{C}_{tac}$, $\mathbb{C}_{proc}$, $\mathbb{C}_{rstr}$ cannot all be simultaneously satisfied in certain special problematic cases, something has to give, and the domination hierarchy determines that it is $\mathbb{C}_{rstr}$. The special situations in which this subordinate constraint is violated are a logical *consequence* of the account, rather than a stipulated environment hand-wired into a specialized repair rule. Recall that the labels on the constraints are intended merely as guides to cross-theory comparison. The root prediction is that all domination orders are possible, yielding a typology of different systems, in which of course, from the operationalist point of view, there would be different constraints and different repairs.

10.3.4 The Theory of Constraints and Repair Strategies

Whereas Myers 1991 argues for restricting the role of constraints in grammar, Paradis has forcefully argued the opposite position. Optimality Theory builds on her work and further promotes the role of constraints, adding a theory of constraint interaction in which lower-ranked constraints are violated in order to satisfy higher-ranked ones. In this section we explicitly consider key aspects of the relation between the approaches, concretely grounding the discussion in a specific illustrative analysis from Paradis 1988 (to which paper page number citations in this section refer).

Like the theories considered previously, Paradis' Theory of Constraints and Repair Strategies (*TCRS*) is a derivational Phonotactics+Repair framework, in which all constraints explicitly treated as such are surface-unviolated phonotactics. Thus the constraints in TCRS cannot conflict in the sense of Optimality Theory. We need to properly understand, then, statements such as the following:

> "All these facts lead me to conclude that phonological processes do not freely violate phonological constraints. Actually, violations occur when there is a *constraint conflict*, which must be solved in some way. I argue that this is accomplished by the PLH."
>
> p. 90, emphasis added.

'PLH' refers to the "phonological level hierarchy … : metrical > syllabic > skeletal > segmental" (p. 89). An example of Paradis' sense of 'constraint conflict' is provided by her analysis of Fula, about which she says

> "the constraint violation … follows from a *conflict of two constraints*: the obligatory Segmental Licensing Convention for skeletal slots … (no floating slot); and the constraint against continuant geminates …"
>
> p. 89, emphasis added.

The derivation she refers to has a now-familiar form:

(317) **Fula Gemination**: An example derivation

Step	Form	Rule	*V:C:	*GEMCONT	FILL^{X}	PARSE^{X}	PARSE^{feat}
0	XXXX **X**X | |/ | | l a w + i	Input			*	******	
1	σ σ |\ | XXXX **X**X | |/ |/ | l a w + i	Spreading, Nucleus Syllabification		*		* * *	
2	σ σ /|\ /| XXXX **X**X | |/ |/ | l a b + i	Feature Changing, Onset Syllabification				*	*
	blocked	Coda Syllabification					
3	σ σ /|\ /| XXXX **X**X | |/ |/ | l a b i	Segmental Delinking, Skeletal Deletion				*	*

Fula has a phonotactic barring geminate continuant consonants. This phonotactic is temporarily violated during the derivation (317), at step 1 (shaded).

The strategy of Repair Analysis applies to this analysis as well. The following table is the counterpart of (311), (314), and (316):

(318) **Fula Gemination**

	Rule	Constraint
Repair/ Phonotactic	$R_{rep}^{\,1}$ = Segmental Delinking and Skeletal Deletion	$\mathbb{C}_{\mathrm{tac}}^{\,1}$ = *V:C:
	$R_{rep}^{\,2}$ = [+cont] → [−cont]	$\mathbb{C}_{\mathrm{tac}}^{\,2}$ = *GEMCONT
Process/ Postcondition	R_{proc} = Spreading	$\mathbb{C}_{\mathrm{proc}}$ = FILL$^{\mathrm{X}}$
Restraining Constraint	*violated by* $R_{rep}^{\,1}$:	$\mathbb{C}_{\mathrm{rstr}}^{\,1}$ = PARSE$^{\mathrm{X}}$
	violated by $R_{rep}^{\,2}$:	$\mathbb{C}_{\mathrm{rstr}}^{\,2}$ = PARSE$^{\mathrm{feat}}$

Ranking: $\{\mathbb{C}_{\mathrm{tac}}, \mathbb{C}_{\mathrm{proc}}\} \gg \mathbb{C}_{\mathrm{rstr}}$ (313)
i.e. $\{\mathbb{C}_{\mathrm{tac}}^{\,1}, \mathbb{C}_{\mathrm{tac}}^{\,2}, \mathbb{C}_{\mathrm{proc}}\} \gg \{\mathbb{C}_{\mathrm{rstr}}^{\,1}, \mathbb{C}_{\mathrm{rstr}}^{\,2}\}$

There are two phonotactics involved; the first is the prohibition on geminate continuant consonants. In the Optimality Theoretic analysis, this is captured in a constraint *GEMCONT which is violated whenever a consonantal root node is associated to the feature [+cont] and to two skeletal positions (regardless of whether those skeletal positions are parsed into syllable structure). The other phonotactic (related to that of the English Closed σ Shortening example discussed in §10.3.3.1) is a prohibition on the sequence V:C:. In the Optimality Theoretic analysis this takes the form of a constraint, *V:C:, which is violated if a V is associated to two Xs, a following C is associated to two Xs, and all four Xs are parsed into prosodic structure. (To focus on differences between the Optimality Theory and TCRS frameworks, rather than substantive phonological assumptions, we stick as close as possible to the substantive assumptions of Paradis 1988.)

The phonological process which is not blocked by the phonotactics is Spreading; in step 1 of (317), this geminates a continuant consonant in violation of *GEMCONT. This violation is repaired by the rule [+cont] → [−cont] which changes *w* to *b* in step 2. A postcondition of the Spreading rule is that the Xs previously unassociated to underlying segments are now so associated: we dub this postcondition FILL$^{\mathrm{X}}$.

At step 2 of the derivation (317) the first X associated to *b* is unparsed (its syllabification being blocked by *V:C:); this violates a 'no floating slot' constraint. The repair for this are rules of Segmental Delinking and Skeletal Deletion. In the TCRS treatment, the '$\mathbb{X}$' in step 3 has been deleted from the representation; in the Optimality Theoretic treatment, this skeletal position is part of the output (as it must be since it is part of the underlying form) but this '$\mathbb{X}$' is not parsed into prosodic structure and therefore the consonant *b* associated to it does not surface long. Thus in the Optimality Theoretic account, the constraint violated by the repair rule of Skeletal Deletion is PARSE$^{\mathrm{X}}$.

The effects of the other repair rule, [+cont] → [−cont], arise naturally within an Optimality Theoretic account following the treatment of segment inventories in §9.1.2. We will assume that the underlying segment denoted *w* in the example (317) is actually a set of features which need to be parsed (associated to a root node) in order to surface. The feature [+cont] would normally be so parsed (to satisfy PARSE^{feat}), except that in the particular problematic case of (317), parsing this feature would violate *GEMCONT, so underparsing is forced; we assume the segment is phonetically realized as [−cont] as a result of the failure to parse [+cont]. This failure to parse [+cont] constitutes a violation of PARSE^{feat}, which is crucially lower-ranked than FILL^{X} (318): this is why in the optimal analysis (the final representation in (317), with '$\mathbb{X}$' present but unparsed into syllable structure), the floating position '**X**' of the suffix is filled (associated to '*b*', with '*b*' having an unparsed [+cont] feature) rather than unfilled (unassociated to a segmental root node).

In the TCRS derivation (317), the segment *w/b* is first associated (step 1) and later delinked (step 3) to a second X. From the perspective of the Optimality Theoretic account, there is another association/dissociation: the feature [+cont] of this same segment is first parsed into the root node (step 0) then unparsed (step 2). The intermediate stage (step 1) at which the phonotactic *GEMCONT is violated is necessary in the derivational account to trigger the repair rule [+cont] → [−cont] which applies only to geminate consonants. This derivation exhibits *opacity* in the sense of Kiparsky 1973a.

How does the Optimality Theory account explain the opaque outcome, where the consonant surfaces as *b* even though it does not surface long? That is, how does it explain the optimality of the final representation in (317)? As follows. High-ranked FILL^{X} forces the floating '**X**' to be filled, which is achieved by associating it to the root node of underlying *w*; *GEMCONT forces [+cont] not to be parsed into this root node, violating lower-ranked *PARSE^{feat}; this explains why *b* surfaces. At the same time (and in fact, logically independently), *V:C: forces '$\mathbb{X}$' not to be parsed, violating lower-ranked *PARSE^{X}; this is why *b* surfaces short. Note that the exact formulation of the constraints *GEMCONT and *V:C: are crucial to this explanation. *GEMCONT, as formulated above, would be violated if [+cont] were parsed into the root node — even though one of the two skeletal positions to which this root node is associated, $\mathbb{X}$, is not itself parsed into syllable structure.

Furthermore, *V:C:, as formulated above, is not violated in the optimal parse precisely because 𝕏 is unparsed.[83]

We now return to the issue of constraint conflict within Paradis' account. As quoted earlier, she states that the constraint 'no floating slot' conflicts with the constraint *GEMCONT; however, in the end, both constraints are satisfied, appropriate repairs having been made. The relevant sense of 'constraint conflict' here centers around the crucial step 1 of the derivation (317): gemination of underlying *w*, in conflict with *GEMCONT, but necessary to trigger [+cont] → [−cont]. The question Paradis raises in her discussion is: why does gemination occur at step 1 rather than Skeletal Deletion

---

[83] The Fula analysis in Paradis (1988) has further complexities, and in (318) we have focussed on the aspects most relevant to the issues under consideration in this section: a general Optimality Theoretic analysis of Phonotactics+Repair theories, with specific attention to what Paradis terms 'constraint conflict' in TCRS. Going beyond the elements displayed in (318), several additional factors involved in the Fula analysis are addressed in the text and footnotes of Paradis 1988. Questions about the derivation (317) which must be addressed include these two:

(a) Why is the free X in the suffix filled by spreading from the preceding consonant rather than from the following vowel?

(b) Why does the long vowel in the stem not shorten in order to satisfy the *V:C: constraint?

Like the core of the TCRS analysis displayed in (318), the core of the Optimality Theoretic analysis presented there must be supplemented in order to address these questions. Given that our objective here is not at all to provide a full treatment of Fula, we content ourselves with a few remarks about how the observations Paradis makes in answer to these questions straightforwardly suggest Optimality Theoretic constraints which satisfactorily complete the treatment of our example (317); properly testing these possible constraints across Fula goes beyond our goals here.

To question (a), Paradis offers two possible explanations: "[i] a nuclear segment does not spread to a non-nuclear position if a non-nuclear segment is available. Moreover, [ii] nouns cannot end with long vowels in Fula." (p. 90; numbers added.) The second possible explanation [ii] refers to a surface-unviolated constraint which could obviously be added to the top of the constraint hierarchy in the Optimality Theoretic account. The first proposed explanation [i] is more interesting, and since it has the 'except when' structure, it is a natural product of an Optimality Theoretic treatment involving constraint conflict. One possible such analysis would involve a constraint against vowels associated to two skeletal positions (which may or may not be parsed into syllable structure): '*V:'. Obviously, this is surface-violated; violations are forced by a dominating constraint FILL^{X}. The constraint *V:, dominated by FILL^{X}, yields explanation [i] as a consequence: it entails that a free X will be associated to a vowel only if a consonant is not available. In (317), a consonant *is* available (even though it happens to be [+cont]).

Paradis does not explicitly address question (b) in the context of the derivation (317), but in a footnote discussing CV:C syllables generally, she observes that in a variety of environments, shortening does not occur to avoid such syllables (note 4). This discussion would suggest a surface-unviolated constraint with the effect that underlying long vowels surface long. One way to achieve this in the Optimality Theoretic analysis would be to assume, as we have done before, that underlying long vowels are distinguished by already being associated to two Xs or two μs in the input; and then to place superordinate in the Fula hierarchy a constraint which requires nuclear skeletal positions to be parsed. In this case, 'PARSE^{X}' in (318) would be specialized to apply to C skeletal positions ('PARSE^{C}') and another constraint PARSE^{V} for V positions would be inserted at the top of the hierarchy.

(which, after all, does occur at step 3)? This is where the Phonological Level Hierarchy is invoked. Because skeletal > segmental in this hierarchy, Paradis observes that

> a slot, which has *priority* over a segment, cannot be deleted because of a segmental restriction (viz. a segmental feature). Therefore the spreading of the continuant consonant seems to be the last resort. It causes a minimal violation, that is a violation of a segmental type, which can be minimally repaired in changing the value of the feature.
>
> p. 90, emphasis added.

Thus the issue comes down to a choice between two repairs: deleting a skeletal position or changing a feature. In the TCRS account, only one repair can apply first, and the choice is crucial — even though, in the end, both repairs will be made. The repair lower in the phonological level hierarchy (feature change) is made first.

As in the previous applications of Repair Analysis, the Optimality Theoretic view of the conflicts inherent in this situation is rather different. In the TCRS account, the 'no floating slot' constraint is met in the surface representation because '$\mathbb{X}$' has been deleted. Thus in this account the constraint is surface-unviolated. In the Optimality Theoretic analysis, however, the constraint PARSE^{X} is in fact violated in the output; only when this constraint is treated as violable can it truly conflict with the other constraints, and be violated as a result. The other victim of conflict, from the perspective of the Optimality Theoretic account, is the constraint PARSE^{feat}, which is therefore also surface-violated.

To summarize: from the TCRS perspective, it is the constraints 'no floating slot' and *GEMCONS which conflict, even though in the TCRS account both are surface unviolated. In our view, such conflict arises only when the former constraint is treated as violable, PARSE^{X}, and is in fact violated in output forms. Furthermore, the mechanism which TCRS invokes in cases regarded as constraint conflict is a mechanism of choosing which repair rule to apply first; it is therefore formally very different from the Optimality Theoretic mechanism of constraint domination for resolving conflicts between constraints which cannot be simultaneously satisfied in a possible output. The Optimality Theoretic conflicts crucially involve other surface-violated constraints such as FILL^{X} and PARSE^{feat} which are not part of the *constraint* component of TCRS, but rather correspond to *rules*, in much the same way as we have seen in the previous analyses of Phonotactics+Repair accounts.

The work of Paradis addresses a number of important and difficult issues which must be resolved in order to render well-defined those derivational theories in which various kinds of rules interact with constraints. We have argued that foregrounding constraints, their interactions, and their conflicts — giving due prominence to the crucial notion that linguistic constraints are violable — makes it possible to formulate phonological analyses which offer fresh substantive insights. The result is a strengthened theory of Universal Grammar, conceived as a set of violable constraints the interactions among which are determined on a language-particular basis. Among the principles of Universal Grammar are cognates of those formerly thought to be no more than loose typological and markedness generalizations. Formally sharpened, these principles now provide the very material from which grammars are built.

Appendix

⟨*incomplete*⟩

A.1 The Cancellation and Cancellation/Domination Lemmas

(216) **Cancellation Lemma.** Suppose two structures S_1 and S_2 both incur the same mark *m. Then to determine whether $S_1 \succ S_2$, we can omit *m from the list of marks of both S_1 and S_2 (`cancel the common mark') and compare S_1 and S_2 solely on the basis of the remaining marks. Applied iteratively, this means we can cancel *all* common marks and assess S_1 and S_2 by comparing only their unshared marks.

(192, 238) **Cancellation/Domination Lemma.** Suppose two parses B and C do not incur identical sets of marks. Then B ≻ C if and only if every mark incurred by B which is not cancelled by a mark of C is dominated by an uncancelled mark of C.

A.2 CV Syllable Structure

(136) **Onset Theorem** (Part). If ONS dominates either PARSE and FILL^{Ons}, onsets are required.

(138) **Coda Theorem** (Part). If −COD dominates either PARSE and FILL^{Nuc}, codas are forbidden.

A.3 Pāṉini's Theorem on Constraint-ranking

Lemma. Suppose a form f is accepted by a constraint hierarchy $\mathbb{CH}$ which includes C. Then either: f satisfies C, or C is not active on the input i

Proof. Consider the filtering of the candidate set at C; f must pass this filter since it passes all of $\mathbb{CH}$. Suppose f violates C. Then f will be filtered out by C unless C is violated by all the candidates which pass the filtering in $\mathbb{CH}$ prior to C, in which case C is not active on i. So if f violates C then C is not active on i. This is equivalent to the statement in the Lemma.

(112) **Pāṉini's Theorem on Constraint-ranking** (PTC). Let $\mathbb{S}$ and $\mathbb{G}$ stand as specific to general in a Pāṉinian constraint relation. Suppose these constraints are part of a constraint hierarchy $\mathbb{CH}$, and that $\mathbb{G}$ is active in $\mathbb{CH}$ on some input i. Then if $\mathbb{G} \gg \mathbb{S}$, $\mathbb{S}$ is not active on i.

Proof. Assume $\mathbb{S}$ so placed in the hierarchy; then any form which is subject to evaluation by $\mathbb{S}$ must have survived the filtering of the top portion of the hierarchy down to and including $\mathbb{G}$; call this subhierarchy T. Let i be an input for which $\mathbb{G}$ is active, which exists by hypothesis. Consider the candidate set Gen(i). By the lemma, any parse f in Gen(i) which survives T must satisfy $\mathbb{G}$. So every parse f which survives to be evaluated by $\mathbb{S}$ satisfies $\mathbb{G}$. But by the contrapositive of the Pāṉinian relationship, satisfying $\mathbb{G}$ entails failing $\mathbb{S}$. Since all the candidates which $\mathbb{S}$ gets to evaluate fail $\mathbb{S}$, $\mathbb{S}$ cannot filter any of them out, and $\mathbb{S}$ is therefore not active on the input i.

References

[Not corrected or updated from Tech Report version, except for one 1995 entry]

Allen, W. S. 1973. *Accent and rhythm*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1972. On Nasalization in Sundanese. *Linguistic Inquiry* 3, 253-268.

Anderson, Stephen R. 1969. *West Scandinavian Vowel Systems and the Ordering of Phonological Rules*. Doctoral dissertation. MIT.

Anderson, Stephen R. 1982. The analysis of French schwa, or how to get something for nothing. *Language* 58, 534-573.

Aoun, Joseph. 1979. Is the syllable or supersyllable a constituent? In *MITWPL 1: Papers on syllable structure, metrical structure, and Harmony processes*, ed. Ken Safir. MIT, Cambridge.

Aoun, Joseph. 1986. *Generalized binding: the syntax and logical form of wh-interrogatives*. *Studies in Generative Grammar 26*. Dordrecht: Foris.

Archangeli, Diana. 1984. *Underspecification in Yawelmani phonology and morphology*. Doctoral Dissertation. MIT, Cambridge, MA. [Garland Press, 1988.]

Archangeli, Diana and Douglas Pulleyblank. 1992. *Grounded phonology*. Ms. University of Arizona and University of British Columbia.

Archangeli, Diana and Douglas Pulleyblank. 1991. African tongue root Harmony: optimal systems. Paper presented at NSF/LSA *Conference on Phonological Feature Organization*. UCSC, Santa Cruz.

Archangeli, Diana and Douglas Pulleyblank. 1992. *Grounded Phonology*. Ms. University of Arizona and University of British Columbia.

Avery, P. and Keren Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6, 179-200.

Bach, Emmon and Deirdre Wheeler. 1981. Montague Phonology: A First Approximation. *University of Massachusetts Occasional Papers in Linguistics* 7. 27-45.

Barton, G. Edward, Robert C. Berwick, Eric Sven Ristad. 1987. *Computational complexity and natural language*. Cambridge, MA: MIT Press.

Bell, Alan. 1971. Some patterns of occurrence and formation of syllable structures. In *Working papers on language universals*, Number 6, 23-137. Language Universals Project, Stanford University.

Bell, Alan and Joan Bybee Hooper. 1978. *Syllables and Segments*. New York: Elsevier−North Holland.

Berg, G. 1991. Learning recursive phrase structure: combining the strengths of PDP and X-Bar Syntax. Technical Report 91−5, Dept. of Computer Science, State University of New York, Albany.

Berwick, Robert, Steven Abney, and Carol Tenny, eds. 1991. *Principle-based parsing*: *computation and psycholinguistics*. Dordrecht: Kluwer Academic Publishers.

Bird, Steven. 1990. *Constraint-based phonology*, Doctoral dissertation, University of Edinburgh.

Bittner, Maria. 1993. Cross-linguistic semantics. Ms. Rutgers University.

Black, H. Andrew. 1991. The optimal iambic foot and reduplication in Axininca Campa. *Phonology at Santa Cruz* 2, 1-18.

Black, Andrew. 1993. Constraint-ranked derivation: truncation and stem binarity in Southeastern Tepehuan. Ms. UC Santa Cruz.

Borowsky, Toni. 1986a. *Topics in the lexical phonology of english*, Doctoral dissertation, University of Massachusetts, Amherst.

Borowsky, Toni. 1986b. Structure preservation and the syllable coda in English. *Natural Language & Linguistic Theory* 7, 145-166.

Bosch, A. 1991. *Phonotactics at the level of the phonological word*. Doctoral Dissertation. University of Chicago.

Broselow, Ellen. 1982. On the interaction of stress and epenthesis, *Glossa* 16, 115-132.

Broselow, Ellen. 1992. Parametric variation in Arabic dialect phonology. In *Perspectives on Arabic Linguistics IV*, ed. E. Broselow, M. Eid and J. McCarthy. Amsterdam: Benjamins.

Broselow, Ellen and John McCarthy. 1983. A theory of internal reduplication. *Linguistic Review* 3, 25-88.

Brousse, O. 1991. *Generativity and Systematicity in Neural Network Combinatorial Learning*. Doctoral dissertion. Department of Computer Science, University of Colorado, Boulder.

Brousse, O. and Smolensky, P. 1989. Virtual memories and massive generalization in connectionist combinatorial learning. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, 380−387. Ann Arbor, MI. Lawrence Erlbaum.

Browning, M. A. 1991. Bounding conditions on representation. *Linguistic Inquiry* 22, 541-562.

Burzio, Luigi. 1987. English Stress. in P.-M. Bertinetto & M. Loporcaro, eds., *Certamen Phonologicum: Proceedings of the 1987 Cortona Phonology Meeting*.

Burzio, Luigi. 1992a. Metrical consistency. Abstract of paper presented at DIMACS Workshop, Princeton University.

Burzio, Luigi. 1992b. *Principles of English stress*, ms., Johns Hopkins University.

Bybee, Joan. 1985. *Morphology: a study of the relation between meaning and form*. Amsterdam: John Benjamins.

Cairns, Charles. 1969. Markedness, neutralization and universal redundancy rules. *Language* 45, 863-885.

Cairns, Charles. 1988. Phonotactics, markedness, and lexical representation. *Phonology* 5, 209-236.

Cairns, Charles and Mark Feinstein. 1982. Markedness and the theory of syllable structure. *Linguistic Inquiry* 13, 193-226.

Calabrese, Andrea. 1988. *Towards a theory of phonological alphabets*. Ph.D. diss., MIT, Cambridge, MA.

Carlson, Lauri. 1978. Word stress in Finnish. Ms. MIT, Cambridge, MA.

Chen, Matthew Y. 1987. The Syntax of Xiamen Tone Sandhi. *Phonology Yearbook* 4. 109-150.

Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. Master's Thesis. University of Pennsylvania, Philadelphia, PA.

Chomsky, Noam. 1992. A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics*. Cambridge, MA: MIT Department of Linguistics and Philosophy.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1989. Some notes on economy of derivation and representation. In *MIT Working Papers in Linguistics* 10, 43-47, ed. Itziar Laka and Anoop Mahajan. Cambridge, MA: MIT Department of Linguistics and Philosophy.

Chomsky, Noam. 1986. *Barriers*, MIT Press, Cambridge, MA.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Churchward, Clerk Maxwell. 1953. *Tongan grammar*. London: Oxford University Press.

Churchyard, Henry. 1991. Biblical Hebrew prosodic structure as the result of preference-ranked constraints. Ms. University of Texas, Austin.

Clements, G. N and S.J. Keyser. *CV Phonology*. Cambridge, MA: MIT Press.

Clements, G. N. 1976. Palatalization: linking or assimilation? *Proceedings of the Twelfth Regional Meeting of the Chicago Linguistic Society*, 96-109.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In *Papers in Laboratory Phonology I*, ed. John Kingston and Mary Beckman, 283-333. Cambridge University Press.

Clements, G. N. 1991. *Place of articulation in vowels and consonants: a unified theory*. In L'Architecture et la géométrie des représentations phonologiques, ed. B. Laks and A. Rialland, Paris: CNRS-Editions.

Cohen, M. A. and Grossberg, S. 1983. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:815–825.

Cole, J. S. 1992. Eliminating cyclicity as a source of complexity in phonology. UIUC-BI-CS Technical Report, Beckman Institute, University of Illinois. To be published in the proceedings of the Illinois conference on *Linguistics and Computation*.

Coleman, John. 1991. *Phonological representations — their Names, Forms, and Powers*. Doctoral Dissertation. University of York.

Davis, Stuart. 1988a. Syllable weight hierarchies and stress assignment. Ms. Indiana Univerity, Bloomington.

Davis, Stuart. 1988b. Syllable onsets as a factor in stress rules. *Phonology* 5, 1-19.

Dell, François. 1973. *Les règles et les sons*. Paris: Hermann.

Dell, François and Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7, 105-130.

Dell, François and Mohamed Elmedlaoui. 1988. Syllabic consonants in Berber: some new evidence. *Journal of African Languages and Linguistics* 10, 1-17.

Dell, François and Mohamed El-Medlaoui. 1989. Quantitative transfer in the nonconcatenative morphology of Imdlawn Tashlhiyt Berber. *Journal of Afroasiatic Languages*.

Devine, A. M. and Laurence Stephens. 1980. Latin prosody and meter: brevis brevians. *Classical Philology* 75, 142-157.

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

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

Dolan, C. P. 1989. *Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning, and Planning*. PhD thesis, Department of Computer Science, University of California, Los Angeles, CA.

Dolan, C. P. and Dyer, M. G. 1987. Symbolic schemata, role binding, and the evolution of structure in connectionist memories. In *Proceedings of the IEEE First International Conference on Neural Networks*, pages II:287–298, San Diego, CA.

Dressler, Wolfgang. 1985. On the predictiveness of natural morphology. *Journal Linguistics* 21, 321-337.

Elman, J. L. 1990. Finding structure in time. *Cognitive Science*, 14:179.

Elman, J. L. 1991. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7:195–226.

Elman, J. L. 1992. Grammatical structure and distributed representations. In Davis, S., editor, *Connectionism: Theory and Practice*. Oxford University Press, Oxford.

Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Doctoral Diss. MIT, Cambridge, MA.

Everett, Daniel. 1988. On metrical constituent structure in Pirahã. *Natural Language & Linguistic Theory* 6, 207-246.

French, Koleen Matsuda. 1988. *Insights into Tagalog reduplication. Infixation and stress from non-linear phonology*. MA Thesis. University of Texas, Arlington. SIL and University of Texas, Arlington, Publications in Linguistics No. 84.

Furby, C. 1974. *Garawa phonology*. Pacific Linguistics Series A, #37. Australian National University, Canberra.

Giegerich, Heinz. 1985. *Metrical phonology and phonological structure*. Cambridge: Cambridge University Press.

Gnanadesikan, Amalia. 1992. The feature geometry of coronal subplaces. Ms. UMass, Amherst.

Golden, R. M. 1986. The "Brain-State-in-a-Box" neural model is a gradient descent algorithm. *Mathematical Psychology*, 30–31:73–80.

Golden, R. M. 1988. A unified framework for connectionist systems. *Biological Cybernetics*, 59:109–120.

Goldsmith, John. 1990. *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.

Goldsmith, John. 1992. Local modeling in phonology. In Davis, S., editor, *Connectionism: Theory and Practice*. Oxford University Press, Oxford.

Goldsmith, John. 1993. Phonology as an intelligent system. In *Bridges between Psychology and Linguistics: A Swarthmore Festschrift for Lila Gleitman*, ed. Donna Jo Napoli and Judy Kegl, 247–267. ,. Hillsdale, NY: Lawrence Erlbaum Associates.

Goldsmith, John and Gary Larson. 1990. Local modeling and syllabification. In *Proceedings of the 26th Meeting of the Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology*, ed. K. Deaton, M. Noske, and M. Ziolkowski. Chicago, IL.

Goodman, Beverley. 1993. The integration of hierarchical features into a phonological system. Doctoral dissertation, Cornell University.

Goodman, Beverley. 1993. *The Integration of Hierarchical Features into a Phonological System*. Doctoral dissertation, Cornell University.

Green, Tom. 1992. Core syllabification and the grid: explaining quantity sensitivity. In *Proceedings of NELS 22*. [in press].

Grimshaw, Jane. 1993. Minimal projection, heads, and inversion. Ms. Rutgers University, New Brunswick, NJ.

Guerssel, M. 1985. Glides in Berber and syllabicity. *Linguistic Inquiry* 17, 1-12.

Haiman, John. 1972. Phonological targets and unmarked structures, *Language* 48, 365-377.

Hale, Kenneth. 1973. Deep-surface canonical disparities in relation to analysis and change: an Australian example, *Current Trends in Linguistics* 11, 401-458.

Hale, Kenneth and A. Lacayo Blanco. 1988. *Vocabulario preliminar del ULWA (Sumu meridional)*. Centro de investigaciones y documentación de la Costa Atlantica, Karawala, Zelaya Sur, Nicaragua and Center of Cognitive Science, MIT, Cambridge, MA.

Halle, Morris. 1973. Stress rules in English, a new version. *Linguistic Inquiry* 4, 451-464.

Halle, Morris and Jean-Roger Vergnaud. 1987. An essay on stress. Cambridge, MA: MIT Press.

Hammond, Michael. 1984. Constraining metrical theory: a modular theory of stress and destressing. Doctoral Dissertation. UCLA.

Hammond, Michael. 1992. Deriving ternarity. Ms. University of Arizona, Tucson.

Hare, M. 1990. The role of similarity in Hungarian vowel harmony: A connectionist account. *Connection Science*, 2:123–150.

Hayes, Bruce. 1980. A metrical theory of stress rules. Doctoral Dissertation. MIT, Cambridge, MA.

Hayes, Bruce. 1982. Extrametricality and English stress. *Linguistic Inquiry* 13, 227-276.

Hayes, Bruce. 1985. Iambic and trochaic rhythm in stress rules. In *Proceedings of the thirteenth meeting of the Berkeley Linguistics Association*, 429-446. University of California, Berkeley.

Hayes, Bruce. 1987. A revised parametric metrical theory. In *Proceedings of the Northeastern Linguistics Society 17*, GLSA, University of Massachusetts, Amherst.

Hayes, Bruce. 1991. *Metrical stress theory: principles and case studies*. Samizdat. Los Angeles. Published as Hayes 1995.

Hayes, Bruce. 1995. *Metrical stress theory: principles and case studies*. The University of Chicago Press. Chicago.

Hewitt, Mark. 1992. *Vertical maximization and metrical theory*. Doctoral dissertation. Brandeis University, Waltham.

Hewitt, Mark and Alan Prince. 1989. OCP, locality, and linking: the N. Karanga Verb. in *Proceedings of the Eighth WCCFL*, ed. E. Jane Fee and K. Hunt. Stanford University, Stanford, CA.

Hooper, Joan Bybee. 1972. The syllable in linguistic theory. *Language* 48, 525-540.

Hopfield, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, 79:2554–2558.

Hopfield, J. J. 1984. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA*, 81:3088–3092.

Howard, Irwin. 1973. A directional theory of rule application. Doctoral Dissertation, MIT. Bloomington: IULC.

Hudson, Grover. 1974. The role of SPC's in Natural Generative Phonology. In *Papers from the parasession on Natural Phonology*, ed. Anthony Bruck, Robert A. Fox, and Michael W. LaGaly, 171-183. Chicago: Chicago Linguistic Society.

Hulst, Harry van der. 1984. *Syllable structure and stress in Dutch*. Dordrecht: Foris.

Hulst, Harry van der. 1992. The independence of stress and rhythm. Ms. Rijksuniversiteit te Leiden.

Hume, Elizabeth V. 1992. *Front Vowels, Coronal Consonants, and Their Interaction in Nonlinear Phonology*. Doctoral dissertation. Cornell University, Ithaca.

Hung, Henrietta. 1992. Relativized suffixation in Choctaw: a constraint-based analysis of the verb grade system. Ms. Brandeis University, Waltham.

Hung, Henrietta. in preparation. *The Rhythmic and Prosodic Organization of Edge Constituents*. Doctoral dissertation. Brandeis University, Waltham.

Hyman, Larry. 1985. *A theory of phonological weight*. Dordrecht: Foris.

Inkelas, Sharon. 1989. *Prosodic Constituency in the Lexicon*. Ph.D. Dissertation. Stanford University.

Itô, Junko. 1986. *Syllable Theory in Prosodic Phonology*, Ph. D. Dissertation, UMass, Amherst.

Itô, Junko. 1989. A Prosodic Theory of Epenthesis, *Natural Language & Linguistic Theory* 7, 217-260.

Itô, Junko. Yoshihisa Kitagawa. and R. Armin Mester. 1992. Prosodic type preservation in Japanese: evidence from *zuuja-go*. SRC-92-05. Syntax Research Center. University of California. Santa Cruz.

Itô, Junko & R. Armin Mester. 1992. Weak layering and word binarity. Report LRC-92-09, Linguistics Research Center. Santa Cruz: University of California.

Itô, Junko and R. Armin Mester. to appear. Licensed segments and safe paths. In *Constraints, violations, and repairs in phonology*, ed. Carole Paradis, Darlene LaCharité. and Emmanuel Nikiema. *Special issue of the Canadian Journal of Linguistics*.

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1987. *Consciousness and the computational Mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray 1991. Musical parsing and musical affect. *Music Perception* 9, 199-230.

Jakobson, Roman. 1962. *Selected writings 1: phonological studies*. The Hague: Mouton.

Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. The Hague: Mouton.

Jain, A. N. 1991. Parsing complex sentences with structured connectionist networks. *Neural Computation*, 3:110−120.

Kager, René. 1989. *A Metrical Theory of Stress and Destressing in English and Dutch*. = *Een metrische theorie over klemtoon en klemtoonverlies in het Engels en het Nederlands*. Dordrecht: ICG.

Kager, René. 1992a. Are there any truly quantity-insensitive Systems?. To appear in *Proceedings of the Berkeley Linguistics Society*, 18.

Kager, René. 1992b. Shapes of the Generalized Trochee. To appear in *Proceedings of the Eleventh West Coast Conference on Formal Linguistics*.

Kager, René. 1992c. Alternatives to the Iambic-Trochaic Law. To appear in *Natural Language & Linguistic Theory*.

Kager, René & Ellis Visch. 1984a. *Een metrische analyse van ritmische klemtoonverschuiving*. MA thesis. Instituut de Vooys voo12r Nederlanse taal- en letterkunde. R. U. Utrecht.

Kager, René & Ellis Visch. 1984b. Syllable weight and Dutch word stress. in *Linguistics in the Netherlands* 1984, ed. Hans Bennis and WUS van Lessen Kloeke, 197-205. Dordrecht: Foris.

Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Doctoral Dissertation. MIT, Cambridge, MA.

Kaye, Jonathan and Jean Lowenstamm. 1981. Syllable structure and markedness theory. In *The theory of markedness in generative grammar*, ed. A. Belletti, L. Brandi, and L. Rizzi, 287-315. Scuola Normale Superiore di Pisa.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1985. The Internal Structure of Phonological Elements: A Theory of Charm and Government, *Phonology Yearbook* 2, 305-328.

Kaye, Jonathan and Jean Lowenstamm. 1984. De la syllabicité, in F. Dell, D. Hirst, and J.-R. Vergnaud, ed., *Forme sonore du langage*, Hermann, Paris.

Kaye, Jonathan. 1990. 'Coda' licensing. *Phonology* 7, 301-330.

Kean, Mary-Louise. 1974. The theory of markedness in Generative Grammar. Doctoral Dissertation. MIT.

Kelkar, Ashok R. 1968. *Studies in Hindi-Urdu I: introduction and word phonology*. Deccan College, Poona.

Kiparsky, Paul. 1973a. Abstractness, opacity, and global rules. Bloomington: Indiana University Linguistics Club.

Kiparsky, Paul. 1973b. Elsewhere in phonology. In *A Festschrift for Morris Halle*, ed. S. Anderson and P. Kiparsky, 93-106. New York: Holt, Rhinehart, and Winston.

Kiparsky, Paul. 1980. Vowel harmony, ms. MIT, Cambridge.

Kiparsky, Paul. 1982. Lexical Phonology and Morphology, in I.S. Yang ed., *Linguistics in the Morning Calm*. LSK, Hanshin, Seoul.

Kiparsky, Paul. 1991. Quantity sensitivity and the nature of templatic parsing. Undated handout. Berkeley Linguistics Society. Berkeley, CA.

Kiparsky, Paul. 1992. Catalexis. Ms. Stanford University, Stanford, CA.

Kirchner, Robert. 1992a. Lardil truncation and augmentation: a morphological account. Ms. University of Maryland, College Park.

Kirchner, Robert. 1992b. *Harmonic Phonology within One Language: An Analysis of Yidin^y*. MA thesis. University of Maryland. College Park.

Kirchner, Robert. 1992c. Yidin^y prosody in Harmony Theoretic Phonology. Ms. UCLA.

Kisseberth, Charles. 1970a. On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291-306.

Kissberth, Charles. 1970b. Vowel elision in Tonkawa and derivational constraints. In *Studies presented to Robert B. Lees by his students*, ed. J. Sadock and A. Vanek. Champain: Linguistic Research Inc.

Kisseberth, Charles. 1972. On derivative properties of phonological rules. In *Contributions to Generative Phonology*, ed. Michael K. Brame., 201-228. Austin, TX: University of Texas Press.

Klokeid, Terry Jack. 1976. *Topics in Lardil grammar*. Doctoral dissertation. MIT, Cambridge, MA.

Kuryłowicz, Jerzy. 1968. *Indogermanische Grammatik. Band II: Akzent. Ablaut.* Heidelberg: Carl Winter.

Lachter, J. and Bever, T. G. 1988. The relation between linguistic structure and associative theories of language learning—a constructive critique of some connectionist learning models. *Cognition*, 28:195–247.

Lakoff, George. 1965. On the nature of syntactic irregularity. Doctoral Dissertation. Harvard University. *Harvard Computational Lab Report NSF-16*. Cambridge, MA.

Lakoff, George. in press. Cognitive Phonology, in John Goldsmith, ed., *The Last Phonological Rule*, University of Chicago Press, Chicago.

LaPointe, Steven and Mark Feinstein. 1982. The Role of Vowel Deletion and Epenthesis in the Assignment of Syllable Structure, in *The Structure of Phonological Representations.Part II*, ed. H.v.d.Hulst and N. Smith, 69-120. Foris, Dordrecht, .

Larson, G. 1990. Local computational networks and the distribution of segments in the Spanish syllable. In *Proceedings of the 26th Meeting of the Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology*, ed. K. Deaton, M. Noske, and M. Ziolkowski. Chicago, IL.

Larson, G. 1992. *Dynamic computational networks and the representation of phonological information.* Doctoral Dissertation. Linguistics Department, University of Chicago.

Legendre, G., Miyata, Y., and Smolensky, P. 1990a. Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 388–395. Cambridge, MA: Lawrence Erlbaum.

Legendre, G., Miyata, Y., and Smolensky, P. 1990b. Harmonic Grammar—A formal multi-level connectionist theory of linguistic well-formedness: An application. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 884–891. Cambridge, MA: Lawrence Erlbaum.

Legendre, G., Miyata, Y., and Smolensky, P. 1990c. Can connectionism contribute to syntax? Harmonic Grammar, with an application. In *Proceedings of the 26th Meeting of the Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology*, ed. K. Deaton, M. Noske, and M. Ziolkowski. Chicago, IL.

Legendre, G., Miyata, Y., and Smolensky, P. 1991a. Unifying syntactic and semantic approaches to unaccusativity: A connectionist approach. In *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, ed. L. Sutton, and C. Johnson (with Ruth Shields). Berkeley, CA.

Legendre, G., Miyata, Y., and Smolensky, P. 1991b. Distributed recursive structure processing. In *Advances in Neural Information Processing Systems 3*, ed. D. S. Touretzky and R. Lippman, 591–597. San Mateo, CA: Morgan Kaufmann.

Legendre, Géraldine, William Raymond, and Paul Smolensky. 1993. Analytic typology of case marking and grammatical voice. To appear in Proceedings of the *Berkeley Linguistics Society*, 19.

Lerdahl, Fred and Ray Jackendoff 1983. *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levelt, Clara. 1991. Samoan reduplication: how to behave as a plural and other good manners. *Proceedings of the Leiden Conference for Junior Linguists 2*, ed. J. van Lit, R. Mulder and R. Sybesma. Leiden University: The Netherlands.

Levin [Blevins], Juliette. 1985. *A Metrical Theory of Syllabicity*. Doctoral Dissertation. MIT, Cambridge.

Liberman, Mark. 1975. *The intonational system of English*. Doctoral Dissertation. MIT, Cambridge.

Liberman, Mark and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249-336.

Levin [Blevins], Juliette. 1985. *A Metrical Theory of Syllabicity*. Ph.D. Dissertation. MIT. Cambridge.

Lombardi, Linda and John McCarthy. 1991. Prosodic circumscription in Choctaw morphology. *Phonology* 8, 37-71.

Lowenstamm, Jean and Jonathan Kaye. 1986. Compensatory lengthening in Tiberian Hebrew, in *Studies in Compensatory Lengthening*, ed. L. Wetzels and E. Sezer. Foris: Dordrecht.

McCarthy, John. 1979. *Formal Problems in Semitic Phonology and Morphology*, Doctoral dissertation, MIT, Cambridge, MA.

McCarthy, John. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17, 207-263.

McCarthy, John. To appear. A case of surface constraint violation. *Canadian Journal of Linguistics*. special issued edited by Carole Paradis. Darlene LaCharité. and Emmanuel Nikiema.

McCarthy, John and Alan Prince. 1986. *Prosodic Morphology*. Samizdat, Amherst and Waltham, MA.

McCarthy, John and Alan Prince. 1988. Quantitative Transfer in Reduplicative and Templatic Morphology. in Linguistic Society of Korea. ed. *Linguistics in the Morning Calm* 2. Hanshin Publishing Co. Seoul. Pp. 3-35.

McCarthy, John and Alan Prince. 1990a. Foot and Word in Prosodic Morphology: The Arabic Broken Plurals. *Natural Language and Linguistic Theory* 8. 209-282.

McCarthy, John and Alan Prince. 1990b. Prosodic Morphology and Templatic Morphology. In *Perspectives on Arabic Linguistics: Papers from the Second Symposium*, ed. M. Eid and J. McCarthy, 1-54. Amsterdam: Benjamins.

McCarthy, John and Alan Prince. 1991a. Prosodic Minimality. lecture presented at University of Illinois Conference *The Organization of Phonology*.

McCarthy, John and Alan Prince. 1991b. *Linguistics 240: Prosodic Morphology*. lectures and handouts from 1991 LSA Linguistic Institute Course. University of California. Santa Cruz.

McCarthy, John and Alan Prince. 1993. *Prosodic Morphology I: constraint interaction and satisfaction*. Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McCarthy, John and Alison Taub. 1992. Review of Carole Paradis and Jean-François Prunet, *The Special Status of Coronals*. *Phonology* [in press].

McClelland, J. L. and Kawamoto, A. H. 1986. Mechanisms of sentence processing: Assigning roles to constituents. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*, ed. D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, ch. 19, 272–325. Cambridge, MA: MIT Press/Bradford Books.

Marcus, Gary, Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, & Fei Xu. 1992. *Overregularization in language acquisition*. *Monographs of the Society for Research in Child Development*, 57 (4, Serial No. 228). Also MIT CCS Occasional Paper #41.

Mester, R. Armin. 1991. Some remarks on Tongan stress. Ms. UCSC, Santa Cruz, CA.

Mester, Ralf-Armin. 1992. The quantitative trochee in Latin. SRC-92-06, Syntax Research Center, UC Santa Cruz. To appear in *Natural Language and Linguistic Theory, 1993.*

Michelson, Karin. 1988. *A comparative study of Lake Iroquoian accent*. Dordrecht: Kluwer.

Miikkulainen, R. and Dyer, M. G. 1991. Natural language processing with modular PDP networks and distributed lexicon. *Cognitive Science*, 343–399.

Mohanan, K. P. 1991. On the bases of radical underspecification. *Natural Language & Linguistic Theory* 9.2, 285-325.
Mohanan, K. P. in press. Fields of Attraction in Phonology, in John Goldsmith, ed., *The Last Phonological Rule* [sic], University of Chicago Press, Chicago.
Myers, Scott. 1987a. *Tone and the structure of words in Shona*. Doctoral dissertation. University of Massachusetts, Amherst.
Myers, Scott. 1987b. Vowel shortening in English. *Natural Language & Linguistic Theory* 5, 485-518.
Myers, Scott. 1991. Persistent rules. *Linguistic Inquiry* 22, 315-344.
Nespor, Marina & Irene Vogel. 1986. Prosodic phonology. Dordrecht: Foris.
Noske, Roland. 1982. Syllabification and Syllable-changing Rules in French, in Harry van der Hulst and Norval Smith, ed., *The Structure of Phonological Representations. Part II.*, Foris, Dordrecht.
Oehrle, Richard. 1971. Some remarks on the role of morphology in the assignment of stress. Ms. MIT.
Paradis, Carole. 1988a. On Constraints and Repair Strategies, *The Linguistic Review* 6, 71-97.
Paradis, Carole. 1988b. Towards a Theory of Constraint Violations, *McGill Working Papers in Linguistics* 5, 1-44.
Paradis, Carole and François Prunet, eds. 1991. *The special status of coronals: internal and external evidence*. As *Phonetics and phonology*, Volume 2. New York: Academic Press.
Payne, David. 1981. *The Phonology and Morphology of Axininca Campa*. Summer Institute of Linguistics. Arlington, TX.
Payne, David. Judith Payne. and Jorge Santos. 1982. *Morfologia. Fonologia. y Fonetica del Asheninca del Apurucayali. Campa—Arawak Preandino.* Ministry of Education. Summer Institute of Linguistics. Yarinacocha. Peru.
Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. NY: Holt, Rhinehart, & Winston.
Piggott, Glyne and Raj Singh. 1985. The phonology of epenthetic segments. *Canadian Journal of Linguistics* 30, 415-453.
Piggott, Glyne. 1992. Satisfying the minimal word. Ms. McGill University, Montréal.
Pinker, Steven. 1984. *Language learnability and language development*. Cambridge: Harvard Univ. Press.
Pinker, Steven and Alan Prince. 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28:73–193.
Pollack, J. B. 1988. Recursive auto-associative memory: Devising compositional distributed representations. In *Proceedings of the Tenth Annual Meeting of the Cognitive Science Society*. Montreal, Canada: Erlbaum Associates.
Pollack, J. B. 1990. Recursive distributed representation. *Artificial Intelligence*, 46:77–105.
Poser, William J. 1985. Cliticization to NP and Lexical Phonology. WCCFL 4, 262-272.
Poser, William J. 1986. Invisibility. *GLOW Newsletter* 16, 63-64.
Prince, Alan. 1976. 'Applying' stress. Ms. University of Massachusetts, Amherst.
Prince, Alan. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11, 511-562.
Prince, Alan. 1983. Relating to the Grid. *Linguistic Inquiry* 14. 19-100.

Prince, Alan. 1986. Improving tree theory. In *Proceedings of the Eleventh annual meeting of the Berkeley Linguistics Society*, ed. Mary Niopokuj, Mary VanClay, Vassiliki Nikiforidou, Deborah Feder, 471-490. University of California, Berkeley.

Prince, Alan. 1990. Quantitative Consequences of Rhythmic Organization. In *CLS 26-II: Papers from the Parasession on the Syllable in Phonetics and Phonology*, ed Karen Deaton, Manuela Noske, and Michael Ziolkowski. Chicago Linguistic Society, Chicago.

Prince, Alan. 1993. In defense of the number *i*: anatomy of a linear dynamical model of linguistic generalizations. Technical Report RuCCS TR-1, Rutgers Center for Cognitive Science.

Prince, Alan and Paul Smolensky. 1991a. Optimality. Talk given at Arizona Phonology Conference. University of Arizona, Tucson

Prince, Alan and Paul Smolensky. 1991b. Notes on connectionism and Harmony Theory in linguistics. Technical Report CU-CS-533-91. Department of Computer Science, Univ. of Colorado, Boulder.

Prince, Alan and Paul Smolensky. 1992. Optimality: constraint interaction in generative grammar. Revised and augmented handout from talk delivered at WCCFL.

Pulleyblank, Douglas. 1983. Tone in Lexical Phonology. Dordrecht: D. Reidel.

Pyle, Charles. 1972. On Eliminating BM's, *CLS* 8.

Rizzi, Luigi. 1990. *Relativized Minimality*. Linguistic Inquiry Monograph 16. Cambridge: MIT press.

Rager, J. and G. Berg 1990. A connectionist model of motion and government in Chomsky's government-binding theory. *Connection Science*, 2:35−52.

Rosenthall, Sam. in preparation. *The phonology of vowels and glides*. Doctoral dissertation. University of Massachusetts, Amherst.

Rosenthall, Sam. in prep. *The Phonology of Vowels and Glides*. Doctoral dissertation. University of Massachusetts, Amherst.

Rubach, Jerzy and Gert Booij. 1985. A grid theory of stress in Polish. *Lingua* 66, 281-319.

Rumelhart, D. E. and J. L. McClelland 1986. On learning the past tenses of English verbs. ch. 18, 216−271. Cambridge, MA: MIT Press/Bradford Books.

Rumelhart, D. E., P. Smolensky, J. L. McClelland, and G. E. Hinton, 1986. Schemata and sequential thought in PDP models. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*, ed. J. L. McClelland, D. E. Rumelhart, and the PDP Research Group, ch. 14, pages 7−57. Cambridge, MA: MIT Press/Bradford Books.

Samek-Lodovici, Vieri. 1992a. Universal constraints and morphological gemination: a crosslinguistic study. Ms. Brandeis University.

Samek-Lodovici, Vieri. 1992b. A unified analysis of crosslinguistic morphological gemination. In *Proceedings of CONSOLE-1.*

Sapir, Edward. 1930. *The Southern Paiute language*. *Proceedings of the American Academy of Arts and Sciences* 65, 1-296.

Schachter, Paul. 1987. Tagalog. In *The world's major languages*, ed. Bernard Comrie, ch. 47, p. 936-958. Croom Helm: London and Sidney.

Schachter, Paul and Fe Otanes. 1972. *Tagalog reference grammar*. Berkeley, CA: Univ. of California Press.

Scobbie, James. 1991. *Attribute Value Phonology*. Doctoral dissertation, University of Edinburgh.

Scobbie, James. 1992. Towards Declarative Phonology, *Edinburgh Working Papers in Cognitive Science* 7, 1-26.

Selkirk, Elisabeth. 1980a. Prosodic domains in phonology: Sanskrit revisited. In *Juncture*, ed. Mark Aronoff and Mary-Louise Kean, 107-129. Anma Libri, Saratoga, CA.

Selkirk, Elisabeth. 1980b. The role of prosodic categories in English word stress, *LI* 11, 563-605.

Selkirk, Elisabeth. 1981. Epenthesis and degenerate syllables in Cairene Arabic. In *Theoretical Issues in the Grammar of the Semitic Languages*, ed. H. Borer and J. Aoun. MIT, Cambridge.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge: MIT Press.

Selkirk, Elisabeth. 1986. On Derived Domains in Sentence Phonology. *Phonology Yearbook* 3. 371-405.

Selkirk. Elisabeth. 1993. The prosodic structure of functional elements: affixes, clitics, and words. Handout of talk presented at the *Signal to Syntax* Conference. Brown University, Providence.

Sherer, Tim. in prep. *Prosodic Phonotactics*. Doctoral dissertation. Univ. of Massachusetts, Amherst.

Singh, Raj. 1987. Well-formedness Conditions and Phonological Theory, in Wolfgang Dressler et al., ed., *Phonologica 1984*, Cambridge University Press, Cambridge.

Sommerstein, A. H. 1974. On Phonotactically Motivated Rules, *Journal of Linguistics* 10, 71-94.

Smolensky, P. 1983. Schema selection and stochastic inference in modular environments. In *Proceedings of the National Conference on Artificial Intelligence*, 378–382, Washington, DC.

Smolensky, P. 1984a. Harmony Theory: thermal parallel models in a computational context. Technical report 8404, Institute for Cognitive Science, Univ. of California at San Diego. In P. Smolensky & M. Riley, *Harmony theory: Problem solving, parallel cognitive models, and thermal physics*.

Smolensky, P. 1984b. The mathematical role of self-consistency in parallel computation. In *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, 319–325. Boulder: Erlbaum.

Smolensky, P. 1986. Information processing in dynamical systems: Foundations of Harmony Theory. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, ed. D. E.Rumelhart, J. L. McClelland, and the PDP Research Group, ch. 6, pages 194–281. Cambridge, MA: MIT Press/Bradford Books,

Smolensky, P. 1987. On variable binding and the representation of symbolic structures in connectionist systems. Technical Report CU-CS-355-87. Department of Computer Science, University of Colorado at Boulder.

Smolensky, P. 1988. On the proper treatment of connectionism. *The Behavioral and Brain Sciences*, 11:1–74.

Smolensky, P. 1990. Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence*, 46:159–216.

Smolensky, P. In press. Constituent structure and explanation in an integrated connectionist/symbolic cognitive architecture. In *The Philosophy of Psychology: Debates on Psychological Explanation*, ed. C. Macdonald and G. Macdonald. Oxford: Basil Blackwell.

Smolensky, P., G. Legendre, and Y. Miyata. 1992. Principles for an integrated connectionist/symbolic theory of higher cognition. Technical Report CU-CS-600-92. Department of Computer Science, University of Colorado at Boulder.

Spring, Cari. 1990. *Implications of Axininca Campa for Prosodic Morphology and Reduplication*. Doctoral dissertation. University of Arizona. Tucson.

St. John, M. and J. L. McClelland. 1990. Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46:217–258.

Stampe, David. 1969. The acquisition of phonetic representation. In *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*.

Stampe, David. 1973. On chapter nine. In *Issues in phonological theory*, ed. Michael Kenstowicz and Charles Kisseberth, 44-52. The Hague: Mouton.

Stampe, David. 1973/9. *A Dissertation in Natural Phonology*, Doctoral Dissertation. NY: Garland Publ. Co.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral Dissertation. MIT, Cambridge, MA.

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

Stevens, Kenneth and S. J. Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65, 81-106.

Stevens, Kenneth, S. J. Keyser, and H. Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. In *Invariance and variability in speech processes*, ed. J. Perkell and D. Klatt. Hillsdale, NJ: Erlbaum.

Stolcke, A. 1989. Unification as constraint satisfaction in structured connectionist networks. *Neural Computation*, 1:559–567.

Stump, Gregory. 1989. A note on Breton pluralization and the Elsewhere Condition. *Natural Language & Linguistic Theory* 7, 261-273.

Touretzky, D. S. 1989. Towards a connectionist phonology: The "many maps" approach to sequence manipulation. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, 188–195, Ann Arbor, MI: Lawrence Erlbaum.

Touretzky, D. S., editor. 1991. Special issue: Connectionist approaches to language learning. *Machine Learning*, 7:105–252.

Touretzky, D. S. and Wheeler, D. W. 1991. Sequence manipulation using parallel mapping networks. *Neural Computation*, 3:98–109.

Vennemann, Theo. 1972. On the theory of syllabic phonology. *Linguistische Berichte* 18, 1-18.

Visch, Ellis. 1989. A metrical theory of rhythmic stress phenomena. (*Een metrische theorie van ritmische klemtoonverschijnselen*). Dordrecht: ICG.

Voltaire. 1759. *Candide, ou l'optimisme*. Revised 1761 [for the worse, in our view]. tr. Lowell Bair as *Candide by Voltaire*, 1959. New York: Bantam Books.

Wertheimer, Max. 1923. Laws of organization in perceptual forms. In A source book of Gestalt psychology, ed. W. D. Ellis. London: Routledge & Kegan Paul.

Wheeler, Deirdre. 1981. *Aspects of a Categorial Theory of Phonology*, Doctoral dissertation, University of Massachusetts, Amherst.

Wheeler, Deirdre. 1988. Consequences of some categorially motivated phonological assumptions. In *Categorial grammars and natural language structures*, ed. R. Oehrle, E. Bach, and D. Wheeler. Dordrecht: Reidel.

Wheeler, D. W. and Touretzky, D. S. 1993. A connectionist implementation of cognitive phonology. In *The Last Phonological Rule* [sic], ed. John Goldsmith. Chicago, IL: Univ. of Chicago Press.

Wilkinson, Karina. 1986. Syllable structure and Lardil phonology. Ms. Univ. of Massachusetts, Amherst.

Wilkinson, Karina. 1988. Prosodic structure and Lardil phonology. *Linguistic Inquiry* 19, 325-334.

Wiltshire, Caroline. 1992. *Syllabification and Rule Application in Harmonic Phonology*. Doctoral Dissertation. University of Chicago.

Woodbury, Anthony. 1981. *Study of the Chevak dialect of Central Yupik Eskimo*. University of California, Berkeley.

Wurzel, Wolfgang. 1985. On morphological naturalness. *Nordic Journal of Linguistics* 7, 145-163.

Yip, Moira. 1988. The obligatory contour principle and phonological rules: a loss of identity. *Linguistic Inquiry* 19, 65-100.

Yip, Moira. 1992. Cantonese loan word phonology and optimality. To appear in *Journal of East Asian Linguistics*.

Yip, Moira. 1993. Phonological constraints, optimality, and phonetic realization in Cantonese. UPenn Colloquium.

Zec, Draga. 1988. *Sonority constraints on syllable structure*. Doctoral Dissertation. Stanford University: Stanford, CA.

Zec, Draga. in prep. Coda constraints and conditions on syllable weight.

Zoll, Cheryl. 1992a. When syllables collide: a theory of alternating quantity. Ms. Brandeis University.

Zoll, Cheryl 1992b. pc. Remarks on the OCP. Internet.

Zoll, Cheryl. 1993. Ghost consonants and optimality. WCCFL, Santa Cruz.

Zonneveld, Wim. 1976. Destressing in Halle's English stress rules. *Linguistic Inquiry* 7, 520-525.