Constraint Conjunction versus Grounded Constraint Subhierarchies in Optimality Theory*

Jaye Padgett - University of California, Santa Cruz (Version July 29, 2001)

1. Introduction

A central assumption of classical Optimality Theory (Prince and Smolensky 1993) is that constraint domination is strict. That is, if constraint A outranks constraints B and C (A >> B, C), then a candidate output that violates A can never be favored over one that satisfies A but violates B and/or C, no matter how many violations of B and C occur. Tableau (1) illustrates the familiar scenario of strict domination: though candidate (1)a has more constraint violations than (1)b, it wins.

(1)

	A	B	C
a. ☞ *cand(a)*		*	*
b. *cand(b)*	*!		

Informally, one might say that the number of times a candidate violates a constraint, or set of constraints, has no bearing on constraint ranking. One could imagine a different world. For instance, suppose A >> B, C *unless both of the latter are violated*, shown in (2). In such a world, both (2)a and (2)b would win over (2)c (the choice between them ultimately depending on the ranking of B and C). However, a candidate (2)d that violates *both* B and C would be *less* harmonic than (2)c. It is as though the two violations together cause these constraints to be 'promoted' over A. In claiming that constraint violations do not add up in this way, Prince and Smolensky set Optimality Theory apart from the connectionist work that formed part of its inspiration.

(2)

	B&C	A	B	C
a. ☞ *cand(a)*			*	
b. ☞ *cand(b)*				*
c. *cand(c)*		*!		
d. *cand(d)*	*!		*	*

But as is well known, Smolensky (1993, 1995, 1997) hypothesizes that just such an interaction of constraints as in (2) exists, known as *local constraint conjunction*.[1] Because it explicitly relinquishes the assumption of strict domination, even if only under some conditions (see below), constraint conjunction represents an important departure from the more restrictive original conception of constraint ranking. Indeed, constraint conjunction faces an important challenge of potential overgeneration. Though there have been several proposals about how to address this challenge, no consensus has emerged. And yet the 'worst of the worst' interaction that conjunction is designed to accommodate, illustrated above, has both an intuitive naturalness and certain empirical support. In this paper, I examine several significant uses of constraint conjunction that have been proposed, and suggest that they all can and should be subsumed under a notion independently required, that of the *universal constraint subhierarchy*. If this is correct, then constraint conjunction is not required of the theory. In addition, following Prince and Smolensky (1993), I take universal subhierarchies to be derived from linguistically relevant scales. Assuming these are phonetically or psycholinguistically grounded, then we have a promising means by which to address the challenge of overgeneration. The intuition guiding all of this paper is that apparent 'worst of the worst' effects are really 'too much of a *single* bad thing'.[2]

2. Local Constraint Conjunction

Smolensky (1995) formulates local constraint conjunction as in (3).

(3) *The Local Conjunction of C_1 and C_2 in domain D*:

a. C_1&C_2 is violated when there is some D in which both C_1 and C_2 are violated.
b. Universally C_1&C_2 >> C_1,C_2

The idea behind constraint conjunction is that two constraints, when violated within some sufficiently local context, add up to more than the sum of their parts, in the specific sense illustrated in (2): the combined effects amount to a separately rankable, dominant constraint. Let us call this a *WOW* effect, for 'worst of the worst'. As an example of a *WOW* interaction, Smolensky suggests neutralization to coronal place in the coda position. This can be seen as the markedness of place of articulation, and of coda position generally, 'ganging up' to rule out all but coronal place in the coda. Assume the familiar universal place markedness hierarchy shown in (4)a (ignoring any possible ranking between *Lab and *Dor). Each of these constraints can be conjoined with NoCoda, giving (4)b. By assumption (3)b, these conjoined constraints outrank their simple constraint components. Itô and Mester (to appear-a) pursue an analogous strategy to account for a range of coda processes affecting German, including syllable-final devoicing, which is handled through the conjunction NoCoda&*[-son, +voice].

(4) a. *Lab, *Dor >> *Cor
b. NoCoda&*Lab >> NoCoda, *Lab *by constraint conjunction*
NoCoda&*Dor >> NoCoda, *Dor
NoCoda&*Cor >> NoCoda, *Cor

If ranked above Ident(Place), as shown below, the constraint NoCoda&*Lab rules out forms such as [map] in favor of [mat]. The form [mat] satisfies the conjoined constraint, it should be noted, assuming that the *domain* of this constraint is a single segment. (See below.) By analogous means, /mak/ would neutralize to [mat] also. Forms like [nat] gratuitously violate Ident(Place), since the conjoined constraint targets only codas.

(5)

Input: /map/	NoCoda&*Lab	Ident(Place)	NoCoda	*Lab	*Cor
a. map	*!		*	**	
b. ☞ mat		*	*	*	*
c. nat		**!	*		**

In order to adequately explain coda neutralization effects, this *WOW* approach actually requires another assumption about rankings besides that made in (3)b. It implies that the universal ranking of the markedness hierarchy *Lab, *Dor >> *Cor continues to exert an effect in the derived constraints. In particular, not only must NoCoda&*Lab dominate its component constraints, but the ranking shown in (6)a must obtain as well. More generally, this use of constraint conjunction and simplex universal hierarchies could be taken to imply the ranking assumption in (6)b, explicitly proposed by Spaelti (1997), called the Universal Conjoined Constraint Ranking Hypothesis. This assumption is also adopted as it applies to self-conjoined constraints by Alderete (1997) and Itô and Mester (to appear-b). Gafos and Lombardi (1999) extend the idea in a particular way to the conjunction of two subhierarchies.

(6) a. NoCoda&*Lab, NoCoda&*Dor >> NoCoda&*Cor
b. If C_n >> ... >> C_2 >> C_1 holds universally, then so does C_n&X >> ... >> C_2&X >> C_1&X, where X is any constraint.

Were (6)a not to hold, we could not explain why coda neutralization is typically to the unmarked. For example, a language neutralizing codas to labial place would be predicted, as shown below. (Lower ranked constraints are omitted.)

(7)

Input: /mat/	NoCoda&*Cor	Ident(Place)	NoCoda&*Lab
a. ☞ map		*	*
b. mat	*!		

Besides the general issue of whether the power of constraint conjunction is necessary to the theory, conjunction raises two pressing questions (Smolensky 1997): what constraints can be conjoined, and what determines the domain of a conjoined constraint? Regarding the first question, Kirchner (1996) offers as an example of an undesirable conjunction one involving a constraint against complex onsets and another banning heavy syllables, predicting a language in which only light syllables can have complex onsets. Examples such as this are not hard to construct. Though there is wide agreement on the importance of the question, there has been remarkably little consensus regarding the answer. For example, Kirchner speculates that constraints can be conjoined only with themselves or with closely related constraints. Fukazawa and Miglio (1998) argue for a less demanding restriction, that conjunction be limited to constraints from the same 'family'. Itô and Mester (to appear-b) argue for more power than this, prohibiting only the conjunction of markedness constraints with faithfulness constraints.[3] Taking a different approach, Łubowicz (1998) claims that constraints can be conjoined only if their formulations share an argument. (Compare Hewitt and Crowhurst 1996, Crowhurst and Hewitt 1997, who employ a different notion of constraint conjunction.) Baković's (1999) notion of 'co-relevance' is similar, but makes different predictions. What all of the proposals just mentioned share is the intuition that conjoined constraints must have some property in common. Unfortunately, it remains entirely unclear what that property is.[4]

The second question, concerning possible domains, is related to the first, since limits on domains might affect which constraints could conjoin and vice versa. Itô and Mester (to appear-b) provide an illustration of the problem: were the domain of the conjoined constraint NoCoda&*[-son, +voi] to be as large as a syllable, then it would rule out forms such as [bat] having a voiced obstruent in the onset and a voiceless coda. There have been fewer proposals to limit the domain of conjunction in any general way, and there is again no agreement on an approach. Łubowicz (1998, 2002) suggests that the domain of a conjoined constraint must be the minimal one possible given the relevant constraints. Nathan (2001) offers what might be considered a formalization of this idea. But this restriction rules out certain prominent uses of constraint conjunction, such as the use of self conjunction of markedness to derive dissimilation, as in Itô and Mester (1996, to appear-b) and Alderete (1997). For example, Itô and Mester argue that while the domain of voiced obstruent dissimilation in Modern Japanese is the stem, in Old Japanese it was the word. Earlier approaches employing the Obligatory Contour Principle also stipulated the domains of dissimilatory constraints, e.g. Yip (1988). Itô and Mester (to appear-b) argue that conjoined constraints must in principle have access to the same range of domains as ordinary constraints, encompassing prosodic and morphological constituents.

To summarize, the worry of overgeneration by conjunction is well understood, but little convincing progress has been made to address it. It is possible that compelling solutions will be found. But the continued elusiveness of answers suggests we consider another possibility, that the problem lies with constraint conjunction. The problem, I suggest, is that the *combinatoric* mechanism of conjunction, by its very nature, does not adequately seek out any *simple*, *unitary* reality underlying apparent *WOW* effects. What underlies at least some of them, I will argue, are universal constraint subhierarchies that are grounded in phonetic and psycholinguistic scales. Scales involve degrees of some simple, unitary notion, and the subhierarchies projected from them therefore penalize degrees of some simple, unitary marked configuration. As we will see, this is what fleshes out the 'too much of a bad thing' intuition.[5]

3. Grounded universal constraint subhierarchies

Since their introduction by Prince and Smolensky (1993), universal constraint subhierarchies have played a central role within Optimality Theory in explaining implicational universals. A defining property of these subhierarchies are the fixed universal rankings. Prince and Smolensky distinguish two kinds, illustrated in (8).

(8) a. Markedness subhierarchies: *Lab, *Dor >> *Cor
b. Subhierarchies from Prominence Alignment: *Nuc/t >> ... >> *Nuc/ɛ >> *Nuc/a

Prince and Smolensky posit markedness hierarchies directly. The other category, (8)b, are derived by means of an operation called Prominence Alignment, which is rooted in the notion of two scales. In the present example the two scales in question are shown in (9)a. These scales are 'aligned' to derive the harmony rankings shown in (9)b. The intuition here is that the most prominent syllable position (nucleus) is best associated with the most prominent kind of sound (most sonorous), and so on. Finally, these harmony rankings project to the constraint rankings shown in (9)c.

(9) Prominence Alignment: Example:

a.	Given two scales	a > ɛ > ... > t Peak > Margin	*Sonority hierarchy* *Syllable prominence hierarchy*
b.	Posit harmony rankings	P/a ≻ P/ɛ ≻ ... ≻ P/t M/t ≻ ... ≻ M/ɛ ≻ M/a	
c.	And constraint rankings	*P/t >> ... >> *P/ɛ >> *P/a *M/a >> *M/ɛ >> ... >> *M/t	

Other work has emphasized that so-called markedness scales such as (8)a are themselves rooted in phonetic scales (e.g., Boersma 1998 and Sanders in progress.). For example, building on Pulleyblank (1989), Cohn (1993), and other work, Walker (1998) infers a universal constraint subhierarchy (10) based on the principle that nasality is aerodynamically incompatible with increasing stricture.[6]

(10) *Nas/obst-stop >> *Nas/fric >> *Nas/liq >> *Nas/glide >> *Nas/vowel

In a formalism pre-dating Optimality Theory, Archangeli and Pulleyblank (1994) similarly posit implicational relationships among feature co-occurrence constraints involving [ATR] and tongue height and backness, inferred from scales of articulatory and perceptual difficulty.

I assume, in fact, that all universal constraint subhierarchies are grounded in scales having some phonetic or psycholinguistic relevance, and that there is no reason to distinguish cases such as (8)a and (8)b. In addition, anticipating other examples to come, it seems likely that Prominence Alignment is too specific a notion for projecting subhierarchies. The simplest cases have more the character of (11), and it is conceivable that most or all cases could be reduced to this. (To simplify, I skip an intermediate 'harmony alignment' step.)[7]

(11) Projection of Universal Constraint Subhierarchies

a. Given a scale of articulatory/perceptual/processing difficulty $\mathcal{D}$: $D_n > D_{n-1} > ... > D_1$
(where '>' means 'more difficult than')

b. Project a universal constraint subhierarchy: $C_n >> C_{n-1} >> ... >> C_1$
(where $C_i = {*}D_i$)

For example, the universal subhierarchy of (10) is based upon a scale of aerodynamic/articulatory difficulty: [d̃] > [z̃] > [l̃] > [j̃] > [ĩ]. The MinDist constraints of Flemming (1995) and Space constraints of Padgett (to appear) require a certain distance between contrasting sounds along some perceptual dimension, where decreasing distance corresponds to increasing difficulty in distinguishing contrasting sounds, implying hierarchies like Space1/3 >> Space1/2 >> Space of Padgett (to appear). Notions like 'articulatory difficulty' and 'perceptual distance' are inherently scalar. Some researchers, including those just cited and Steriade (1997), Kirchner (2000), Boersma (1998), Sanders (to appear), among others, hypothesize that many or all constraints are grounded in 'difficulty' along articulatory, perceptual, or processing dimensions. If this is true, then perhaps the great majority of constraints in fact occupy universal subhierarchies, and the latter are even more prevalent in constraint hierarchies than is typically assumed.

The relevance of this to the question of constraint conjunction can now be considered. Conjunction is posited in a situation where some constraint C_x demonstrably outranks two other constraints C_1 and C_2, yet violation of *both* C_1 and C_2 trumps C_x. This is depicted in (12)a. Compare this to (12)b, a scenario in which C_x dominates C_{i-1} of some universal constraint subhierarchy, while C_i of the same subhierarchy dominates C_x. These two scenarios have more in common than this rather trivial ranking similarity. In both cases, the relation between the dominant constraints (C_1&C_2 and C_i respectively) and their subordinate counterparts (C_1 and C_2 on the one hand, C_{i-1} on the other) represents in some sense 'more of the same'. This raises the question whether both notions are independently necessary.

(12) Constraint conjunction and universal constraint subhierarchies

a. $C_1\&C_2$ >> C_x >> C_1, C_2
b. C_i >> C_x >> C_{i-1}

The similarity between constraint conjunction and constraint subhierarchies becomes most obvious when we consider the special case of *self*-conjunction. Smolensky (1995) points out that the operation of conjunction might be recursively applied to a single constraint, creating what he calls a 'power hierarchy', as shown in (13).

(13) Power hierarchy

C^n >> C^{n-1} >> ... C^2 >> C *from recursive application of* C&C

To the extent that C can be interpreted as disfavoring any configuration of articulatory, perceptual, or processing difficulty, then it is entirely natural to assume that two violations of C in some context correspond to *more* difficulty of the *same* sort, and so on. Therefore, a power hierarchy is like a universal constraint subhierarchy in the sense of being grounded in a *scale* of 'difficulty'. If this is true, then the special case of self-conjunction derives a hierarchy that is in crucial respects equivalent to a universal constraint subhierarchy as in (12)b. This is true, it bears emphasizing, because of the inference of a scale over some *unitary* property, made possible by the assumption that two instances of a 'difficult' thing are more difficult in the *same* way than one. I assume, therefore, that constraint self-conjunction can be directly subsumed by the independently required notion of a universal constraint subhierarchy.

The similarity between (12)a and (12)b is less obvious precisely because the substantive basis of a conjunction of two different constraints has remained unclear. C_1 and C_2 may or may not be grounded in the same scale of difficulty. It is this absence of a simple, unitary, substantive basis, I claim, that causes problems for constraint conjunction. In the following section I attempt to show why constraint subhierarchies do better.

4. Illustrations

The goal of this section is to survey some proposals involving local constraint conjunction, in order to consider whether they might be reinterpreted along the lines suggested above. I choose for consideration only proposals which amount to central claims of the works containing them, rather than uses of constraint conjunction which play more minor roles in their works, or even the role of analytical expediencies. I also consider only proposals involving phonology. Even with these limitations in mind, the idea is not to provide exhaustive survey, but rather to consider a few representative cases.[8] Nor do I intend to explore to their limits the alternative proposals involving universal constraint subhierarchies. Each of these deserves a paper itself, and in some cases these papers already exist. Rather, the idea is to show that universal constraint hierarchies offer a plausible alternative approach to apparent *WOW* effects in general, and one worth exploring.

The table below summarizes the cases to be examined.

(14)

Constraints conjoined	Facts explained	References
NoCoda & featural markedness	Coda neutralization to the unmarked	Smolensky (1995), Itô and Mester (to appear-a)
Self-conjunction of markedness	Dissimilation ('OCP effects')	Itô and Mester (1996, to appear-b) Alderete (1997)
Feature-specific Ident & feature-specific Ident	Chain shifts	Kirchner (1996)
Anti-spreading and featural markedness	Restrictions on triggers of assimilation	Smolensky (1997)
Markedness and faithfulness	Derived environment effects	Łubowicz (1998, 2002), Itô and Mester (to appear-a)

4.1 NoCoda and featural markedness

Alternative approaches to coda neutralization, involving universal constraint subhierarchies, are already well known. The most common approach appeals to the idea of positional faithfulness (Selkirk 1994, Casali 1996, 1997, Beckman 1997, 1998). According to this approach, faithfulness constraints such as Ident are subdivided into those that refer to 'strong' positions and those that do not. One 'strong' position is the syllable onset. (See, besides the references just cited, Jun 1995, Padgett 1995, Lombardi 1999. Some of these, and Steriade 1997, argue that onset versus coda is not the right distinction, but I put this aside for the sake of simplicity. The point holds regardless.) Researchers often then posit a universal ranking $\text{Ident}_{\text{Strong}}$ >> Ident, for example $\text{Ident}_{\text{Onset}}$ >>

Ident (though see below). The most thorough discussion of the kind of phonetic scales underlying onset/coda asymmetries occurs in Steriade (1997), where it is argued that the number and quality of cues to phonological contrasts differ according to context. The cues associated with onset position are more numberous, and perceptually more robust, than those associated with coda position. Similar reasoning extends to place of articulation. (See Jun 1995 and Padgett 1995.)

To see how positional faithfulness is an instance of schema (11), consider (15). It seems reasonable that underlying the need for faithfulness is a processing difficulty: maintaining a correspondence relation between words is harder the more different they are. This is presumably why abstract underlying representations are sometimes reanalyzed, and why related output forms are sometimes made more similar (the latter referring to 'output-output' and paradigm leveling effects). Now, since the cues to onset place are more numerous and robust than those to coda place, the perceptual *difference* between, say, [da] and [ba] is greater than that between [ad] and [ab]. Therefore, we can infer a scale of processing difficulty as in (15)a. Here '$\Leftrightarrow$' stands for a correspondence relation, in this case between input S_I and output S_O. '$\neg\mathscr{F}(X)$' is shorthand for a lack of faithfulness in that correspondence relation, along feature (class) X; in other words, the corresponding segments differ in X. What (15)a says is that a correspondence relation between segments differing in place is more difficult when the relevant output segment is in onset position. This is what justifies (15)b. To help see this, a statement of Ident(Place) is provided in (15)c. It is equivalent to standard formulations, but makes comparison with (11) easier: depending on what position it is relativized to, Ident(Place) rules out one of the configurations in (15)a.

(15) a. $\neg\mathscr{F}(\text{Place})(S_I \Leftrightarrow S_{O\text{-Onset}}) > \neg\mathscr{F}(\text{Place})(S_i \Leftrightarrow S_O)$
b. Ident_{Ons}(Place) >> Ident(Place)
c. Ident(Place): Corresponding input and output segments do not differ in place

The success of positional faithfulness in accounting for positional neutralization is well documented. Here is the hypothetical example dealt with earlier by NoCoda&*Lab, recast in terms of positional faithfulness:

(16)

Input: /map/	Ident_{Ons}(Place)	*Lab	Ident(Place)	*Cor
a. map		**!		
b. ☞ mat		*	*	*
c. nat	*!		**	**

Some works, including Zoll (1996, 1997, 1998), Lombardi (2001), Smith (2002), and Walker (2002), have argued for positional markedness constraints in addition to positional faithfulness. (Steriade's 1997 proposals are also cast as positional markedness.) To the extent that it is rooted in phonetic or psycholinguistic matters, positional markedness is in principle equally compatible with the universal constraint subhierarchy idea. For example, Smith (2002) proposes pairs of constraints M/strong and M, where M is a markedness constraint and M/strong is relativized to a strong position. The relativized constraints are grounded in the requirement that strong positions have certain prominence-enhancing features. One such pair of constraints is

HTone/σ́, requiring that stressed syllables bear high tone, and its non-positional version HTone. These ideas seem consistent with universal subhierarchies like HTone/σ́ >> HTone, rooted in scales of phonetic difficulty. In the case at hand, the scale in question presumably involves degrees of perceptibility of prominence, depending on how many prominent features are provided.

The positional faithfulness subhierarchy, like those that follow, avoids the problem of overgeneration facing local constraint conjunction. It does this by grounding apparent *WOW* effects in genuine phonetic or psycholinguistic scales. Recall the worrisome possibility of a constraint NoCoda&*[+voice, -son] with a domain of the syllable, raised by Itô and Mester (to appear-b), which rules out [bat]. Such an example cannot arise in the positional faithfulness approach, because it is precisely the quality of cues associated with a specific segment that bear on faithfulness to that segment. This is no ad hoc stipulation of the theory either: the cue based approach is rooted in a rich phonetics literature documenting the perceptual relevance of the proposed cues. It is this same fact that prevents the equivalent of NoHeavySyllable&*[+voice, -son], to take just one example, where the problem is with the constraints conjoined as much as with the domain of application. On the assumption that there is no plausible scale of phonetic or psycholinguistic difficulty that makes obstruent voicing more difficult in heavy syllables than elsewhere, such a case cannot be derived.

The approach employing universal constraint hierarchies has one more advantage. Recall that in order to maintain the generalization that coda neutralization is to the unmarked, the local conjunction approach must rely on an additional stipulation given in (6), the Universal Conjoined Constraint Ranking Hypothesis. But the fact that labials remain more marked in the coda, compared to coronals, follows in the positional faithfulness approach *directly* from the independently necessary assumption *Lab >> *Cor. The full typology predicted by the relevant four constraints is shown in (17). The results depend on two universal constraint subhierarchies: one of faithfulness constraints, the other of markedness constraints. Given these constraints and fixed rankings, a language neutralizing to labial place in the coda is impossible. This fact requires no separate stipulation.

(17)	*Lab >> *Cor >> IdentOns >> Ident	*No coronals or labials anywhere*
	*Lab >> IdentOns >> *Cor >> Ident	*No labials anywhere, coronals in onset*
	IdentOns >> *Lab >> *Cor >> Ident	*Labials and coronals in onset only*
	*Lab >> IdentOns >> Ident >> *Cor	*No labials anywhere; coronals everywhere*
	IdentOns >> *Lab >> Ident >> *Cor	*Labials in the onset; coronals everywhere*
	IdentOns >> Ident >> *Lab >> *Cor	*Labials and coronals everywhere*

In this discussion of positional faithfulness and markedness, I assumed that the implicational relationship between the relevant constraints, such as $Ident_{Onset}$ and Ident, is captured by means of universal ranking, $Ident_{Onset}$ >> Ident. But there is another way to capture this relationship, in fact assumed by some work on positional faithfulness: posit that every violation of $Ident_{Onset}$ implies a violation of Ident. That is, assume that violations of Ident are a superset of violations of $Ident_{Onset}$—that Ident is more 'stringent' (Prince 2001, de Lacy 2001). Given a stringency relation between the constraints, we capture the implicational relationships without stipulated rankings. This predicts generally the same typologies, and the difference between assuming universal rankings and assuming a stringency relationship is subtle. (See McCarthy 2002a and the references just cited for some discussion.) It may be that one approach can subsume the other, and that the phonetic scales assumed throughout this paper project not universal constraint subhierarchies, but universal implicational relationships involving violation marks. To keep the discussion simple I assume the former approach throughout this paper.

4.2 Dissimilation as self-conjunction of markedness constraints

Itô and Mester (1996, to appear-b), and Alderete (1997) argue that dissimilation should be understood as the local self-conjunction of simple markedness constraints, rather than as following from a constraint such as the Obligatory Contour Principle (Leben 1973, Goldsmith 1976), which penalizes repetition per se. The fact that Yamato Japanese words cannot contain two voiced obstruents, for example, is a result of the self-conjunction of *[+voice, -son]. This analysis is illustrated in (18).

(18)

Input: /gaze/	*[+voi, -son]^2	Ident(voice)	*[+voi, -son]
a. gaze	*!		**
b. ☞ kaze		*	*
c. ☞ gase		*	*
c. kase		**!	

As we saw, local self-conjunction is virtually equivalent to the use of grounded universal constraint subhierarchies. A prohibition on voiced obstruents is likely grounded in articulation, since voicing is difficult to maintain with inhibited airflow. Picking up the earlier discussion of self-conjunction, it is reasonable to assume that two voiced obstruents involve more of this same difficulty than one, implying the scale shown in (19)a. This scale entails (19)b, given our notion of subhierarchy projection. I avoid the notation [+voice, -son]^2 since this is associated with the operation of conjunction.

(19) a. [+voice, -son]...[+voice, -son] > [+voice, -son]
b. *[+voice, -son]...[+voice, -son] >> *[+voice, -son]

On the other hand, the phonetic underpinnings of this kind of dissimilation in fact seem poorly understood, in comparison to those underlying positional neutralization of voice and place. Ohala (1981) argues that there is a class of dissimilatory phenomena that are motivated by ambiguity in the perceived speech signal, but long-distance dissimilations involving obstruent voicing or consonantal place of articulation do not fit into this class. It remains to be seen whether a simple notion of increasing articulatory difficulty, as suggested above, is at play, or something else, and whether this can shed light on which features undergo dissimilation, and within what domain. Indeed, the very question whether this sort of dissimilation involves nothing more than a compounding of simple markedness, or whether there is still a need for constraints on sequences per se, remains worth exploring. (See Suzuki 1998 on the latter view.)

4.3 Chain shifts as conjunction of Ident constraints

Chain shifts have often been handled by means of counterfeeding rule ordering in derivational frameworks, and they fall therefore under the category of a derivational opacity problem for a fully parallel Optimality Theory. Kirchner (1996) shows that chain shifts can be accounted for by means of the local conjunction of feature-specific Ident constraints.[9] Consider Kirchner's analysis of a synchronic three-step chain shift in Nzɛba, a Bantu language. Under certain morphological conditions, the vowels [a,ɛ,e,i] surface as [ɛ,e,i,i] respectively. That is, each vowel except /i/ raises by just one degree. Assuming Kirchner's assignment of features to these segments, shown below, the reader can see that each 'step' involves a change in one feature value.

(20)		low	high	ATR
	i	-	+	+
	e	-	-	+
	ɛ	-	-	-
	a	+	-	-

Kirchner's formulation of Raising, seen in the tableau below, is 'maximize vowel height', such that a vowel receives a violation for each 'step' (i.e., feature value) by which it falls short of [i]. The tableau considers the fate of each possible input vowel separately, with the notation a → e indicating an input-output mapping, following Kirchner. The important insight is that one violation of Ident is tolerated by this raising process (though the low-ranked simple Ident constraints are not shown here), while *two or more* violations are not. Kirchner accounts for this by means of the local conjunction of Ident constraints shown here.

(21)

Input: /a ɛ e i/	Ident(low)& Ident(ATR)	Ident(hi)& Ident(ATR)	Raising
a → a			***!
☞ a → ɛ			**
a → e	*!		*
a → i	*!	*!	
ɛ → a			**!*
ɛ → ɛ			**!
☞ ɛ → e			*
ɛ → i		*!	
e → a	*!		***
e → ɛ			*!*
e → e			*!
☞ e → i			
i → a	*!	*!	***
i → ɛ		*!	**
i → e			*!
☞ i → i			

The avoidance of too much vowel movement makes this sort of case very similar to that of positional faithfulness. Recall that faithfulness violations in strong positions imply a greater difference between corresponding forms than identical violations in weak positions, since distinctions in strong positions are more robustly cued. The scale of processing difficulty posited in (22)a is entirely analogous to that seen for positional faithfulness: the more different the corresponding forms, the greater the burden of keeping them in correspondence. Indeed, it is conceivable that (15)a and (22)a reduce to one notion. This leads to the universal hierarchy (22)b. For clarity of comparison with Kirchner's analysis, I assume that Ident(Height), stated in (22)c, reckons violations over [high], [low], and [ATR]. I assume also that these form a feature class in the sense of Padgett (2002a), corresponding precisely to the 'Height' node in feature geometry (Odden 1991), but nothing crucial hinges on this.

(22) a. $\neg\mathscr{F}(\text{3Height})(S_i \Leftrightarrow S_O) > \neg\mathscr{F}(\text{2Height})(S_i \Leftrightarrow S_O) > \neg\mathscr{F}(\text{1Height})(S_i \Leftrightarrow S_O)$

b. Ident(3Height) >> Ident(2Height) >> Ident(Height)

c. Ident(Height): Corresponding input and output segments do not differ in height

A distinction among faithfulness constraints similar to (22)c is in fact proposed by Gnanadesikan (1997), precisely to account for chain shifts. Padgett (to appear) argues for such a distinction as well (though not referring to vowel height), rooted in differences in the perceptual distance of faithfulness violations. The application of this hierarchy to the Nzɛba data is illustrated below. A mapping such as /a/ → [e] violates Ident(2Height), since these vowels differ by two of the three features, [high, low, ATR]. All faithfulness violations are counted in this way. (Since any height shift greater than one is ruled out anyway, Ident(3Height) is not shown in the tableau.)

(23)

Input: /a ɛ e i/	Ident(2Height)	Raising
a → a		***!
☞ a → ɛ		**
a → e	*!	*
a → i	*	
ɛ → a		**!*
ɛ → ɛ		**!
☞ ɛ → e		*
ɛ → i	*!	
e → a	*!	***
e → ɛ		*!*
e → e		*!
☞ e → i		
i → a	*	***
i → ɛ	*!	**
i → e		*!
☞ i → i		

Moreton and Smolensky (2002) argue that some chain shifts involve segmental deletion. For example, in Catalan, [n] in stressed word-final syllables of nouns and adjectives deletes, e.g. /plan/ → [pla] 'plane'. By a more general process, word-final stops delete after nasals, e.g., /kant/ → [kan] 'singing'. But the nasal is retained in the latter forms, *[ka]. Moreton and Smolensky propose to handle this by means of conjoined Max constraints (penalizing segmental deletion). Thus, one violation of Max is permissible, but two (or more), as in *[ka], are not. Since this is a form of self-conjunction, it is possible to recast it directly in terms of a universal subhierarchy. Again, increasing differences between corresponding forms implies increasing processing difficulty. In this case, the differences involves the presence versus absence of segments. This is shown in (24). Of course, the strings [kant] etc. (24)a stand in for any strings, and strings can differ by other than the endmost segments.

(24) a. $\neg\mathscr{F}(\text{kant}_i \Leftrightarrow \text{k}_O) > \neg\mathscr{F}(\text{kant}_i \Leftrightarrow \text{ka}_O) > \neg\mathscr{F}(\text{kant}_i \Leftrightarrow \text{kan}_O)$

b. *n*-Max > ... > 2-Max > Max

Having said this, it is not clear that the Catalan case represents a true chain shift. What makes a chain shift a compelling one is its apparent unitary nature. First, the shifts involved all occur along what might plausibly be considered the same phonetic dimension. This might be the dimension of vowel height, as in Nzɛba, or it might be that of sonority, for example, as in Irish eclipsis, where /t/ becomes [d] while /d/ becomes [n] (see Ní Chiosáin 1991 and Gnanadesikan 1997, for example). It is not obvious that increasing deletion involves such a scale, though (A.D. Green, p.c.) suggests a scale of recoverability is at work: the input /kant/ is increasingly recoverable from the outputs [k], [ka], [kan], and [kant]. Second, the processes involved otherwise seem to be 'one'. This seems dubious in Catalan, because deletion of /n/, unlike the other deletion process, affects only nouns and adjectives. What we seem to have are simply two independent processes that interact in a derivationally opaque manner.

Moreton and Smolensky (2002) also argue that a local conjunction account of chain shifts predicts correctly the absence of shifts involving both deletion and insertion, e.g. hypothetical /AxB/ → [AB] but /AB/ → [AyB].[10] (They note, though, that other accounts for chain shifts employing correspondence theory predict this too.) The question, in our terms, is whether a scale such as (25) is phonetically or psycholinguistically plausible. In all of the cases seen so far involving differences between corresponding forms, the properties by which forms differ seem plausibly scalar, that is, involving monotonic increases in some arguably unitary property: robustness of cues (positional faithfulness), vowel height (Nzɛba chain shift), and possibly recoverability (Catalan).[11] But in this sense (25) does not seem to be a genuine scale at all. [AxB] is simply not guaranteed to be perceptually more different from [AyB] than either of these is from [AB]. This becomes obvious if the segments [x] and [y] are similar. A hypothetical example is /badi/ → [bad], /bad/ → [badɪ]. The would-be most distinct forms, [badi] and [badɪ], are perceptually closer. There might be examples that work better, but it is clear that (25) has no general validity. Moreover, if chain shifts are required more specifically to involve a single *phonetic* dimension, such as vowel height or sonority, then (25) certainly does not qualify: insertion and deletion are not different degrees of the same thing. In sum, the requirement that universal constraint subhierarchies be plausibly grounded succeeds in ruling out this non-existent chain shift as well as constraint conjunction does.

(25) $\neg\mathscr{F}(\mathrm{AxB_i} \Leftrightarrow \mathrm{AyB_O}) > \neg\mathscr{F}(\mathrm{AxB_i} \Leftrightarrow \mathrm{AB_O}), \neg\mathscr{F}(\mathrm{AB_i} \Leftrightarrow \mathrm{AyB_O})$

This restrictive potential of grounding, in addition, goes beyond the ability to rule out (25). There are many other implausible chain shifts that constraint conjunction alone unfortunately does allow, but which can be excluded by appeal to grounded scales. One hypothetical example: /be/ → [bo], /bo/ → [bõ]. It is possible to model such a shift by a conjunction of Ident(back) and Ident(nasal).[12] But if chain shifts must involve a unitary phonetic scale, as backing and nasalization do not, they cannot be modeled using grounded subhierarchies.

4.4 Restrictions on triggers of assimilation through conjunction of anti-spreading and featural markedness constraints

This case is due to Smolensky (1997), who analyzes a complex range of facts involving [ATR] spreading in Lango, a Nilotic language, taking as a point of departure proposals of Archangeli and Pulleyblank (1994). Lango is interesting because, though it contrasts [ATR] across all vowel types (having [+ATR] [i,e,ə,o,u] versus [-ATR] [ɪ,ɛ,a,ɔ,ʊ]), restrictions on the cooccurrence of [ATR] and other vowel features manifest themselves in the harmony processes. Smolensky employs constraint conjunction in order to account for several facets of the data. Here I focus primarily on one of these facets, the failure to spread [-ATR] from high vowels. The examples in (26), from Archangeli and Pulleyblank, illustrate rightward harmony targeting the first person singular possessive suffix [-ə/a] ([ə] being the [+ATR] variant). Harmony occurs in (26)a-b, but not in (26)c, where the trigger is non-high. Since the default realization of the suffix vowel is [a], harmony may or may not be occurring in (26)d. It is not a coincidence that all of these examples involve geminates, but I discuss this later. (Tones are not transcribed.)

(26)	a.	piggə	'my juice'	b.	atɪnna	'my child'
		opukkə	'my cat'		lʊtta	'my stick'
	c.	gwenna	'my chickens'	d.	bɛlla	'my wheat'
		dokka	'my cattle'		kɔmma	'my chair'
		ñəɳɳa	'my crocodile'		calla	'my picture'

Archangeli and Pulleyblank argue persuasively that restrictions on targets and triggers of [ATR] spreading reflect universal preferences involving the co-occurrence of [ATR] with other vowel place features. These are grounded in articulatory and acoustic (in)compatibility. There is a cross linguistic preference for [+ATR] vowels to be high, for example, and vice versa, based on the antagonism between tongue raising and pharyngeal constriction. Translated into Optimality Theoretic terms, this corresponds to a universal subhierarchy *[+ATR, -high] >> *[+ATR, +high]. Archangeli and Pulleyblank argue in addition that all conditions on rules must reflect these patterns of grounding. In the present case, this implies a preference for [+ATR] to *spread* from high vowels and not non-high vowels, just as in Lango.

Smolensky's proposal is to account for the latter fact by locally conjoining the constraint *[+ATR, -high] and another penalizing spreading, called *HD (for 'No head domain', referring to the source of a feature spreading domain). The domain of the conjunction is stipulated to be a single segment, so that the effect is that [+ATR] spreading cannot occur from a non-high vowel. This analysis is illustrated below. ('Agree' is the constraint driving spreading.)

(27)

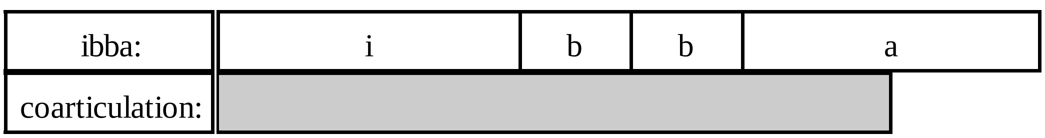
Input: /pigga/	*[+ATR, -high]&*HD	Agree(ATR)
a. pigga		*!
b. ☞ piggə		
Input: /gwenna/		
a. ☞ gwenna		*
b. gwennə	*!	

Can failure to spread from a non-high vowel be seen as reflecting not a conjunction of marked states, but a particularly bad spot on a unitary scale of difficulty? There is good evidence that vowel harmony has its basis at least in part in phonetic vowel-to-vowel coarticulation (Ohala 1993, 1994, Bessell 1998, Majors 1998, Beddor et al. 2001, to appear, and Przezdziecki 2002).

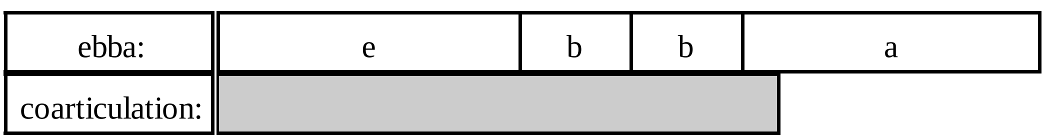
The essence of vowel-to-vowel coarticulation is that vowel gestures overlap (most) consonant gestures to a large enough extent that vowels separated by consonants affect each other's acoustic profile: high vowels slightly raise neighboring non-high vowels, and so on, more so at the latter's neighboring edge (Öhman 1966, see Farnetani and Recasens 1999 for an overview). Languages differ in the strength of leftward versus rightward coarticulation and other factors, but it is ubiquitous in languages to some extent. There is also some evidence for the view that the greater the degree of coarticulation, the more likely harmony is to result (Beddor et al. 2001). The feature co-occurrence constraints of Archangeli and Pulleyblank (1994) are interesting in this light: if [+ATR] and [-high] involve antagonistic articulations to some degree, then we might expect [+ATR] gestures of high vowels to be greater in degree. McKay (1976) found just this in a study of English (though the identification of the English tense/lax distinction and ATR is debatable). It is therefore reasonable to expect high [+ATR] vowels to cause more coarticulation in neighboring vowels—and so be more likely triggers of harmony.

Consider the diagrams below. (28)a considers a hypothetical form [ibba]. The top row of the diagram is a schematic representation of the vowel and consonant acoustic phases in time: vowel formant information is present during the periods labeled by the vowels, and masked during the consonantal closure phases. The second row is intended to suggest an extent in time during which the first vowel affects the formants of the second. The actual extent for Lango is not known; but we can infer that coarticulation affects more of the audible portion of the second vowel in (28)a than in (28)b, given the discussion above, and that is what matters for our purposes.

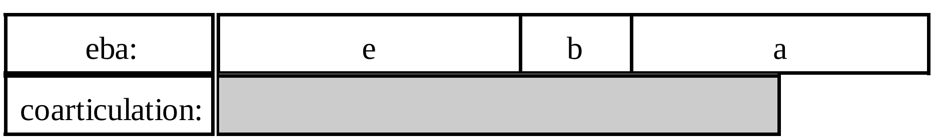
(28) a.



b.



c.



Consider (28)c now. Since the vowel is again mid, we expect an [+ATR] gesture of equal magnitude as in (28)b, all else equal. But the audible effect on the second vowel is nevertheless greater, for obvious reasons: there is only one consonant intervening. If the extent of *audible* coarticulation is really a factor in conditioning harmony, we expect the possibility that harmony will occur in cases like (28)a and (28)c, but not in (28)b. And this is what occurs in Lango; harmony spreads from all vowels across singleton consonants (data again from Archangeli and Pulleyblank):[13]

(29)	a.	ciɳə	'my hand'	b.	yɪba	'my tail'
		ɳutə	'my neck'			
	c.	ɳetə	'my side'	d.	lɛba	'my tongue'
		wodə	'my son'		bwɔma	'my wing'
					waɳa	'my eye'

The intuition that unifies these cases is straightforward: assimilation is preferred when the target vowel is audibly affected more by coarticulation. In principle we might formalize this in two different ways. First, we could posit a family of Spread constraints relativized to degree of coarticulation. Second, we might see this as another manifestation of positional faithfulness. The second approach rests on the idea that the [ATR] value of the vowel [a] in (28)a is in a less prominent 'position', that is, it resides in a vowel that has suffered a good deal of coarticulation. (Compare the proposal of Jun 1995 and Padgett 1995 to distinguish Ident(Place) constraints for nasals versus obstruents, since place is known to be less well-cued in the former.) In the absence of arguments one way or the other, I assume the first approach here. In the scale in (30)a, $[\ ]_{[ATR]}$ denotes a span of [ATR]. Given the above, it is articulatorily more difficult to avoid harmony when the target vowel is highly affected by coarticulation, indicated by [~V]. This implies a hierarchy like that of (30)b, where $\text{Spread}_{\text{Strong}}$ is relativized to this more difficult context.

(30) a. $[V]_{[ATR]}C_0[\sim V]_{[ATR]} > [V]_{[ATR]}C_0[V]_{[ATR]}$
b. $\text{Spread}_{\text{Strong}}$(+ATR) >> Spread(+ATR)

The application of this alternative analysis to the same data, and to words with singleton consonants, is shown in (31).

(31)

Input: /pigga/	$\text{Spread}_{\text{Strong}}$(ATR)	Ident(ATR)	Spread(ATR)
a. pigga	*!		*
b. ☞ piggə		*	
Input: /gwenna/			
a. ☞ gwenna			*
b. gwennə		*!	
Input: /ɳeta/			
a. ɳeta	*!		*
b. ☞ ɳetə		*	

In order to account for the different behavior of words with geminate versus singleton consonants, Archangeli and Pulleyblank (1994) must write two independent rules. Smolensky (1997) unifies the processes, by further appeal to constraint conjunction: the dominant constraint he employs is actually a three-way conjunction, *[+ATR, -high]&*HD&*$[CC]_{ATR}$, where *$[CC]_{ATR}$ penalizes an [ATR] span across geminates. The approach advocated here, however, makes a more direct and explanatory connection between the high vowel condition and the geminates. It retains Smolensky's insight that spreading from non-high vowels across geminates is somehow the 'worst of the worst', but it does this once again by appeal to a *unitary* underlying generalization. An analogous approach can also explain further patterns of [ATR] harmony in Lango. One of these, for example, is a regressive [+ATR] spreading process that is always limited to high triggers. When two consonants intervene, the trigger must be a high *front* vowel. Just as [+ATR] and [-high] are antagonistic, [+ATR] and [+back] are antagonistic.

Universal hierarchies like (30)b are restricted by the requirement that they be grounded in independently motivated phonetic scales, as Archangeli and Pulleyblank argue. This is once again the response to the worry of overgeneration. And once again the domain of the effect, extending from the source segment to encompass whatever the harmony span is, follows from the phonetic explanation of spreading itself, rather than requiring separate stipulation. The claimed advantage regarding overgeneration might appear threatened by the appeal above to degrees of coarticulation. If we could make reference to arbitrarily fine degrees of auditory effect, then presumably there would be no restrictive advantage to the account at all. But we needn't assume this. It seems likely that appeal to a very small number of distinctions (only two were needed above) would be all that is required to capture observed harmony effects. In comparison, constraint conjunction appears to make an infinite number of derived constraints available to the theory. And a great deal of recent work argues that phonology must indeed make reference to more phonetic distinctions than assumed by distinctive feature theory (e.g., Browman and Goldstein 1986 et seq., Steriade 1997, 2000, Kirchner 1997, 2000, Flemming 1995, 2001, Boersma 1998, Zhang 2000, and Ní Chiosáin and Padgett 2001). These added distinctions allow us to capture phonological generalizations otherwise missed, as just seen above. The worry of overgenerating contrast is handled by appeal to output constraints that, in one way or another, restrict logically possible contrasts to the right set. For example, assuming there are simply no faithfulness constraints regulating differences in degree of coarticulation, it can never form the basis of a contrast (Kirchner 1997). For further discussion of contrast and output constraints, see Flemming (1995, 2001), Ní Chiosáin and Padgett (2001), and Padgett (2002b, to appear).

4.5 Derived environment effects through conjunction of markedness and faithfulness

The last example is due to Łubowicz (1998, 2002), who argues that derived environment effects can be accounted for by means of the conjunction of markedness and faithfulness constraints. (See also Baković 1999, Walker 2002, and Itô and Mester to appear-a.) Consider her example from Polish, originally due to Rubach (1984). Two processes are relevant to the argument, a morphologically governed velar mutation, and a process spirantizing [ʤ] to [ʒ]. The first is responsible for alternations such as [krok] 'step' ~ [kroʧ-ek] (diminutive) and [droŋg] 'pole' ~ [drõwʒ-ek] (dim.). (The alternation between [oN] and [õw] occurs for independent reasons.) The second, spirantization, explains why we find [drõwʒ-ek] instead of *[dronʤ-ek]: compare [k] ~ [ʧ], where the output is an affricate. Spirantization deaffricates the voiced postalveolar affricate. However, this affects only [ʤ] derived by velar mutation. Words with underlying /ʤ/ do not spirantize: compare [brɪʤ] 'bridge' ~ [brɪʤ-ek]. Therefore Łubowicz (following Rubach) treats this as an instance of a process, spirantization, conditioned by a phonologically derived environment.

The analysis by means of conjunction goes as follows. Given the failure of spirantization of underlying /ʤ/, the ranking Ident(cont) >> *ʤ can be inferred. (Łubowicz takes affricates to be strident stops, lacking any [+continuant] value.) To explain the spirantization in [drõwʒ-ek], the extra faithfulness violation incurred by a velar-mutated form is capitalized on. Assuming mutation involves a change from [+dorsal] to [+coronal], it violates a constraint Ident(Coronal). Łubowicz posits a conjoined constraint Ident(Cor)&*ʤ, which will be violated precisely by [ʤ] derived from velars. The analysis is illustrated below. This tableau presupposes that mutation must take place, so that no candidate like *[droŋgek] is considered.

(32)

Input: /brɪʤ-ek/[14]	Ident(Cor)&*ʤ	Ident(cont)	*ʤ
a. ☞ brɪʤek			*
b. brɪʒek		*!	
Input: /droŋg-ek/			
a. dronʤek	*!		*
b. ☞ drõwʒ-ek		*	

Łubowicz argues that all derived environment effects can be analyzed in this way, including those derived morphologically rather than phonologically. This proposal faces some worries not discussed here and has several competitors, old and new. (For some discussion of the challenges for such an account, see Anttila 1999, Blumenfeld 2002 and McCarthy 2002b.)[15] However, one alternative approach, due to Burzio (1999, 2000), seems compatible with treatment by means of universal constraint subhierarchies, and so should be considered here.[16]

In a far-ranging proposal, Burzio argues that much of phonology is governed by a principle of 'gradient attraction'. The generalization is said to lie behind output-output effects, contrast dispersion, and other phenomena, but its relevance to derived environment effects is of interest here. The intuition behind gradient attraction is that the more similar two forms are, the more pressure they are under to get even more similar, a generalization said to be fundamental to connectionist work. As one consequence, faithfulness violations are more 'difficult', or avoided, when the forms in question are independently more similar. The less they are alike, on the other hand, the less speakers are inclined to relate them to one another and so keep them similar. If this is correct, then a scale as in (33)a is at work. Put simply, the more similar a form W_i is to its respective base B_i, the worse it is for them to differ. As always, a universal subhierarchy is implied, (33)b. For reference, the Ident constraint is state in (33)c.

(33) a. Given words W_1, W_2, ..., W_n, morphologically related to base forms B_1, B_2, ..., B_n, respectively, where W_1 is least similar to B_1, W_2 more similar to B_2, etc.:
$\neg\mathscr{F}(W_n \Leftrightarrow B_n) > \neg\mathscr{F}(W_{n-1} \Leftrightarrow B_{n-1}) > ... > \neg\mathscr{F}(W_2 \Leftrightarrow B_2) > \neg\mathscr{F}(W_1 \Leftrightarrow B_1)$

b. $\text{Ident}_n\text{-OO(F)} >> \text{Ident}_{n-1}\text{-OO(F)} >> ... >> \text{Ident}_2\text{-OO(F)} >> \text{Ident}_1\text{-OO(F)}$

c. Ident-OO(F): Corresponding output segments do not differ in feature F

Consider how this applies to the case at hand. The form [brɪʤek] has as its morphological base form [brɪʤ]. Suppose the Ident constraint relating these output forms, called Ident_n-OO(cont) here, dominates the constraints forcing spirantization, which I encapsulate here as 'Spirantize'. Then spirantization will be blocked, as shown below. Notice that both candidates (34)a-b strongly resemble the base form, putting the suffix aside; (34)b differs by one [continuant] value. Now compare (34)c-d. By assumption, velar mutation must apply, so that the only live candidates are those whose underlying velar has been altered to postalveolar. But these candidates are therefore less similar to their base forms already, compared to (34)a-b, independently of spirantization. If their relation to the base is therefore governed by a separate, lower-ranked Ident constraint, according to (33), then spirantization will go through, as shown.

(34)

Input: /brɪʤ-ek/ Base: [brɪʤ]	Ident_n-OO(cont)	Spirantize	Ident_{n-1}-OO(cont)
a. ☞ brɪʤek		*	
b. brɪʒek	*!		*
Input: /droŋg-ek/ Base: [droŋk]			
c. dronʤek		*!	
d. ☞ drõwʒek			*

These ideas are obviously preliminary. The obvious challenge is to make the crucial notion 'independently similar' explicit and test it against more cases. Further, McCarthy (2002b) notes some challenges for the constraint conjunction approach to derived environments which must be faced by this approach equally. For example, phonologically derived environments are always relevant to the process in question in the specific sense of feeding it. Thus, spirantization of [ʤ] to [ʒ] is contingent upon the independent application of velar mutation; the latter mutates /g/ to [ʤ]. As McCarthy notes, we would not want to allow spirantization to affect the first segment of hypothetical /ʤoʤ-ek/ → *[ʒoʒ-ek] simply because mutation has applied somewhere in the stem. If the suggestion here is to fly, it must be clarified why 'independently similar' is locally restricted in this way.

But the point is to suggest that an alternative account for derived environment effects based on grounded subhierarchies (adapting Burzio's gradient attraction proposal) is conceivable, and worth investigation. While conjunction of markedness and faithfulness captures the desired effect (at least for some cases), it is unclear *why* markedness and faithfulness should combine in this way. This is the problem of overgeneration in another guise: we must rule out many logically possible constraint conjunctions, and simultaneously justify those that remain. Whether gradient attraction reflects a truth about phonology or not, it is rooted once again in a simple, unitary, underlying generalization.

5. Conclusion

I have argued that constraint conjunction might be eliminated from the repertory of devices allowed within Optimality Theory. I have also argued that at least some apparent 'worst of the worst' effects can be reanalyzed by appealing to a notion independently required by the theory, the universal constraint subhierarchy. If this turns out to be correct, then Optimality Theory's original claim that constraint domination is strict will stand unmoderated. The intuition behind using constraint subhierarchies is that explanations should be rooted in simple, unitary generalizations, while the mechanism of conjunction, by its very nature, captitalizes on complex generalizations. In fact, this paper's motto might be that the only 'worst of the worst' effects are those we might call 'too much of a (single) bad thing'. In addition, the pressing but unsolved questions about restrictions on conjunction are recast as questions about restrictions on universal subhierarchies. If universal subhierarchies are grounded in phonetic or psycholinguistic scales, then the theory will be constrained by the requirement that posited hierarchies appeal to such scales, and that these be empirically verifiable at least in principle.

Given the many uses to which constraint conjunction has been put, it seems highly unlikely that all cases could be, or should be, reanalyzed precisely along these lines. But the considerations of section 3, where the underlying similarities between 'worst of the worst' and 'too much of a bad thing' were examined, suggest that many of them should be. In any case, our larger conclusions suggest that *all* cases should be reconsidered one way or another.

The question constraint conjunction raises, whether linguistic generalizations obey strict domination, or whether constraints can instead have additive effects roughly as posited by connectionist models, is surely one of the most important and far-reaching questions facing Optimality Theory. Whatever the answer turns out to be, I hope to spark further discussion of the issue.

Notes

* For feedback that led to much improvement of this paper, I am indebted to Caroline Féry, Linda Lombardi, Máire Ní Chiosáin, Nathan Sanders, Adam Ussishkin, Rachel Walker, Andy Wedel, and the audience of a talk at the University of Potsdam on July 16.

1. Smolensky and Moreton (2002) attribute constraint conjunction also to Green (1993).

2. Hewitt and Crowhurst (1996) and Crowhurst and Hewitt (1997) propose a kind of 'constraint conjunction' that is quite different from that addressed here.

3. Itô and Mester (to appear-a) revise this position, arguing that some (but not all) kinds of markedness-plus-faithfulness conjunction are required.

4. Another problem involves making precise what notions like 'same family', or 'constraint argument' mean. (On the latter see also discussion in Anttila 1999 and Blumenfeld 2002.) Since I'm advocating a different approach altogether, I don't pursue this question.

5. Fukazawa and Lombardi (2002) similarly discuss the worry of overgeneration for conjunction, and also argue that the theory must appeal to phonetic grounding in order to address the problem. However, they assume conjunction within a constraint 'family' exists (following Fukazawa and Miglio 1998), and argue, unlike the proposal here, that grounded constraints might be formally arbitrary or complex.

6. It is likely that perceptual factors also play a role in disfavoring at least some of these segment types. Nasality reduces the perceptual distance between vowels, for example (Wright 1986, Beddor 1993).

7. Whether the sonority hierarchy fits into this schema depends on the prior question of just what underlies the sonority hierarchy, the phonetic bases of which are famously elusive. Ohala and Kawasaki-Fukumori (1997) argue that the real primitives underlying 'sonority' involve perceptual salience of phonetic cues. If this is true then recasting (9) as (11), in a way similar to cases discussed below, would seem promising.

8. Among the significant proposals not addressed are Itô and Mester's (2001, to appear-a) proposal to employ conjunction to account for certain problems of derivational opacity, and Downing's XXX conjunction of alignment and Onset to derive the failure to reduplicate onsetless syllables.

9. Kirchner actually employs the Parse formulation of faithfulness introduced by Prince and Smolensky (1993), but the logic carries over to correspondence-theoretic Ident.

10. See the authors cited for ideas about why chain shifts involving only insertion do not seem to occur.

11. 'Unitary' does not have to mean unanalyzable. For example, 'robustness' of phonetic cues is a matter both of the number, and the strength, of cues, and the cues involved may be various, as they certainly are for place and voicing contrasts. But the resulting degree to which a contrast is perceptible is nevertheless a reasonably unitary property itself.

12. To do this requires separate constraints demanding vowel backing and nasalization. But it is equally true that the Catalan example involves separate constraints forcing the two components of the shift, as noted. In this sense, neither example is a true chain shift.

13. The account here predicts that spreading will be restricted to high vowel triggers whenever two consonants intervene, and not simply when a geminate intervenes. However, all of the relevant data provided by Archangeli and Pulleyblank (1994) involve geminates, and their rule likewise generalizes to any CC sequence.

14. Many take the actual input to be more abstract, among other things having two 'yer' vowels, /brɪʤ+ĭk+ĭ/, but this can be ignored for our purposes.

15. The Polish case is worrisome also because the cited examples of non-derived [ʤ] are all in borrowed items like [brɪʤ] and [banʤo] 'banjo'. They therefore seem amenable to an analysis in terms of lexical strata as in Itô and Mester (1999).

16. Though my adaptation is not one that Burzio himself is likely to endorse. He suggests, contrary to the line pursued here, that additive constraint interaction (as roughly captured by constraint conjunction) is the norm, and strict domination the exception.

References

Alderete, John. 1997. Dissimilation as local conjunction. In *Proceedings of the North East Linguistics Society 27*, edited by K. Kusumoto. Amherst, MA: GLSA, Dept. of Linguistics, University of Massachusetts:17-31.

Anttila, Arto. 1999. Derived environment effects in colloquial Helsinki Finnish: Ms., Boston University.

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

Baković, Eric. 1999. Assimilation to the unmarked: Ms., Pennsylvania State University.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* 14.1:1-46.

Beckman, Jill. 1998. *Positional faithfulness*: Ph.D. dissertation, University of Massachusetts, Amherst.

Beddor, Patrice. 1993. The perception of nasal vowels. In *Nasals, nasalization, and the velum*, edited by M. K. Huffman and R. A. Krakow. New York: Academic Press:171-196.

Beddor, Patrice, R. Krakow, and S. Lindemann. 2001. Patterns of perceptual compensation and their phonological consequences. In *The Role of Speech Perception in Phonology*, edited by E. Hume and K. Johnson. New York: Academic Press:55-78.

Beddor, Patrice S., J. Harnsberger, and S. Lindemann. to appear. Language-particular patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of phonetics*.

Bessell, Nicola J. 1998. Local and non-local consonant-vowel interaction in Interior Salish. Phonology 15.1:1-40.

Blumenfeld, Lev. 2002. Russian palatalization in stratal OT: morphology and [back]. Berkeley, CA: Paper presented at the Trilateral Phonology Weekend, UC Berkeley.

Boersma, Paul. 1998. *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Browman, Catherine, and Louis Goldstein. 1986. Towards an articulatory phonology. Phonology Yearbook 3:219-252.

Burzio, Luigi. 1999. Surface-to-surface constraints: when your representations turn into constraints. Baltimore: Ms., Johns Hopkins University.

Burzio, Luigi. 2000. Segmental contrast meets output-to-output faithfulness. *The linguistic review* 17.2-4:367-384.

Casali, Roderic F. 1996. Resolving hiatus. Los Angeles: Ph.D. dissertation, UCLA.

Casali, Roderic F. 1997. Vowel elision in hiatus contexts: which vowel goes? Language 73:493-533.

Cohn, Abigail. 1993. The status of nasalized continuants. In *Nasality*, edited by M. Huffman and R. Krakow. San Diego: Academic Press:329-367.

Crowhurst, Megan, and Mark Hewitt. 1997. Boolean operations and constraint interactions in Optimality Theory. Chapel Hill, NC and Waltham, MA: Ms., University of North Carolina and Brandeis University.

de Lacy, Paul. 2001. Conflation and scales. In *Proceedings of NELS 32*, edited by M. Hirotani:93-112.

Farnetani, Edda, and Daniel Recasens. 1999. Coarticulation models in recent speech production theories. In Coarticulation: theory, data, and techniques, edited by W. J. Hardcastle and N. Hewett. Cambridge: Cambridge University Press:31-65.

Flemming, Edward. 1995. Auditory representations in phonology. Doctoral Dissertation, UCLA, Los Angeles, CA.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. Phonology 18.1:7-44.

Fukazawa, Haruka, and Linda Lombardi. 2002. Complex constraints and linguistic typology in Optimality Theory: Ms., University of Maryland, College Park.

Fukazawa, Haruka, and Viola Miglio. 1998. Restricting conjunction to constraint families. In *Proceedings of Western Conference On Linguistics 9*:102-117.

Gafos, Adamantios, and Linda Lombardi. 1999. Consonant transparency and vowel echo. In *Proceedings of NELS 29, vol. 2: papers from the poster sessions*, edited by P. Tamanji, M. Hirotani and N. Hall. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts:81-95.

Gnanadesikan, Amalia. 1997. *Phonology with ternary scales*: Doctoral dissertation, University of Massachusetts, Amherst.

Goldsmith, John. 1976. *Autosegmental phonology*: PhD dissertation, MIT [published 1979, Garland: New York].

Green, T. 1993. The conspiracy of completeness. New Brunswick, NJ: Paper presented at the Rutgers Optimality Workshop 1, Rutgers University.

Hewitt, Mark, and Megan Crowhurst. 1996. Conjunctive constraints and templates in Optimality Theory. In *Proceedings of NELS 26*, edited by J. Beckman. Amherst, MA: GLSA, Department of Linguistics, University of Massachusetts:101-116.

Itô, Junko, and Armin Mester. 1996. Structural economy and OCP interactions in local domains: Handout from the 1996 Western Conference on Linguistics (WECOL), University of California, Santa Cruz.

Itô, Junko, and Armin Mester. 1999. The phonological lexicon. In The handbook of Japanese linguistics, edited by N. Tsujimura. Malden, MA: Blackwell:62-100.

Itô, Junko, and Armin Mester. 2001. Structure preservation and stratal opacity in German. In *Segmental phonology in Optimality Theory: constraints and representations*, edited by L. Lombardi. Cambridge: Cambridge University Press:261-295.

Itô, Junko, and Armin Mester. to appear-a. On the sources of opacity in OT: coda processes in German. In *The optimal syllable*, edited by C. Féry and R. v. d. Vijver. Cambridge: Cambridge University Press.

Itô, Junko, and Armin Mester. to appear-b. *Markedness and word structure: OCP effects in Japanese*. Cambridge, MA: MIT Press.

Jun, Jongho. 1995. *Perceptual and articulatory factors in place assimilation: an Optimality Theoretic approach*: Unpublished Ph.D. dissertation, UCLA.

Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27:341-350.

Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14.1:83-111.

Kirchner, Robert. 2000. Geminate inalterability and lenition. *Language* 76.3:509-545.

Leben, Will. 1973. Suprasegmental Phonology. Doctoral dissertation, MIT.

Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural language and linguistic theory* 17.2:267-302.

Lombardi, Linda. 2001. Why place and voice are different: constraint-specific alternations in Optimality Theory. In *Segmental phonology in Optimality Theory: constraints and representations*, edited by L. Lombardi. Cambridge: Cambridge University Press:13-45.

Łubowicz, Anya. 1998. Derived environment effects in Optimality Theory. Amherst, MA: Ms., University of Massachusetts [ROA 239].

Łubowicz, Anya. 2002. Derived environment effects in Optimality Theory. Lingua 112:243-280.

Majors, Tivoli. 1998. Stress dependent harmony: phonetic origins and phonological analysis. Austin, TX: Doctoral Dissertation, University of Texas, Austin.

McCarthy, John J. 2002a. A thematic guide to Optimality Theory. Cambridge: Cambridge University Press.

McCarthy, John J. 2002b. Comparative markedness. Amherst, MA: Ms., University of Massachusetts.

McKay, I. 1976. Ultrasonic investigation of anterior pharyngeal wall (tongue root) position in syllables containing tense and lax vowels. Cincinnatt, OH: Ph.D. dissertation, University of Cincinnati.

Moreton, Elliot, and Paul Smolensky. 2002. Typological consequences of local constraint conjunction. In Proceedings of WCCFL 21, edited by L. Mikkelsen and C. Potts. Cambridge, MA: Cascadilla Press.

Nathan, Lance. 2001. Constraint conjunction and the OTP. Cambridge, MA: Ms., MIT.

Ní Chiosáin, Máire. 1991. Topics in the phonology of Irish. Doctoral dissertation, University of Massachusetts, Amherst.

Ní Chiosáin, Máire, and Jaye Padgett. 2001. Markedness, segment realization, and locality in spreading. In Segmental phonology in Optimality Theory: constraints and representations, edited by L. Lombardi. Cambridge: Cambridge University Press:118-156.

Odden, David. 1991. Vowel geometry. *Phonology* 8:261-289.

Ohala, John J. 1981. The listener as a source of sound change. In *Chicago Linguistic Society: papers from the parasession on language and behavior,* edited by C. S. Masek, R. A. Hendrick and M. F. Miller. Chicago, IL: Chicago Linguistic Society:178-203.

Ohala, John. 1993. Coarticulation and phonology. Language and speech 36:155-170.

Ohala, John J. 1994. Towards a universal, phonetically-based, theory of vowel harmony. In *Proceedings of ICSLP 94*:491-494.

Ohala, John J., and Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Language and its ecology: essays in memory of Einar Haugen, edited by S. Eliasson and E. H. Jahr. Berlin: Mouton de Gruyter:343-365.

Öhman, S.E.G. 1966. Coarticulation in VCV utturances: spectrographic measurements. Journal of the acoustical society of america 39:151-168.

Padgett, Jaye. 1995. Partial class behavior and nasal place assimilation. In *Proceedings of the South Western Optimality Theory Workshop 1995*. Tuscon, AZ: The University of Arizona Coyote Papers:145-183.
Padgett, Jaye. 2002a. On the characterization of feature classes in phonology. Language 78.1:81-110.
Padgett, Jaye. 2002b. Russian voicing assimilation, final devoicing, and the problem of [v]. Santa Cruz: Ms., University of California, Santa Cruz.
Padgett, Jaye. to appear. Contrast and post-velar fronting in Russian. *Natural language and linguistic theory*.
Prince, Alan. 2001. Invariance under re-ranking: Handout of paper presented at WCCFL 20, University of Southern California.
Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: constraint interaction in generative grammar*: Ms., Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder.
Przezdziecki, Marek. 2000. Vowel-to-vowel coarticulation in Yorùbá: the seeds of ATR vowel harmony. In Proceedings of WCCFL 19, edited by Billery and Lillehaugen. Somerville, MA: Cascadilla Press:385-398.
Pulleyblank, Douglas. 1989. Patterns of feature cooccurrence: the case of nasality. In *Arizona phonology conference 2: coyote working papers 9*. Tuscon, AZ: University of Arizona:98-115.
Rubach, Jerzy. 1984. *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.
Sanders, Nathan. in progress. Opacity and sound change in the Polish lexicon. Doctoral dissertation, University of California, Santa Cruz.
Sanders, Nathan. to appear. Preserving synchronic parallelism: diachrony and opacity in Polish. In CLS 37: The main session. Papers from the 37th meeting of the Chicago Linguistic Society, Vol. 1, edited by M. Andronis, C. Ball, H. Elston and S. Neuvel. Chicago: Chicago Linguistic Society [Also handout of paper presented at the third Generative Linguistics in Poland conference, Warsaw, Poland, 2001.].
Selkirk, Elisabeth. 1994. The constraint question for feature structure: a review of some answers: Course notes, University of Massachusetts, Amherst.
Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*. Amherst, MA: Ph.D. dissertation, University of Massachusetts.
Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at the Rutgers University Optimality Workshop, New Brunswick, NJ.
Smolensky, Paul. 1995. On the internal structure of the constraint component *Con* of UG: Handout of talk given at University of Arizona.
Smolensky, Paul. 1997. Constraint interaction in generative grammar II: local conjunction, or random rules in Universal Grammar: Handout of talk given at Hopkins Optimality Theory Workshop/Maryland Mayfest, Baltimore.
Spaelti, Philip. 1997. *Dimensions of variation in multi-pattern reduplication*: Ph.D. dissertation, University of California, Santa Cruz.
Steriade, Donca. 1997. Phonetics in phonology: the case of laryngeal neutralization. Ms., UCLA.
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Papers in laboratory phonology V: acquisition and the lexicon, edited by M. B. Broe and J. B. Pierrehumbert. Cambridge: Cambridge University Press:313-334.

Suzuki, Keiichiro. 1998. *A typological investigation of dissimilation*. Tuscon: Ph.D. dissertation, University of Arizona.

Walker, Rachel. 1998. *Nasalization, neutral segments, and locality*: Ph.D. dissertation, University of California, Santa Cruz [published by Garland Press, New York, 2000].

Walker, Rachel. 2002. Weak triggers in Veneto metaphony. Los Angeles: Ms., University of Southern California.

Wright, J. T. 1986. The behavior of nasalized vowels in the perceptual vowel space. In *Experimental Phonology*, edited by J. J. Ohala and J. J. Jaeger. Orlando: Academic Press:45-67.

Yip, Moira. 1988. The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19:65-100.

Zhang, Jie. 2000. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: Max[F] or Ident[F]? Phonology 17.3:427-478.

Zoll, Cheryl. 1996. Parsing below the segment in a constraint based framework: PhD Dissertation, University of California, Berkeley.

Zoll, Cheryl. 1997. Conflicting directionality. *Phonology* 14:263-286.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. Cambridge, MA: Ms., MIT.