About this document

This article is a draft of my chapter on “Markedness and Faithfulness Constraints” in the Blackwell’s *Companion to Phonology* (chapter 74).

The published version is 11,000 words long; this version is 14,300 words long. The published version is missing sections 3.4, 4.3, 5, has a shorter section 1.1, and has half as many references.

The title is “Markedness and Faithfulness Constraints”. However, I found dealing with the topic to be so difficult due to its breadth and complexity I decided focus on a very specific issue – what the Constraint Definition Language is (explained inside).

Given the difference between the versions, I have given this one the title “Markedness and Faithfulness Constraints: CDLs”.

- Paul de Lacy.
Tuesday, April 06, 2010

Markedness and Faithfulness Constraints: CDLs[*]

Paul de Lacy

Revised: 5 April 2010
Revised: 15 January 2010
Original: 1 April 2009

1. Markedness and Faithfulness Constraints

Objects and mechanisms called 'constraints' have featured in many theories of the phonological and syntactic modules. However, the explicit bifurcation into 'markedness and faithfulness' constraints is specifically found in Optimality Theory (OT – Prince & Smolensky 1993/2002/2004) and its developments (especially McCarthy & Prince 1995/1997/1999), as well in as theories based on OT (Stochastic OT – Boersma 1998, Boersma & Hayes 2001; Targeted Constraint Theory – Wilson 2001; OT With Candidate Chains – McCarthy 2007; Stratal OT – Kiparsky to appear, Bermúdez-Otero to appear; Harmonic Grammar – Pater 2009b and references cited therein).[1] So, this chapter focuses on the things called 'constraints' in OT (specifically the 'classical OT' of P&S and M&P). In particular, it focuses on OT constraints in the phonological module; there are also OT theories of the syntactic module and OT theories of morphology – they will not be discussed here.

This chapter's aim is to examine the basic syntax and semantics of constraints. On the syntax side: what is the form of constraints? What is the 'constraint construction language'? On the semantics side: how are constraints 'interpreted' – i.e. how are constraints used to assess a candidate's violation marks?

This chapter focuses on the basics of constraints, so it does not aspire to identify every constraint theory or list every constraint and constraint generator (§4.2) that has been proposed; for that, see the ongoing *ConCat* project at http://jones.ling.indiana.edu/~concat/.

An OT constraint is commonly treated as a function that takes a candidate and returns 'violation marks' (see P&S p.75ff). Violation marks are discrete elements; they are usually written as a string of asterisks with one asterisk per unique element (but violation marks in their formal implementation are not necessarily a string). For example, a constraint *DORSAL returns one violation mark for each instance of the representational element [dorsal] in an output representation (constraint names are usually written in small capitals). So for a candidate that includes an output representation [pakak], *DORSAL returns ** because there are two [dorsal] features in the form (one for each [k]). 'Candidates' are sets of forms including at least an output and input representation, but probably many other related representations as well as relations between them (§3.3; P&S p.5).

[*] My thanks to Marc van Oostendorp, Keren Rice, Matt Wolf, and an anonymous reviewer for their extensive comments.

[1] I will refer to Prince & Smolensky (2002) as P&S below; it is available for free on Rutgers Optimality Archive (http://roa.rutgers.edu/) #537. McCarthy & Prince (1995) will be referred to as M&P; it is Rutgers Optimality Archive #60.

Constraints fit into the overall phonological system thus (following P&S). An input is drawn from the lexicon; the input consists of phonological material with morphological and syntactic annotation/structure. A generation mechanism (GEN) produces many (perhaps an infinite number of) candidates. One or more candidates is selected from the array; the selection process (EVAL) involves constraints generating violation marks for candidates and an algorithm that uses the violation marks and other factors to determine the winning candidate(s). EVAL refers to 'ranking' – a total order on constraints. Ranking does not influence how violation marks are calculated; however, ranking is crucial in discovering the winning candidate. The Phonetic module then takes (one of) the winner(s) and realizes it (i.e., converts it into articulatory movements that produce speech sound).

In short, Constraints are just one part of many mechanisms that work together to determine the winning output representation. Constraints do not determine winners on their own.

The term 'markedness and faithfulness' as applied to constraints was coined in P&S§1.4. 'Markedness constraints' return violation marks based solely on the form of the output representation. *DORSAL above is a markedness constraint. Unfortunately, the term 'markedness' can cause confusion because it seems to imply an inherent connection to theories of Markedness (see ch.73 [HUME–MARKEDNESS]). However, theories of Markedness are expressed in OT via both markedness and faithfulness constraints (e.g. de Lacy 2006). The term 'output constraint' is therefore less confusing than 'markedness constraint' so I will use it here. However, the phrase 'markedness constraint' is in such widespread use that I fear 'output constraint' will never catch on (in spite of my efforts in this chapter) (to add to the confusion there are constraints called 'output-output constraints', which are actually faithfulness constraints (see §3.3)).

As originally used, 'faithfulness constraints' are those that return violation marks based on comparison of the output representation with the input (P&S§1.2; though strictly speaking the output *included* the input, discussed in §3.1). Later work, especially M&P, broadened the term to include any constraint that assigned violations by comparing any pair of inter- or intra-representational forms (e.g., the base of reduplication and the reduplicant – M&P; the derivational base and the output – Benua 1997; a designated form and the output – McCarthy 1999). The majority of work in OT now uses M&P's Correspondence Theory, so in these cases it is accurate to refer to 'Correspondence constraints' – i.e. those that use Correspondence relations in their calculation of violation marks. However, non-correspondence faithfulness constraints exist in some versions of OT (e.g. Containment theories – §3.1), so 'faithfulness constraints' is still a usefully broad term.

This chapter focuses on a few important issues about constraints. Section 2 discusses constraint form in output constraints: what are constraints made of, and how do they return violation marks? Section 3 deals with faithfulness and Correspondence constraints. Section 4 discusses the source of constraints – whether they are innate and how/whether they relate to external sources. Section 5 discusses how theories of constraints influence and are influenced by other mechanisms in the phonological module.

1.1 "Constraint" in other theories

The term 'constraint' is used in many different ways in many different theories. In some rule-based theories there are objects called 'constraints' or 'filters' that − if their conditions are met − doom the derivation or block rules from applying. If an input I undergoes a series of rules to create a representation ϕ and there is a filter *ϕ, the derivation is doomed (i.e. input I has no corresponding output). See Chomsky & Lasnik (1977) for early examples in syntax, and Ito (1986) for conditions on syllable structure. The filter/condition concept does not have a direct analogue in OT − OT constraints assign violations; EVAL is the source of (relative) doom for candidates.

The term 'constraint' is also met in 'morpheme structure constraints' (Chomsky & Halle 1968:171ff − also called 'morpheme structure conditions' or 'lexical redundancy rules'). These constraints limit the form of phonological structures stored in lexical entries. In P&S's OT, there are no constraints on lexical items. Phonological regularities across lexical items are treated as a side-effect of the learning process in Tesar & Smolensky (1996) (see Tesar 2007 for references).

Occasionally 'constraint' is used to refer to side-effects of conditions on representational primitives and to restrictions on the algorithm that generates output candidates (GEN). Obviously, output representations can only be constructed out of objects and relations that are available (i.e. prosodic nodes, features, planes, tiers; precedence, dominance). For example, there is no candidate in which a node can both precede and follow another node because the phonological precedence relation is asymmetric (i.e. if *a* and *b* are on the same tier and $a<b$ (i.e. *a* precedes *b*) and $b<a$ then $a=b$) (see chapter 101 CAIRNS for discussion of precedence). One could informally call the asymmetry of phonological precedence a 'constraint', but it is not an OT constraint.

As another example, many authors writing in metrical theory have assumed that there is no output structure in which a Ft node dominates more than one σ node (ie, no 'ternary' or 'unbounded' feet). This restriction is not an OT constraint, but rather a condition on the candidate generation algorithm GEN.

2. Output Constraints

A constraint takes a candidate and returns violation marks.[2] For example, the constraint *DORSAL returns a violation mark for [ka], two for [kax], and so on. A constraint's violation assignment can be described in informal terms: e.g. "*k returns a violation mark for each [dorsal] segment". This informal description is useful, but far from being a formal definition. A formal definition of a constraint must be couched in a Constraint Definition Language (CDL). A comprehensive CDL specifies representational primitives and relations and restrictions on their combination in constraints.

[2] It is common to see comments like "constraints impose a partial order on the candidate set" (Samek-Lodovici & Prince 1999:9), "this constraint dooms the candidate", and so on. These comments are meant as a quick way of describing the complex process of determining order among candidates; the process involves constraints, VR (§2.2), ranking, and EVAL's mechanisms; constraints are merely one part of the process of establishing the winning candidate.

The same distinction can be made for rule-based theories like Chomsky & Halle (1968). Suppose we observe a rule R that takes an input representation /ak/ and converts it to the representation [aʔ]. R could be described as "change /k/ into a [ʔ] word-finally". However, R must be *defined* in terms of a Rule Definition Language (RDL); a RDL is the elements and relations that can be used to construct a rule and limits on their combination. Most of Chomsky & Halle (1968) is devoted to developing such an RDL; R is defined as /k/→[ʔ]/_[−seg,−FB,+WB].

Some work in OT uses informal descriptions to talk about constraints. Often formal objects (representational elements like prosodic nodes, features, etc.) are mentioned in the informal descriptions, but the constraints are nevertheless not defined in terms of an overarching CDL. There have been attempts to develop a comprehensive CDL (Eisner 1997a,b,c, Potts & Pullum 2002), but most work has either focused on particular groups of constraints, or treated constraints as 'black boxes'.

To explain, it is possible to fruitfully investigate some (perhaps many) aspects of OT theories without knowing the precise definition of constraints, but only knowing which violation marks constraints assign in which situations. After all, the winning candidate is not directly determined by constraints, but by their violation marks. So, if the violation marks are known then the winner can be determined – the exact means by which the violation marks came to be assigned is often not crucial. For this reason I believe it is fair to say that there has been less focus on developing a CDL in OT theories on developing RDLs in rule-based theories like Chomsky & Halle (1968).

Even so, there have been detailed proposals of CDLs for groups of OT constraints (see section 4 below), and some proposals about aspects of the general CDL. In my own (co-)work, for example, Bye & de Lacy (2000) propose restrictions on how constraints can refer to constituent edges; de Lacy (2006) proposes that constraints cannot include both prosodic nodes and segmental features in their definitions. Work on connectives in constraint definitions includes Crowhurst & Hewitt (1997), Balari et al. (2000), and Wolf (2007).

An explicit CDL is both useful and ultimately essential to a complete OT theory. A CDL can tell us which constraint formulations are valid, and thus set a bound on which constraints can and cannot exist.

The following subsections will discuss a CDL. There is a strong uniformity in constraint descriptions and definitions that suggests broad agreement about certain aspects of the CDL. For expository reasons, I will start with the CDL for output constraints. As a word of warning, due to space limitations I will discuss only a few CDLs, and focus on the basic components of just one. The CDL discussed below deals with a broad set of output constraints that I believe every phonologist would accept as possible constraints. I will not attempt to comprehensively discuss all extant CDLs or aspects of CDLs, but instead focus on basic CDL properties. For a specific explicit and recent CDL, see Riggle (2004).

2.1 Output constraints: Representation

There is not a lot of explicit discussion about how constraints work in OT. It seems to me that the majority of work in OT treats constraints as functions from candidates to violation marks. Output constraints inspect the output representation in a candidate, and return a string of violation marks. So, *DORSAL returns one violation for candidates with an output representation [ka], two for [kag], three for [gaxikan] and so on, leading to the description "Assign one violation for each dorsal segment."

We seek a CDL in which *DORSAL can be formulated. One issue to address is the CDL's representational primitives. For example, *DORSAL might be cast in terms of a representational theory in which velar consonants are [+back,+high] (Chomsky & Halle 1968:303), or in one in which a Place node dominates a dorsal node which dominates [+back] and [+high] terminal nodes (Sagey 1986), or one in which an oral cavity node dominates a C-place node which dominates a dorsal terminal node (Clements & Hume 1995) (see Hall 2007b for an overview of feature theories). There are many extant representational theories, but for the purposes of this chapter I will adopt Clements & Hume's model (there is no widespread consensus on which representational theory is correct, though; even less now than ten years ago, I suspect).

The CDL also must specify how representational elements can be combined in constraints. For example, a relatively lax CDL could allow several different versions of *DORSAL, as in (1). I use ♦ to stand for 'root node' – the lowest node that dominates all segmental features. The symbol ↓ stands for the immediate dominance relation: a↓b means *a immediately dominates b* (i.e. a dominates b and there is no c such that a dominates c and c dominates b); a↓b↓c means *a immediately dominates b and b immediately dominates c*. Dominance is an asymmetric transitive relation that holds between nodes on different autosegmental tiers. The descriptions are given in (1a), (b), (c); the violation marks the constraint assigns are shown for [k], [kː] (assuming a one-root geminate theory – Hyman 1985, cf. Selkirk 1991), and [ŋk] (assuming obligatory feature sharing for adjacent elements – Schein & Steriade 1986).

(1) *DORSAL *versions*

Return a violation for…	[k]	[kː]	[ŋk]
(a) each distinct root node ♦ s.t. ♦↓CPlace↓[dorsal]	*	*	* *
(b) each distinct [dorsal] feature	*	*	*
(c) each distinct prosodic node *p* s.t. p↓♦↓CPlace↓[dorsal]	*	* *	* *[3]

I do not know which of the constraints (1a), (1b), and (1c) exist. Suppose it turns out that we need only (1a). The non-existence of (1b) and (1c) could be achieved by putting restrictions on the CDL so that all output constraints must refer to a root node in their definition and no constraint may mention both prosodic nodes and segmental features. Every extant theory of representation provides a CDL with a great deal of potentially expressive power. So, it is highly likely that any CDL theory will have to incorporate extensive limitations on permissible representations in constraints; it is probably too hopeful that all limits on constraints will be a side-effect of inherent limitations in representations (see §4).

[3] (1c) could return one violation if [ŋ] and [k] were both dominated by the same μ or σ node (e.g. as in $[oiŋk]_{\sigma}$).

The CDL must also specify how the representation is used to assess violation marks. For example, suppose a constraint mentions the structure [•↓CPlace↓[dorsal]]. How is this structure used to assess violation marks relative to some candidate? In the constraint description (1a), I assumed that the constraint searches the candidate's output representation and for each distinct structure that has the form [•↓CPlace↓[dorsal]], one violation mark is returned. However, could there be a constraint which returns one violation regardless of how many [•↓CPlace↓[dorsal]] structures there were? Such a constraint would return * for [ka], [kax], and [kaxga]. Let's turn to this issue now.

2.2 Function or representation?

I asserted without comment above that a constraint is a function: i.e. it takes a candidate as an input and returns violation marks. Conceiving of constraints as independent functions opens up the possibility that different constraints could assign violation marks in very different ways.

For example, the ALIGN schema from McCarthy & Prince (1993a) takes four arguments and assesses violation marks with respect to designated prosodic constituents. The violation marks from ALIGN(Ft,R,PrWd,R) are the sum of the number of syllables between the right edge of each foot and the right edge of the PrWd. So, the constraint returns 9 violations for [(σσ)(σσ)(σσ)σ] (see McCarthy & Prince 1995:15,16 for details; to understand this constraint see esp. p.10 and definitions 14, 15, 16). It is clear that the way in which this ALIGN constraint assesses violation marks is quite different from the way in which violations of *DORSAL are assessed. (cf McCarthy 2003).

If constraints are self-contained algorithms that return violation marks and the CDL is sufficiently powerful, we might see pairs of constraints that refer to the same representational structure but differ in how they calculate violation marks. For example, there could be a pair of constraints *∀DORSAL and *∃DORSAL, where *∀DORSAL returned as many violation marks as there are [dorsal] features in an output representation, but *∃DORSAL returned only one violation mark regardless of how many [dorsal] features there are in a form, as long as there is at least one (for relevant discussion, see Wolf 2007). The two constraints refer to the same structure – [•↓CPlace↓[dorsal]] – and differ only in terms of how that structure is used to assess violation marks from a candidate. A pair of constraints like this – i.e. that refer to the same representational structure but differ only in their quantification – would be strong evidence that each constraint is an independent algorithm that assigns violations (or at least that there are several groups of constraints that differ in how they assign violations).

However, my impression is that the constraints-as-functions approach is too powerful. The output constraints that have been proposed in phonological literature are often very similar: they essentially have the form *R, where R is a representation; one violation mark is assigned for each distinct occurrence of R in a candidate's output representation. *DORSAL is an example of such a constraint. The apparent uniformity in how constraints assess violations suggests that it is worthwhile considering an alternative theory of constraints in which constraints are not functions but solely representations.

In such an approach, there would be a single algorithm VR (for Violation assigneR). VR takes as its input an output constraint and a candidate and returns violation marks. VR works the same way for all constraints, so imposing uniformity in how violation marks are assigned. So, the constraint *DORSAL is really a representation [•↓CPlace↓[dorsal]]; *DORSAL itself does not assess violation marks.

There are many ways to formulate a VR algorithm that does the job described above. For example, one could take the set of all subrepresentations of a candidate's output representation and compare each member of the set to a constraint representation; the number of violation marks returned for a particular constraint would be the number of subsets that were equivalent to the constraint's representation. I will instead discuss a somewhat more efficient algorithm that does a similar job. See http://www.pauldelacy.net/VR for software which allows you to see the VR below in action and try out various constraints and representations.

(2) *Violation AssigneR (VR): Outline*

Inputs: • A constraint C

• A candidate that includes an output representation R

Output: A set of violation marks (a set of unique identifiers) indexed to C and the candidate.

(a) Take a node c in C

(b) For each node r in R

If r is the same type and value as c

then check whether r is connected to a structure equivalent to C

• If it is, return a violation mark, and continue to the next r

(c) Add no other violation marks to the Result.

For example, take a constraint *[+voice] which consists of one [voice] node with a value of '+'. Each node in the candidate's output representation is checked. If a node is a [voice] node and has the value '+', a violation mark is added to the result.

The VR might seem straightforward but it has interesting complexities, particularly in the procedure that checks whether a node is "connected to a structure equivalent to [the constraint] C".

Take a more complex constraint – one that involves two nodes: e.g. $*\sigma_{\mu\mu}$ "Don't have bimoraic syllables." $*\sigma_{\mu\mu}$ has three nodes (σ, μ_1, μ_2) and three relations ($\sigma\downarrow\mu_1$, $\sigma\downarrow\mu_2$, $\mu_1<\mu_2$). A node is selected from the constraint (it doesn't matter which one) – let's say σ in this case. The output representation is searched for σ nodes. When one is found, the next step is to check whether σ is connected to a structure equivalent to the constraint. The implementation of this checking procedure is that σ is checked to see if it is in any of the *relations* mentioned by the constraint: i.e. does the particular σ in the representation dominate two different μ nodes? If it does, then the μ nodes that are dominated by σ are then checked to see whether their relations have equivalents in the constraint. After nodes and relations are found in the output representation that are equivalent to those in the constraint, a violation mark is returned.

The procedure that checks whether *n* is connected to a structure equivalent to C means that constraints cannot be unconnected. For example, suppose that there is a constraint $*\mu_1,\mu_2$ which is violated if a word contains two (not necessarily adjacent) moras; these moras are unconnected in this constraint – there is no precedence relation between them, nor is there a node that dominates them both. VR can evaluate such a constraint, but the constraint will never return any violation marks. VR checks *relations* between nodes: i.e. VR will search for a μ_1 node, and then check its relations. Since μ_1 has no relations (i.e. no precedence relation between μ_1 and μ_2). At this point, though, no structure in R has been found that is equivalent to the structure described by the constraint, so no violation marks are returned.

So, the VR algorithm itself, through how it compares the constraint's structure to structures in the output representation, imposes a weak connectedness requirement on constraints. For a constraint C to ever return a violation mark, every node in C must be connected to every other node. Nodes *x* and *y* are 'connected' here if it is possible to trace a direct route through precedence and dominance relations from *x* to *y*.[4]

The connectedness requirement that results from VR is *weak* because a much stronger requirement could be imagined and implemented: i.e. in a constraint, every pair of nodes on the same tier must be in a precedence relation, and every pair of nodes on different tiers must be in a dominance relation. The difference between weak and strong connectedness can be seen in a less-connected version of $*\sigma_{\mu\mu}$: where $\sigma{\downarrow}\mu_1$ and $\sigma{\downarrow}\mu_2$, but there is no precedence relation between μ_1 and μ_2. Such a constraint would not be evaluated by a VR that imposes strong connectedness because it has two nodes on the same tier (μ_1,μ_2) that are not in a precedence relation. However, it is perfectly acceptable in the weak-connectedness VR described above because every node is connected to every other node.

Weak connectedness allows constraints of the form $*[x\ldots x]_D$, where there are two nodes of type *x* within a particular domain D (e.g. Suzuki 1998; also 'local conjunction' – §4.2).

The connectedness side-effect of VR is desirable – as far as I am aware, no-one has proposed constraints that have completely unconnected elements.

The larger point here is that the nature of the algorithm that assigns violation marks is crucial in any theory of constraints. The algorithm not only determines how violation marks are assigned, but whether particular constraints will ever assign violation marks (i.e. it effectively puts restrictions on constraint form just as VR means that constraints must contain connected representations).

If constraints are representations and there is a single VR, there should be broad uniformity in the way that violation marks are assigned. For example, the constraint [•↓CPlace↓[dorsal]] will assign violations for each occurrence of its representation in the candidate's output representation. In contrast, there is no way to formulate a constraint like *∃dorsal: if the constraint consists of the representation [•↓CPlace↓[dorsal]], then the VR will return a violation for *each* occurrence of [•↓CPlace↓[dorsal]]; it cannot be limited to assigning one violation regardless of the number of occurrences of [•↓CPlace↓[dorsal]].

So, the VR theory means that output constraints should all assign violations in fundamentally the same way, while the constraints-as-functions theory allows for significant differences. It is even possible that there is a middle ground: there could be

[4] Let's define a transitive symmetric relation ⊕; $x \oplus y$ if $x<y$ or $x{\downarrow}y$. Nodes *x* and *y* are connected if $x \oplus y$. For example, in the constraint $*\{\sigma{\downarrow}\mu_1, \sigma{\downarrow}\mu_2\}$, μ_1 and μ_2 are connected because $\sigma \oplus \mu_1$ and $\sigma \oplus \mu_2$, so $\mu_1 \oplus \sigma$ (by symmetry) and $\mu_1 \oplus \mu_2$. There is no need for a direct implementation of ⊕ in VR; connectedness follows as a side-effect of the procedure that matches the constraint to the representation.

several violation assignment algorithms, with VR being just one of them. With several VRs, we would expect to see uniformity in how violations are assigned but only within particular groups of constraints.

So, which view is correct? How much uniformity in violation mark assignment is there?

2.3 Regularities and irregularities in violation assignment

There is a great deal of regularity in the way that violation marks are assessed in output constraints. In fact, there is so much regularity that I am sure there would be no confusion among phonologists about how a constraint like *ʃɾ works (even though I concocted it just now): it would return a violation mark for each [ʃ] immediately followed by [ɾ] in the output representation. There is an implicit ‘quantificational’ regularity here: one violation mark is returned for *each* distinct [ʃɾ] (i.e. [ʃɾaʃɾi] returns two violations), not for just having some [ʃɾ] (where [ʃɾaʃɾi] returns just one violation).

As mentioned above, such ‘quantification’ regularity is expressed straightforwardly in VR. The process of iterating through every node in the representation means that a violation will be returned *each time* a node is found that is part of a representation that matches the constraint’s.

There are other widespread similarities in constraints. Many output constraints refer to tier-contiguous representations. For example, a constraint like *ŋk refers to *adjacent* root nodes; *ŋk returns one violation for [aŋko], but [aŋok] gets none (ie, *ŋk is ŋ<k, where ‘<’ is used here to mean ‘immediately precedes’). I have not yet seen a constraint *ŋ…%…k proposed which is violated whenever [ŋ] and [k] appear in the same output in any order (i.e. one violation for each of [aŋko], [aŋok], [akŋo], [akoŋ], etc.) (the ‘%’ notation is from Bach 1968).

The particular implementation of VR discussed above has intra-tier-contiguity as a side-effect of the way that the algorithm works. When a node is encountered in the output representation, the node’s relations (i.e. precedence and dominance) are checked against the constraint’s. If the node is in a precedence or dominance relation with some other node, then that other node’s relations are then checked, and so on. This method of checking means that only nodes that are connected in the representation will be examined. So, for a constraint *ŋ…%…k, a node [ŋ] might be found in the representation, but ŋ does not precede any other node so [k] will never be found by the VR procedure, and the constraint’s representation will never be matched.

Some authors have proposed constraints with weakly contiguous representations to deal with processes like long-distance assimilation and dissimilation (e.g. Alderete 1997, Suzuki 1988). These constraints refer to segmental nodes that are not in a precedence relation, but are related by dominance (i.e. with the form $*[x\ldots\%\ldots y]_D$, where x and y are nodes dominated by constituent D. Such constraints always require or ban identical elements (not non-identical ones, as above), leading some recent work to recast such processes as involving local Correspondence relations (e.g. Rose & Walker 2004) or to restrict the domain of such constraints so that the representations they refer to are mostly always tier-contiguous (e.g. Łubowicz 2005). Such a move might mean that VR should implement a strong contiguity requirement, where within a constraint every node on a particular tier is in a precedence relation with every other node on that tier.

There is a widely accepted constraint type that does not always easily fit with VR: 'positive' constraints. A positive constraint penalizes a representation for *lacking* a particular property, while negative constraints like *DORSAL return violations for *having* a particular structure. A well-known constraint that is often positively formulated is ONSET: "Syllables must have onsets" (see other definitions of this constraint, including negative ones, at *Concat* http://tinyurl.com/concat-onset). Another example is PLACE→DORSAL "Every Place node must dominate a [dorsal] node". Because positive constraints do not mention a representation but rather the *lack* of a representation, they cannot be fed into the VR. That is to say, PLACE→DORSAL does not mention the representation [♦↓Cplace↓[dorsal]]. Instead, it requires that for every root node ♦ there must be some member of the set of all subrepresentations of a candidate that includes ♦ and has the structure [♦↓Cplace↓[dorsal]]. It is not hard to formulate a VR algorithm that can handle positive constraints (VR^{POS}), but such an algorithm is different from the one that handles negative constraints (VR^{NEG}).

Are both VR^{NEG} and VR^{POS} necessary? Many positively formulated constraints can be reformulated negatively, and many negative ones can be reformulated positively. However, the success of reformulation depends entirely on the particular CDL used. For example, within fairly standard representational theories it is not possible to reformulate *DORSAL as a positive constraint or even a set of positive constraints without resulting in the 'pile-up problem' (de Lacy 2002:90ff, 2006:ch.2; cf. Yip 2001). Similarly, the success of reformulating positive constraints as negative ones depends on the CDL's representational primitives and restrictions. For example, can ONSET be reformulated as $*[_\sigma V$ "Return a violation mark for each vowel at the left edge of a syllable" (Kager 1999)? Yes, but *only* if the CDL allows reference to syllable edges in this way and the VR can process it. Whether such boundary reference is possible depends on entirely on the CDL and VR.

There are a variety of popular positive constraints that seem to defy easy negative reformulation; e.g. FTBIN "Assign a violation for each Foot node that does not dominate exactly two moras or two syllables" (after McCarthy & Prince 1986, P&S). Elías (2006) proposes breaking the constraint down into constraints that put upper and lower bounds on moraic content. A constraint against more than two sub-foot elements is straightforward using VR: $*\{Ft_i{\downarrow}x_a, Ft_i{\downarrow}x_b, Ft_i{\downarrow}x_c\}$ where x_a, x_b, x_c are all distinct prosodic nodes (i.e. μ or σ) and ↓ here is *dominance*, not *immediate dominance*. A constraint against having *fewer* than two sub-foot daughter nodes is more difficult: how do we penalize a mono-moraic/-syllabic foot structure using a representation ($Ft{\downarrow}x$) and the VR without also penalizing larger structures? Perhaps the VR could be tweaked, making it seek out complete constituents in evaluating constraints. Or it may be possible to capitalize on boundaries, banning a sub-foot prosodic node that is both a left and right foot boundary (e.g. $*[_{Ft}x]_{Ft}$), depending on how boundary-reference is permitted in the CDL.

The literature is full of examples of positively and negatively formulated constraints. If there is a single VR, a big challenge is to figure out how to recast the positive constraints into negative ones (or vice versa, depending on one's VR), or alternatively identify sets of negative constraints that take over the effect of the positive ones.

The VR has other interesting side-effects, such as its inability to deal with ALIGN-type constraints and 'gradient' constraints generally (McCarthy 2003).

So, which option is right? Is there a single VR algorithm and output constraints are really representations? Or is every constraint its own violation mark calculation algorithm, opening the possibility that constraints could assign violation marks in wildly different fashions? The fact that there is such widespread similarity in how constraints work makes me hope for the former option. However, there is no doubt that recasting every extant output constraint in VR-friendly terms is extremely challenging. It is also quite possible that there are several VR-like algorithms, and that constraints are indexed for which violation calculation algorithm they undergo.

The issues become more complicated when we consider the other major set of constraints – Correspondence constraints.

3. Faithfulness Constraints

Output Constraints refer to the output representation in candidates when violation marks are calculated. However, if only output constraints exist, each grammar would produce the same output regardless of what the input is.[5] So, constraints are needed to regulate inter- and intra-representational relations. For example, there must be constraints that keep outputs from being altered relative to the input form (see McCarthy 2007 for discussion). There is after all a significant difference between [ʔaki] and [ki] in terms of how they relate to the input /aki/. How are constraints used to decide which one of these competitors wins? In more common parlance, how do constraints regulate the *faithfulness* of the output to the input?

3.1 Containment and Correspondence

Prince & Smolensky (1993/2002/2004) proposed a theory – Containment – that expresses differences in the output relative to the input in terms of output symbols. In the process of generating output representations from an input, a segment can be marked as unparsed (e.g. <k>) which means that it will not be phonetically realized. A segment can also be epenthesized – added to the output – and marked as such (e.g. <u>k</u>).[6]

Containment theory provides a straightforward way to ban deletion and epenthesis. Constraints against deletion are simply output constraints that ban unparsed

---

[5] Chomsky (1995) suggests with only output constraints, [ba] would always emerge as the winner. There are actually *many* possible outputs because different rankings of markedness constraints favor different output configurations, among them [ʔáʔə], [ʔáː], [ʔáʔa], [tátə], [táː], and [táʔ]. [ba] could never be the least marked output in many theories of Markedness because of its subminimal size and voiced labial consonant (see de Lacy 2006 and Rice 2007 for overviews of Place Markedness in OT), and Kager (2007) on prosodic markedness.)

[6] P&S adopted the theory that epenthetic segments are empty prosodic positions. However, this proposal is no longer widely accepted; I have updated the description of Containment theory accordingly.

segments: *<♦> bans deletion of root nodes, and *♦̲ bans epenthesis. Such constraints fit straightforwardly into the VR algorithm. More recent Containment theories can be found in Goldrick (2000) and van Oostendorp (2007).

However, McCarthy & Prince (1995/1997/1999) argue for a very different approach. Their proposal – Correspondence Theory (CT) – is now the most widely used theory of inter- and intra-representational identity.

CT proposes a relation – ‘Correspondence’ – that holds between root nodes (at least, see §3.2). For example, the root node /k/ in the input /aki/ can be in a Correspondence relation with the root node [k] in the output [ʔaki]. However, in the process of making output candidates from an input, GEN has a free hand in generating Correspondence relations. So, there will also be a candidate consisting of the input /aki/ and even an output [ʔaki] where /k/ corresponds to [ʔ]. Correspondence relations are often written with subscript numerals when it is not obvious which elements correspond (e.g. /a$_1$k$_2$i$_3$/: [ʔa$_1$k$_2$i$_3$] vs. [ʔ$_2$a$_1$k$_2$i$_3$]).

With a free assignment of Correspondence relations, there are many candidates consisting of the input /aki/ and output [ʔaki] where the only difference is Correspondence. However, due to the way Correspondence constraints work the vast majority of such candidates will never win under any ranking.

Constraints that regulate the Correspondence relations between input and output root nodes are responsible for keeping outputs looking like inputs. For example, IO-MAX returns a violation for each input root node that is not in a Correspondence relation with some output root node. So, candidates consist of (at least) an input representation, an output representation, and the Correspondence relations that hold between them.

Correspondence was a new concept in phonological theory. In SPE and its rule-based successors, input-output faithfulness is an epiphenomenon of rule non-application; an output form is perfectly faithful to its input if no rules apply to it. The more rules that apply, the more likely the output is to become less faithful (though not necessarily: later rules could undo the effect of preceding rules; see Pullum 1976). However, at any point in the derivation of a rule-based theory except for the first rule application, the original input is *not accessible*; the only accessible representation at any point for a rule theory is the current one (i.e. the input to the rule). In contrast, for every candidate in CT, the input representation is always directly accessible via Correspondence: constraints regulate how the output candidate fares relative to the input via Correspondence relations.

McCarthy & Prince (1995) proposed a framework of Correspondence constraints that has remained substantially unchanged in most subsequent work (see §3.3); see their Appendix A for detailed formulations. I discuss a select few here.

IO-MAX returns a violation mark for every input root node that has no output correspondent. So, IO-MAX returns a violation for /p$_1$a$_2$n$_3$/→[p$_1$a$_2$] because /n/ has no output correspondent; it returns no violation for /p$_1$a$_2$n$_3$/→[pã$_{2,3}$] because /n/ corresponds to [ã]. Notice that Correspondence is not a function; an output segment can correspond to more than one input segment, and vice versa.

IO-DEP returns a violation for every output root node that has no input correspondent; IO-DEP is violated by epenthesis: one violation for /i$_1$/→[ʔi$_1$], but none for /i$_1$/ → [j$_1$i$_1$].

IO-IDENT[F] regulates feature change; it returns a violation for each input root node whose output correspondent does not have the same value for feature F. There are probably individual IO-IDENT[F] constraints for every feature, and possibly for groups of features too (see §4.3). For example, IO-IDENT[continuant] is violated once in /k_1a/→[x_1a].

Other Correspondence constraints regulate multiple Correspondence (UNIFORMITY – preventing coalescence, INTEGRITY – preventing breaking/diphthongization/split), preservation of precedence and adjacency (CONTIGUITY, LINEARITY), and preservation of edge proximity (ANCHOR).

A caveat is in order here. It's common in informal discussion to say that a constraint *does* some particular task: "MAX bans deletion"; "DEP militates against epenthesis", "UNIFORMITY eschews coalescence", and so on. It is dangerous to take such statements too seriously. MAX does not ban deletion; it simply returns violations if an input root node has no output correspondent. In fact, MAX has effects that do not fit easily into the traditional concept of 'deletion'. For example, in the mapping /$p_1a_2i_3$/→[$p_1e_{2,3}$] there has not been any violation of MAX, though a standard SPE account would say that there has been deletion relative to the input (/pai/→[pei]→[pe]). In the mapping /k_1a_2/→[ʔa_2] it looks like /k/ has changed into [ʔ] – not deleted, yet there is a violation of MAX (in an SPE-type analysis there *is* a step with deletion: /ka/→[a]→[ʔa]). See Gouskova (2008) for detailed discussion of this point in relation to DEP.

Correspondence constraints essentially encourage identity – i.e. preservation of a particular input property in the output, or vice versa. IO-MAX and IO-DEP require identity of the number of root nodes; IO-IDENT[F] requires identity of feature values, and the other constraints preserve precedence and position.

M&P's Correspondence constraints have a simplicity and symmetry that is surprising compared to the complexities of output constraints. For example, IO-DEP is actually IO-MAX with an output-oriented focus (i.e. IO-DEP is just "OI-MAX"). There also may be constraints of the OI-IDENT[F] form; they differ from IO-IDENT[F] constraints in terms of coalescence and breaking (see Pater 1999).

3.2 CDLs for Correspondence

There are many ways to conceive of a CDL for Correspondence constraints. Regardless of the CDL, though, it seems challenging to conceive of how the VR algorithm discussed in section 2.2 could process Correspondence constraints straightforwardly; Correspondence constraints demand identity in a way that is not inherent to VR. It is therefore possible that each faithfulness constraint is its own function. However, there is such an overwhelming regularity in terms of how they behave that it is very tempting to assume that there is one algorithm that calculates violations for Correspondence constraints, even if it is different from VR.

So, the temptation is to set up a separate violation mark calculation algorithm just for Correspondence constraints (C-VR). However, it is also possible to combine Containment and Correspondence so that Correspondence constraints are expressed in terms that the VR can process, and regulation of identify follows from the VR. In fact, this is essentially how P&S's Containment theory worked. Here I will discuss an updated version as an example.

For IO-MAX, suppose that there is a computable function K over Correspondence that is included in each candidate. For every root node in the input, K returns the output root node it corresponds to. If the root node does not correspond to any output root node, K returns a unique phonetically uninterpreted element ψ that is designated as belonging to the output (ψ is unconnected to the output representation, though).[7] Much like the other Containment theories, then IO-MAX is $*\psi^{O}$, effectively incurring a violation for each output ψ returned by K. IO-DEP is $*\psi^{I}$. See http://www.pauldelacy.net/VR for a partial implementation of this approach. More complex constraints like IO-MAX-Vowel, which preserve segments that are vowels in the input, can be expressed as $*\{\blacklozenge_{i}K\psi^{O},\blacklozenge_{i}\downarrow[+\text{vocalic}]\}$.

IO-IDENT[F] poses a greater challenge in a such a quasi-Containment implementation. However, the Containment approach essentially replaces identity requirements with a ban on dissimilarity, and this method can be capitalized on for IDENT. Let us define a relation G so that if a root node α dominates feature F, then any α-corresponding β (including α itself, assuming that Correspondence is reflexive) is in the relation βGF. The structure in (3) shows two corresponding root nodes α and β. It provides their G relations.

(3) Input:	α_1	Output: β_1
	|	|
	[−voice]	[+voice]
	αG[−voice], αG[+voice]	βG[−voice], βG[+voice]

The constraint IO-IDENT[voice] can then be expressed as $*\{\blacklozenge^{I}G[-\text{voice}], \blacklozenge^{I}G[+\text{voice}]\}$ – i.e. a ban on an input segment having different instances of the same feature with different values. The virtue of this approach is that inter-representational identity here is recast as avoidance of dissimilarity, which is what many output constraints strive to do (e.g. assimilation constraints – de Lacy 2002:ch.7, Baković 2007).

However, there have been proposals that support the idea that Correspondence constraints are independent functions (or at least that there are several FVRs). Alderete (2001) and Struijke (2000) propose that there are Correspondence constraints that differ only in terms of how they assess violations in relation to quantification.

Struijke (2000) argues that there are counterparts to the Correspondence constraints that differ solely in terms of 'quantification'. So, while IO-IDENT[nasal] returns a violation for each pair of corresponding input-output segments that have different values for [nasal], IO-∃IDENT[nasal] returns a violation for each input segment for which there is *some* output that has a different value for [nasal]. The difference is seen when an input segment corresponds to more than one output segment. For example, if /$\tilde{a}_1$/ splits to become [a_1n_1], IO-IDENT[nasal] returns one violation because there is an IO pair that disagrees in [nasal]: </ã/,[a]>. However, IO-∃IDENT[nasal] does not return

[7] There is a strong similarity in this approach to Wolf & McCarthy's (2005) 'string-based' correspondence, where the input string /x/ is deleted if /x/ corresponds to an empty string *e* (Correspondence is between strings not root nodes in this view). Here, Correspondence is still between root nodes, with an additional uninterpretable root node ψ thrown in (akin to Containment Theory's ◊ and □).

any violations because for every input segment (i.e. /ã/) there is *some* pair that preserves the [nasal] value (i.e. </ã/, [n]>).

Alderete (2001) argues that at least some faithfulness constraints have 'anti-faithfulness' counterparts. For example, OO-IDENT[voice] returns a violation for each corresponding segment that has different values of [voice] (for the OO- part, see §3.3). However, ¬OO-IDENT[voice] returns a violation if there is no pair of correspondents that disagrees on [voice] values. The constraints differ in terms of how they assess violations rather than the representations they refer to.

If Struijke's (2000) and Alderete's (2001) proposals are correct, they pose a serious challenge to an approach that seeks to find a single F-VR algorithm. Their proposals mean that there are sets of constraints that differ only in terms of the procedure of violation mark assignment, not in the representation and relations they refer to.

3.3 Developments in Correspondence Constraints

The theory of Correspondence constraints has been reduced, altered, and extended since M&P.

For example, Keer (1999) argues that UNIFORMITY does not exist, with the effect that coalescence is obligatory in certain situations. Similarly, a number of authors have argued that DEP does not exist (e.g. Myers 1997, Bernhardt & Stemberger 1998, 2007:592,593; Causley 1999, Urbanczyk 2007). They observe that output constraints do a similar job to IO-DEP; output constraints (usually) prefer less representational structure over more, and so does IO-DEP. However, Gouskova (2008) argues that the effects of IO-DEP can be distinguished from structural constraints.

Most work has focused on extending Correspondence to new nodes (§3.3.1) and dimensions (§3.3.2).

3.3.1 Loci of Correspondence

McCarthy & Prince (1995:14) proposed that Correspondence holds between segments. I have adopted a particular theory of representation (autosegmental theory) in this chapter that does not provide an easy way to define 'segment'; in this theory, the most natural understanding of M&P's proposal is to say that Correspondence holds between 'root nodes'. M&P also suggests that Correspondence could hold between other nodes: tonal nodes, prosodic nodes, and terminal and non-terminal feature nodes.

Myers (1997) develops this idea for tonal nodes. For example, IO-MAX-T requires that every input tone node correspond to some output tone node. The most significant effect of the proposal is that tones can survive even when their segmental sponsors are deleted. See Yip (2007) for an introduction to tone constraints.

McCarthy (2000) argues for a variety of constraints that require (at least some) prosodic nodes to be in Correspondence. Since (most) prosodic structure is apparently absent in inputs, evidence for Correspondence between syllables and feet comes from identity across other forms (e.g. base-reduplicant, base-derivative; see §3.3.2). (Also see McCarthy & Prince 1993b, 2001 on STROLE).

A widely discussed extension has been feature Correspondence (McCarthy & Prince 1995:71, LaMontagne & Rice 1995, Causley 1997, Walker 1997, Zhang 2000, Lombardi 2001). Lombardi (2001) argues that feature-to-feature Correspondence is essential in explaining differences between how place features and voice features differ in their behavior. Causley (1997) argues that it is essential in dealing with coalescence.

However, coalescence can be achieved without featural Correspondence (e.g. Pater 1995, de Lacy 2002:ch.8), and a general concern with MAX-feature approaches is the lack of observed feature autonomy. In several theories, features do not seem to have the same kind of independence as tones: while tones can survive if their sponsors are deleted, there may not be similar effects for features (though see the special cases of featural morphemes – see Akinlabi 1996, Wolf 2005§2.2 for recent discussion).

3.3.2 Dimensions of Correspondence

The discussion above has focused on Correspondence between inputs and outputs. However, there have been many proposals that extend the reach of Correspondence. The proposals fall into two categories: intra-representational Correspondence and inter-representational Correspondence.

M&P proposed that intra-representational Correspondence is found in reduplication. A reduplicant morpheme has no input content, but its output segments can correspond to certain other output segments (the reduplicant's 'base'). For example, one of the reduplicated forms of Māori [parau] 'baffled' is [paɾapaɾau]. The reduplicated segments correspond to other output segments thus: [$p_1a_2ɾ_3a_4p_1a_2ɾ_3a_4u_5$].

M&P argue that constraints on Base-Reduplicant (BR) Correspondence have the same form as constraints on IO Correspondence. BR-MAX requires every base element to have some correspondent in the reduplicant (violated once in [$p_1a_2ɾ_3a_4p_1a_2ɾ_3a_4$u]), BR-DEP requires every reduplicant segment to have a correspondent in the base, and BR-IDENT[F] regulates featural identity between base and reduplicant.

What is surprising about the extension of Correspondence to the Base-Reduplicant dimension is that there is essentially one formal mechanism that accounts for both the input→output relation and reduplication. Other theories of reduplication conceive of the phenomenon as involving templates or a type of long-distance assimilation, perhaps through autosegmental spreading (see e.g. McCarthy & Prince 1986). See Urbanczyk (2007) for an overview of BR-Correspondence and reduplication.

Other intra-representational Correspondence relations have been proposed. Kitto & de Lacy (1999) argue that epenthetic segments can correspond to other output segments, resulting in 'copy epenthesis': e.g. Winnebago [boːp<u>ũ</u>nũs] 'hit at random' (Miner 1992). The reason for proposing Correspondence here is 'overapplication': nasal vowels only occur after nasal consonants in Winnebago, except when epenthetic vowels copy a post-nasal vowel, as above. Such 'overapplication' is expected with Correspondence, since featural identity of corresponding elements can trump phonotactic restrictions; and is also found in reduplication and other types of Correspondence (see Urbanczyk 2007 for discussion of under- and over-application in reduplication; and Benua 1997 for output-output Correspondence).

Hansson (2001) and Rose & Walker (2004) go further in arguing that any output segment can correspond to another output segment. The effect is seen in long-distance agreement. For example, in Chumash sibilants agree in anteriority within a word: /s-ilakʃ/ → [ʃilakʃ] ‘it is soft’ cf. [s-ixut] ‘it burns’.

There have been many proposals about inter-representational Correspondence, too. Benua (1997) proposes that segments in the output representation can correspond to segments in the ‘trans-derivational base’ of that output. The trans-derivational base of a word is basically the word minus its structurally outermost affix. So, the base of *original* [[origin]al] is *origin*. *Original* itself is the base of *originality*, and *origin* is also the base of *originate*.

OO-Correspondence is argued to explain why some morphologically complex words do not follow expected phonological patterns, but instead remain similar to their base. For example, in my idiolect (and in many other English-based idiolects) the head foot avoids final syllables in nouns, but otherwise is drawn to the right edge of a PrWd: [əd(mə́sə)bʊ] *admissible*, [ədməsə(bə́lə)ɾi] *admissibility* (incidentally, [ə] can be stressed in my dialect, and /l/→[ʊ] outside onsets). However, with some affixes the foot does not get drawn rightward as expected: [əd(mə́sə)bʊnəs] *admissibleness*, *[ədmə(sə́bʊ)nəs]. When *ness* appears in a word, it subjects the candidate to an OO-faithfulness requirement that has the effect of forcing the correspondent of the base’s head syllable to also be a head. So, *[ədmə(sə́bʊ)nəs] loses to [əd(mə́sə)bʊnəs] because the corresponding head syllable in the base [əd(mə́sə)bʊ] is [mə], not [sə].

Further work on inter-word relationships has argued that candidates should consist of entire output paradigms of related word-forms. See McCarthy (2005) for references and discussion (cf. Bobaljik 2008).

Yet other work has proposed Correspondence relations between the output representation and another output representation that is identified by a special selection mechanism, with the aim being to account for phonological opacity (see ch.86 [BAKOVIC]). See McCarthy (1998, 1999), Jun (1999), and Bye (2002) for discussion.

3.4 Mixing Correspondence and Output constraints

Some constraints have been proposed that mix output and Correspondence conditions. The constraints penalize output structures if there is some kind of unfaithfulness.

For example, Archangeli & Suzuki (1997) propose a constraint that returns a violation mark for each round vowel in the output that differs in height from its input counterpart (i.e. /a/→[u] returns a violation mark but /i/→[u] does not).

Łubowicz (2002) also argues that there are constraints that returns violation marks for particular output forms, but only if they are crucially unfaithful in some way. Specifically, there is a constraint that returns a violation mark for the mapping /g/→[d͡ʒ], but not for /d͡ʒ/→[d͡ʒ]. The constraint cannot simply be a faithfulness constraint because it is not violated in the mapping /g/→[ʒ]. Łubowicz (2002) uses a mechanism called ‘local conjunction’ to construct the combined output and Correspondence conditions (see section 4.3). For example, [*ǰ & IDENT(coronal)]$_{Seg}$ returns a violation for a segment that is both [d͡ʒ] (a voiced postalveolar affricate) and differs from its input in its specification for coronal place of articulation.

McCarthy (2002) proposes a class of constraints that mix output conditions and Correspondence relations (also see Baković 1999). For example, $_{\text{N}}$NoVcdOb returns a violation for each output voiced obstruent that does not correspond to a voiced obstruent. So $_{\text{N}}$NoVcdOb returns one violation for /aka/→[aga], but none for /aga/→[aga]. The idea behind this class of 'comparative markedness' constraints is that there are constraints that penalize particular output structures only if those output structures are 'new' – i.e. they were not in the input form. To be a bit more precise, the constraints do not refer to Correspondence relations between the output and input but between the output and the 'fully faithful form'; see McCarthy (2002§6.2) for details.

Łubowicz (2002) and McCarthy (2003) propose that their constraints are needed to deal with derived environment effects, and comparative markedness is also argued to account for counterfeeding opacity, 'grandfather effects' (blocking a structure S unless S is faithful to the underlying form), and a diverse array of other phenomena (see McCarthy 2003§4.3 for discussion of Łubowicz's (2002) local conjunction theory, see Łubowicz (2005) for a critique of Comparative Markedness).

There is some controversy over whether such constraints are really necessary. de Lacy (2006§8.6) argues that Comparative Markedness constraints make incorrect predictions for grandfather effects; also see Blumenfeld (2003b), Hall (2006, 2007) and the commentaries on comparative markedness in *Theoretical Linguistics 29*. The local conjunction approach is criticized by Inkelas (1999§2.3); also see Blaho (2003), Blumenfeld (2003a), and Wolf (2008).

4. Possible and impossible constraints

Even after defining the CDL's representational elements, relations, and violation mark assignment algorithm(s), there remains the question of which constraints actually exist.

I wish I could list all the phonological constraints that exist in the human brain here. Unfortunately, there is no agreed-upon list. Many constraints have been proposed, and many algorithms too. Given the rapid changes in phonological theories and variety of constraint proposals, it is more useful to discuss general intrinsic and extrinsic restrictions on theories of constraints.

Every CDL imposes intrinsic limits on possible constraints. The nature of the elements and relations by which constraints are defined means that some imaginable constraints could not occur. For example, suppose the CDL has no disjunction operator. A constraint that assigns a violation to a segment if it is either [+voice] or [labial] is then not possible – it is impossible to formulate using the CDL's syntax. Similarly, the VR itself may impose 'restrictions' on constraint in the sense that certain constraints might be well formed in the CDL, but not assign violation marks. In the VR discussed above, constraints that had unconnected representational elements would not assign violation marks; so, while such unconnected constraints could exist, they are effectively inert and will never be observed to have an effect on selecting winning candidate. Other cases were discussed in §1.1 and §2.2.

It is also possible (even likely) that there are extrinsic limits on constraints. An extrinsic limit is a restriction on particular types of constraint even though the constraints would have a well-formed syntactic structure in the CDL. For example, suppose there was a CDL that made it possible to define a constraint that banned syllable onsets yet such a constraint did not exist: an extrinsic limit would have to be responsible. The alternative is to suppose that there are almost no significant extrinsic limits: the set of constraints includes every constraint definable using the CDL (up to a certain level of complexity). The issue of extrinsic limits is a very difficult one. The first issue addressed below is methodological: how can we tell whether there are extrinsic limits on constraints (§4.1)? Sections 4.2 and 4.3 discuss where those extrinsic limits come from.

4.1 Evidence for extrinsic restrictions

The majority of OT theories and sub-theories do propose many extrinsic restrictions on constraints, specified by the CDL (cf. discussion in Blevins 2004). The evidence comes from restrictions and requirements that cannot be attributed to non-cognitive mechanisms.

To explain, suppose we never observe a particular phonological property in any human language, like an epenthetic [k] (e.g. /iti/ never surfaces as [kiti] in any grammar). It is possible that the lack of [k] epenthesis is due to constraints. For [k] to be epenthetic, there has to be a (set of) output constraint(s) that returns violation marks for every segment except [k]. Without such a constraint, epenthetic [k] won't occur (see e.g. de Lacy 2006:ch.3).

However, there potentially are non-CDL reasons why epenthetic [k] is never observed. Some other part of the phonological component could be responsible (see section 5). There is also luck – i.e. war, pestilence, and plague, which may have accidentally wiped out all speakers of languages with epenthetic [k]. After all, every theory of phonology predicts many tens of thousands of distinct phonological systems, and only a few thousand have existed and will ever exist.

For epenthetic [k] to be observed, it also must be learnable. Actuation of a phonological change comes about through learner misperception or misarticulation. So if epenthetic [k] cannot come about through such a situation, it won't be observed. Even if a sound change can be actuated easily, if it cannot be transmitted effectively it will quickly disappear. In this particular case, though, there is evidence that [k] can be misperceived to occur in vowel hiatus environments (Kingston & de Lacy 2006).

So, if it can be shown that a particular phonological situation never occurs and this cannot be ascribed to (a) non-constraint grammatical mechanisms and (b) extra-phonological restrictions do not prevent it from never (or very rarely) occurring in natural language, phonological extrinsic restrictions on constraints are responsible. For epenthetic [k], it is easy to come up with a set of constraints that penalize everything except [k] (e.g. *labial, *coronal, *glottal), so the CDL must not permit this set of constraints (or at least, this set of constraints with free ranking – see section 5).

There are several other methods of determining that a particular phonological situation is due to constraints. See Kingston & de Lacy (2006§3.3) and references cited for discussion.

4.2 Origins and Universality of Constraints

If there are restrictions on possible constraints, where do those restrictions come from? There are fundamentally two different proposals: (a) innateness and (b) constraint-construction mechanisms that refer to phonology-external structures.

The innateness view is that constraints are hard-wired into the brain (i.e. part of our genetic make-up). The 'hard-wired' view comes in two versions. One is that each constraint is specified independently. In this version, only those constraints that are hard-wired into the brain exist, so extrinsic limits on constraints boil down to genetics. The other version is that there are hard-wired algorithms that automatically generate constraints – 'constraint generators' (sometimes called 'schemas'). For example, there would be an 'IDENT[F]' constraint generator that produces constraints with the form D-IDENT[F] where D is a pair of dimensions (input-output, base-reduplicant, etc.), and F is a subsegmental node. The generator is 'complete' in that it would generate constraints for every D and every F (Green 1993). The constraint-construction algorithms determine which constraints exist.

The alternative is to propose mechanisms that are derived from phonology-external mechanisms, or at least can take phonology-external factors into account. A growing body of work argues that there are many algorithms that take phonetic factors like ease of articulation and perceptual distinctiveness into account in evaluating which phonological constraints to generate. In this view, limits on constraints are a combination of the inherent limits of the constraint-construction algorithm *and* the restrictions imposed by the phonology-external factors that those algorithms refer to.

For example, Hayes (1995) discusses phonological constraints on voiced stops. Phonetic voicing in stops is hard to maintain; the further back the stop is, the more difficult it is to maintain voicing during the closure phase: it's harder to maintain voicing during the closure phase of [g] than for [d], and its harder for [d] than for [b]. Suppose there is a constraint generator that produces constraints on voicing in stops. It could imaginably generate many constraints *g, *d, *b, *g/d, *g/b, *d/b, *g/d/b (where **x/y* means "Return a violation for any segment that is *x* or *y*"). However, if the mechanism referred to articulation in a way that reflected voicing difficulty, the constraints would be winnowed down to *g, *g/d, *g/d/b. Hayes (1995) further observes that the CDL's intrinsic representational restrictions could impose further limits on the possible constraints: *g/d, for example, is not definable in some feature theories as there is no feature that [g] and [d] share to the exclusion of [b]; with such representational theories, the only constraints generated by the mechanism would be *g and *g/d/b (i.e. *[+voice,–continuant, –nasal]).

To summarize, the majority of work in OT adopts the idea that there are constraint generators. However, there is ongoing disagreement over whether constraint generators can refer to phonology-external factors like ease of articulation and perceptual difficulty. Gordon (2007) provides discussion and references; for recent work, see Flack (2007) and Hayes & Wilson (2008).

A related issue is constraint universality. A constraint is 'universal' if it exists in every grammar. A constraint can exist in every grammar because it is hard-wired into CON (the set of constraints), or because it is produced by a constraint generator (see

§4.3) that produces the same constraints in the same way for every grammar. A 'language-specific' constraint is one that exists in only some languages; it must be learned. For specific discussions of constraint universality, see P&S, Green (1993), Samek-Lodovici (1996, 1998a,b), Ellison (2000), McCarthy (2002a§1.2.1, §3.1.5.2), and de Lacy (2003).

There is an important nuance to constraint universality/language-specificity. It is possible that constraints are not universal, but rather *constraint generators* are. For example, ALIGN is a constraint generator that exists in every grammar. However, if ALIGN is allowed to take individual morphemes (or morphs) as arguments, it could produce language-specific constraints like ALIGN($[um]_{Af}$, L, Stem, L) "The affix *um* occurs stem-initially, is a prefix" for Tagalog, and ALIGN($[ka]_{Af}$, L, Ft', R) "The affix *ka* follows, is a suffix to, the head foot" for Ulwa (McCarthy & Prince 1995). So, while ALIGN($[um]_{Af}$, L, Stem, L) does not exist in every language, the constraint generator that created it does. The point could be extended to other constraint generators, and even those that refer to phonology-external factors. If a constraint generator refers to an articulatory or acoustic factor that varies among speakers, it could be that the same constraint generator will produce speaker-specific constraints.

4.3 Constraint Generators

Many constraint generators and elaborations or emendations to constraint generators have been proposed. A couple mentioned so far are the D-IDENT[F] generator and the ALIGN generator of McCarthy & Prince (1995). The ALIGN schema generates a great number of constraints that have been used to deal with areas as diverse as intonation (Gussenhoven 2007), syntax-sensitive prosodic phrasing (Truckenbrodt 1999, 2007), syllabification (McCarthy & Prince 1995; Zec 2007), metrical structure (Kager 2007), and morphology-sensitive prosodification (Ussishkin 2007).

A well-established development of Correspondence constraint generators is Beckman's (1998) 'positional faithfulness' theory. Beckman (1998) proposes that there are Correspondence constraints that refer to particular environments. There are only a few possible environments that can be referred to, called 'prominent positions': onset, root-initial syllable, roots, and the head syllables of feet. These environments can combine freely with faithfulness constraints: i.e. for every IDENT[F] constraint there is an onset-IDENT[F] constraint, and there are versions of MAX and DEP for each prominent position (also see Alderete 1995 for DEP). Onset-IDENT[voice] returns a violation for each output segment that is both in an onset and corresponds to an input segment that has a different value for [voice] (as in /pa/ → [ba], but not /ap/→[ab]).

A constraint generation mechanism that is somewhat different from IDENT, ALIGN and positional faithfulness is local conjunction (Smolensky 1993, Smolensky & Legendre 2006). Local conjunction combines *constraints* to form new constraints. For example, NOCODA and *LABIAL can be combined to form $*[\text{LABIAL\&NOCODA}]_{\text{SEG}}$ where this is violated only when a labial is dominated by a coda node (e.g. [ap]). Local conjunction is not simply a conjunction of violation marks; *LABIAL&NOCODA is not violated by [pak], even though [pak] violates *LABIAL and NOCODA. Instead, local conjunction carries a condition that the violation is localized to a specific domain: i.e. the segment that is the source for the violation of *LABIAL must be the segment that is the source for the violation of NOCODA. Local conjunction is a controversial mechanism; see Pater (2009a) for discussion.

5. Constraints interact

No constraint is an island. The effect of any individual constraint on the output of a phonological system depends on the other constraints in CON, the candidates GEN creates, the constraint's ranking, how EVAL determines the winner, and how the interface to the phonetic module processes the candidate.

For example, feet with more than two syllables apparently do not occur (e.g. Hayes 1995 and many others). It is possible that the set of constraints ensures that no winner has a foot with more than two syllables. However, it is also possible that GEN is incapable of generating such oversize feet (Hyde 2002:318). So, whether there needs to be a constraint against trisyllabic feet depends on one's theory of GEN.

The theory of constraints also interacts with the theory of EVAL – the component that determines winners. EVAL doesn't tally the number of constraint violations for each candidate and pick the candidate with the fewest violation marks; it refers to both constraint violations and *ranking* to determine the winner. In P&S's OT, ranking is a total order on constraints. Suppose candidate A violates constraint C_1 but not C_2, candidate B violates C_2 and not C_1, and C_1 outranks C_2. The winner of the A vs. B competition will be B because A violates the higher ranked constraint C_1 but B does not. Restrictions on ranking can limit the kind of constraints needed in CON.

Let's consider some typological empirical generalizations: /k/ and /p/ can become [ʔ], but /k/ never becomes [p], /p/ never becomes [k], and /ʔ/ never becomes [k] or [p] (de Lacy 2006:ch.3 and references cited therein). There's a way to express these observations using both constraint form *and* ranking. Suppose ranking can be universally fixed between constraints, as proposed by P&S. The fixed rankings would be (at least) *k » *ʔ and *p » *ʔ. It's not enough to just fix these rankings, it is also essential that there is no other constraint C where C favors [p] or [k] over [ʔ]. The two sets of fixed rankings mean that /k/→[ʔ] and /p/→[ʔ] but never /ʔ/→[p], /ʔ/→[k]. So, the systems that the phonological module can produce are partly determined by constraints and their violations, but also by ranking and by the nature of EVAL itself.

In the 'fixed ranking' example, the results were achieved by two things: a particular theory of CON that included *k, *p, *ʔ *and* a theory of EVAL that allowed universally fixed rankings. However, suppose that EVAL does not allow universally fixed rankings. The empirical generalizations above are immediately beyond reach with a CON that has freely rankable *k, *p, and *ʔ. Instead, without fixed ranking it is necessary to invoke a CDL theory that employs constraints that are 'stringently' related in terms of their violation mark assignments. Suppose that there are no *p, *ʔ constraints but instead, there is *k, *k/p, *k/p/ʔ, where **x*/*y* returns a violation for each segment that is either *x* or *y*. So *k/p is violated twice in each of [kak], [kap], [pap], and once in each of [kat] and [pat]. There is no way to rank these constraints to make /ʔ/ become [k] or [p] because [ʔ] always incurs a proper subset of the violations that [k] and [p] incur.

So, theories of constraints depend closely on the theory of EVAL, and vice versa. The empirical generalizations about neutralization can be achieved by non-stringent constraints with fixed ranking, or by stringent constraints without fixed ranking (or even by stringent constraints with a particular fixed ranking); however, it cannot be achieved by non-stringent constraints and no fixed ranking. (For further discussion, see de Lacy 2004, 2006). In short, there's a connection between the CDL and what the algorithm that determines winners can do. (The interconnectedness of the theory of constraints and theory of EVAL is particularly clear in proposals for EVALs that differ from P&S's (e.g. Boersma 1998, Wilson 2001, McCarthy 2007).)

The theory of constraints is also influenced by the theory of the phonology-phonetics interface. If GEN can create phonological representations that are phonetically uninterpretable, candidates with uninterpretable structures must be eliminated by either constraints or by the mechanism that sends constraints to the phonetic module. For discussion, see de Lacy (2007).

Finally, it is crucially important to evaluate the adequacy of a constraint by examining how it interacts with all the other constraints in CON. A constraint might seem eminently reasonable on its own, but when its effect among all the other constraints is examined bizarre and undesirable effects can emerge.

For example, there are two well-known and widely accepted sets of constraints that have a surprising effect when they interact. P&S propose a set of constraints that favor syllables with highly sonorous nuclei over syllables with less sonorous nuclei (the *NUC/*x* constraints). Beckman (1998) proposed that there are constraints that preserve features in stressed syllables (σ́-IDENT[F]) (see §4.3 above). Both sets of constraints have ample support, and their effects as individual sets have been well studied: the nucleus-sonority constraints are crucial in accounting for syllabification (P&S 1993, Zec 2007); the stress-faithfulness constraints are needed to explain why features survive in stressed syllables but not elsewhere.

However, when these sets of constraints interact, they have a surprising effect: if the stress-faithfulness constraints outrank the nucleus sonority constraints, which in turn outrank all constraints on the relation between unstressed syllables and sonority, the resulting system neutralizes all vowels to *high* sonority segments in unstressed syllables: e.g. /pitiku/ → [pítako]. Such a system is very surprising; vowel neutralization in unstressed syllables is supposed to create *less* sonorous vowels like [ə], or disperse vowels so that higher sonority vowels become more sonorous and lower sonority vowels become less sonorous (Crosswhite 1999). This particular weird result has a happy ending, by the way: Ibibio has exactly this type of system, so validating the proposals (Akinlabi & Erua 2002, Akinlabi 2006). (For more on the theoretical point, see de Lacy 2006:315ff.) The general point, though, is that what constraints (or groups of constraints) seem to do on their own can be very different from what they actually do when they interact with all the other constraints in CON.

6. Summary

This chapter has left a vast number of issues about constraints untouched and only barely skimmed over a few others. However, a few points about constraints emerge. A formal theory of constraint form – a 'Constraint Definition Language' – provides valuable insight into which constraints can and cannot exist. There is fairly widespread (if tacit) agreement on many aspects of such a CDL, but also many disagreements about both fundamental issues and the details. Constraints are only one part of a complex system that determines phonological winners; constraints themselves do not determine winners; and no constraint or set of constraints has any predictive power on its own. Only when the entire collection of constraints, GEN, EVAL, and the phonetic module interface are examined together can anything be asserted about the predictive power or restrictive nature of the theory.

References

Akinlabi, Akinbiyi (1996). "Featural affixation." *Journal of Linguistics* 32.2: 239-289.

Akinlabi, Akinbiyi (2006). *Ibibio vowel distribution*. ms. Rutgers University.

Akinlabi, Akinbiyi & Eno E. Urua (2002). Foot structure in the Ibibio verb. *Journal of African Languages and Linguistics* 23: 119-160.

Albright, Adam (2002). The identification of bases in morphological paradigms. Doctoral dissertation, UCLA.

Alderete, John (1997). Dissimilation as local conjunction. In K. Kusumoto (ed.) *Proceedings of NELS 27*. Amherst, MA, GLSA**:** 17-32.

Alderete, John (2001). Dominance effects as transderivational anti-faithfulness. *Phonology* 18:201-253. (ROA 407)

Archangeli, Diana and Keiichiro Suzuki (1997). The Yokuts challenge. In Iggy Roca (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press, pp.197-226.

Bach, Emmon (1968). Two proposals concerning the simplicity metric in phonology. *Glossa* 2:2-49.

Baković, Eric (1999). Assimilation to the unmarked. In Jim Alexander, Na-Rae Han, and Michelle Minnick Fox (eds.) *University of Pennsylvania Working Papers in Linguistics* 6.1. http://www.ling.upenn.edu/papers/v6.1-contents.html [29 March 2009]. Also Rutgers Optimality Archive 340.

Baković, Eric (2007). Local assimilation and constraint interaction. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.335-351.

Balari, Sergio, Rafael Marín, and Teresa Vallverdú (2000). Implicational constraints, defaults, and markedness. GGT Report de Recerca GGT-00-8, Universitat Autònoma de Barcelona. Rutgers Optimality Archive 396.

Beckman, Jill (1998). Positional faithfulness. Doctoral dissertation, UMass. Also Rutgers Optimality Archive #234.

Benua, Laura (1997). Transderivational identity: phonological relations between words. Doctoral dissertation, UMass. Rutgers Optimality Archive #259.

Bermúdez-Otero (to appear). *Stratal Optimality Theory*. Oxford University Press.

Bernhardt, Barbara and Joseph P. Stemberger (1998). Handbook of phonological development: From the perspective of constraint-based nonlinear phonology. San Diego, Academic Press.

Bernhardt, Barbara and Joseph P. Stemberger (2007). Phonological impairment in children and adults. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.575-594.

Blaho, Sylvia (2003). Derived environment effects in Optimality Theory: The case of presonorant voicing in Slovak. In Wayles Browne, Ji-yung Kim, Barbara H. Partee, and Robert A. Rothstein (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The Amherst Meeting 2002*. Ann Arbor: Michigan Slavic Publications, pp. 103-120.

Blaho, Sylvia (2008).

Blevins, Juliette (2004). Evolutionary Phonology: the emergence of sound patterns. Cambridge, CUP.

Blumenfeld, Lev (2003a). Russian palatalization and stratal OT: Morphology and [back]. In Wayles Browne, Ji-yung Kim, Barbara H. Partee, and Robert A. Rothstein (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The Amherst Meeting 2002*. Ann Arbor: Michigan Slavic Publications, pp. 141-158.

Blumenfeld, Lev (2003b). Counterfeeding, derived environment effects, and comparative markedness. *Theoretical Linguistics* 29: 89-99.

Bobaljik, Jonathan (2008). Paradigms (Optimal and Otherwise): a case for skepticism. In Asaf Bachrach and Andrew Ira Nevins (eds.) *Inflectional Identity*. Oxford University Press, pp. 29-54. Rutgers Optimality Archive 856.

Boersma, Paul (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. PhD dissertation, University of Amsterdam. LOT International Series **11**. The Hague: Holland Academic Graphics. [http://www.fon.hum.uva.nl/paul/diss/]

Boersma, Paul & Bruce Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* **32**: 45-86.

Burzio, Luigi (2002). Surface-to-surface morphology: when your representations turn into constraints. In P. Boucher (ed.) Many morphologies. Somerville, MA, Cascadilla Press, pp.142-177. [Also ROA 341].

Bye, Patrik and Paul de Lacy (2000). Edge asymmetries in Phonology and Morphology, In Ji-Yung Kim and Masako Hirotani (eds.) *Proceedings of NELS (North-eastern Linguistics Society conference) 30.* Amherst, MA: GLSA Publications, pp.121-135. Rutgers Optimality Archive 328.

Bye, Patrik (2002). Virtual phonology: rule sandwiching and multiple opacity in North Saami. Doctoral dissertation, University of Tromsø. Rutgers Optimality Archive 498.

Causley, Trisha (1997). Identity and featural correspondence: the Athapaskan case. In Kiyomi Kusumoto (ed.) *Proceedings of NELS 27*. Amherst, MA: GLSA, pp. 93-105.

Causley, Trisha (1999). Complexity and markedness in Optimality Theory. Doctoral dissertation, University of Toronto.

Chomsky, Noam (1995). *The Minimalist Program*. MIT Press.

Chomsky, Noam and Morris Halle (1968). The sound pattern of English. New York, Harper & Row.

Chomsky, Noam and Howard Lasnik (1977). Filters and control. Linguistic Inquiry 8: 425-504.

Clements, G. N. and Elizabeth V. Hume (1995). The internal organization of speech sounds. In John Goldsmith (ed.) The handbook of phonological theory. Oxford, Blackwell, pp.245-306.

Crosswhite, Katherine M. (1999). Vowel reduction in Optimality Theory. Doctoral dissertation, UCLA.

Crowhurst, Megan, and Mark Hewitt (1997). Boolean operations and constraint interactions in Optimality Theory. Ms., University of North Carolina, Chapel Hill, and Brandeis University, Waltham, MA. Rutgers Optimality Archive 229.

de Lacy, Paul (2002). The formal expression of Markedness. Doctoral dissertation, UMass. Rutgers Optimality Archive 542.

de Lacy, Paul (2003). Constraint universality and prosodic phrasing in Māori. In Angela Carpenter, Andries Coetzee, and Paul de Lacy (eds.) *Papers in Optimality Theory II*. UMOP 26. Amherst, MA: GLSA Publications, pp.59-79. [Also Rutgers Optimality Archive #561]

de Lacy, Paul (2004). Markedness conflation in Optimality Theory. Ph 21: 145-199.

de Lacy, Paul (2006). Markedness: Reduction and preservation in phonology. Cambridge, CUP.

de Lacy, Paul (2007). Freedom, Interpretability, and the Loop. In Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.). *Freedom of Analysis?*. Studies in Generative Grammar 95. New York: Mouton de Gruyter, pp. 86-118.

Eisner, Jason (1997a). Decomposing FootForm: Primitive constraints in OT. Rutgers Optimality Archive 205.

Eisner, Jason (1997b). Efficient generation in Primitive Optimality Theory. Rutgers Optimality Archive 206.

Eisner, Jason (1997c). What constraints should OT allow? Rutgers Optimality Archive 204.

Elías Ulloa, José (2006). *Theoretical aspects of Panoan metrical phonology: Disyllabic footing and contextual syllable weight*. Doctoral Dissertation, Rutgers University. Rutgers Optimality Archive 804.

Ellison, T. Mark (2000). The universal constraint set: Convention, not fact. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford: Oxford University Press. [ROA-987]

Flack, Kathryn (2007). Inducing functionally-grounded constraints. *In Michael Becker (ed.), University of Massachusetts Occasional Papers in Linguistics* 36: *Papers in Theoretical and Computational Phonology*. Amherst: GLSA. Rutgers Optimality Archive 930.

Goldrick, Matthew (2000). Turbid output representations and the unity of opacity. In M. Hirotani (ed.) Proceedings of the North East Linguistic Society. Vol.30. Amherst, MA, GLSA, pp.231-246. Rutgers Optimality Archive 368.

Gordon, Matthew (2007). Functionalism in phonology. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.61-78.

Gouskova, Maria (2008). DEP: Beyond epenthesis. Rutgers Optimality Archive 948.

Green, Thomas (1993). The conspiracy of completeness. Rutgers Optimality Archive 8.

Gussenhoven, Carlos (2007). Intonation. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.253-280.

Hall, T.A. (2006). Derived environment blocking effects in Optimality Theory. *Natural Language and Linguistic Theory* 24: 803-856.

Hall, T.A. (2007a). Comparative Markedness vs. standard faithfulness theory: A typological comparison. Ms., Indiana University. Rutgers Optimality Archive 942.

Hall, T.A. (2007b). Segmental features. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.311-333.

Hansson, Gunnar (2001). *Theoretical and typological issues in consonant harmony*. PhD Dissertation, University of California, Berkeley.

Hayes, Bruce (1995). Metrical Stress Theory: principles and case studies. Chicago, University of Chicago Press.

Hayes, Bruce, and Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, pp, 379-440.

Hyde, Brett (2002). A restrictive theory of metrical stress. *Phonology* 19: 313-359.

Hyman, Larry (1985). A theory of phonological weight. Dordrecht, Foris.

Inkelas, Sharon (1999). Phonotactic blocking through structural immunity. In Barbara Stiebels & Dieter Wunderlich (eds.) *Lexicon in focus*. Akademie Verlag, pp.7-40. Rutgers Optimality Archive 366.

Ito, Junko (1986). Syllable theory in prosodic phonology. Doctoral dissertation, UMass.

Jun, Jongho (1999). Generalized Sympathy. Rutgers Optimality Archive 297.

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

Kager, René (2007). Feet and metrical stress. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.195-227.

Keer, Edward (1999). Geminates, the OCP and the Nature of CON. Doctoral Dissertation, Rutgers University. Rutgers Optimality Archive 350.

Kingston, John and Paul de Lacy (2006). Synchronic explanation. ms. University of Massachusetts Amherst and Rutgers University. http://equinox.rutgers.edu/~delacy/research/publications

Kiparsky, Paul (to appear). *Paradigms and Opacity*. CSLI Publications.

Kitto, Catherine and Paul de Lacy (1999). Correspondence and epenthetic quality. In Carolyn Smallwood and Catherine Kitto (eds.) Proceedings of AFLA VI: The Sixth Meeting of the Austronesian Formal Linguistics Association. Toronto, University of Toronto, pp.181-200.

Lombardi, Linda (2001). Why Place and Voice are different: Constraint-specific alternations in Optimality Theory. In Linda Lombardi (ed.) Segmental Phonology in Optimality Theory: Constraints and Representations. Cambridge, CUP, pp.13-45. Rutgers Optimality Archive 105.

Łubowicz, Anna (2002). Derived Environment Effects in Optimality Theory. *Lingua* 112: 243-280. Rutgers Optimality Archive 765.

Łubowicz, Anna (2005). Locality of conjunction. Rutgers Optimality Archive 764. Published in John Alderete, Chung-hye Han, and Alexei Kochetov (eds.) *Proceedings of the 24th West Coast Conference on Formal Linguistics*. Cascadilla Press, pp. 254-262. http://www.lingref.com/cpp/wccfl/24/index.html

McCarthy, John (1999). Sympathy and phonological opacity. *Phonology* 16: 331-399.

McCarthy, John (2000). The prosody of phase in Rotuman. *Natural Language and Linguistic Theory* 18: 147-197.

McCarthy, John (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

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

McCarthy, John (2005) Optimal paradigms. In Laura Downing, Tracy Alan Hall, and Renate Raffelsiefen (eds.) *Paradigms in Phonological Theory*. Oxford: Oxford University Press. pp. 170-210. Rutgers Optimality Archive 485.

McCarthy, John (2007). *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

McCarthy, John & Alan Prince (1986). *Prosodic morphology*. Rutgers Technical Reports TR-32. Rutgers University Center for Cognitive Science.

McCarthy, John J. and Alan Prince (1993a). Generalized alignment. In Geert Booij and Jaap van Marle (eds.) Yearbook of Morphology. Dordrecht, Kluwer, pp.79-153. Rutgers Optimality Archive 7.

McCarthy, John and Alan Prince (1993b). Prosodic morphology I: Constraint interaction and satisfaction. Rutgers Technical Report TR-3. New Brunswick, Rutgers University Center for Cognitive Science. Rutgers Optimality Archive 482.

McCarthy, John & Alan Prince (1997). Faithfulness and Identity in Prosodic Morphology. Rutgers Optimality Archive 216.

McCarthy, John & Alan Prince (1999). Faithfulness and Identity in Prosodic Morphology. In Harry van der Hulst, René Kager, and Wim Zonneveld (eds.) *The Morphology- Prosody Interface*, Cambridge University Press, pp. 218-309.

Miner, K. (1992). Winnebago accent: The rest of the data. In *Indiana University Linguistics Club twenty-fifth anniversary volume*. Bloomington, Ind.: Indiana University Linguistics Club. pp. 28-53.

Myers, Scott (1997). OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15.4: 847-892. Rutgers Optimality Archive 6.

Pater, Joe (1999). Austronesian nasal substitution and other NC effects. In Rene Kager, Harry van der Hulst and Wim Zonneveld (eds.) The prosody morphology interface. Cambridge, MA, CUP, pp.310-343.

Pater, Joe. (2009a). Review of Smolensky and Legendre *The Harmonic Mind*. Rutgers Optimality Archive 983.

Pater, Joe (2009b). Weighted constraints in generative linguistics. *Cognitive Science* 33: 999-1035. Rutgers Optimality Archive 982.

Potts, Christopher & Geoffrey K. Pullum (2002). Model theory and the content of OT constraints. *Phonology* 19.3: 361-393.

Prince, Alan and Paul Smolensky (1993). Optimality Theory: constraint interaction in generative grammar, Rutgers Technical Reports TR-2.

Prince, Alan and Paul Smolensky (2002). Optimality Theory: Constraint interaction in Generative Grammar. Rutgers Optimality Archive 537.

Prince, Alan and Paul Smolensky (2004). Optimality Theory: Constraint interaction in generative grammar. Oxford, Basil Blackwell.

Pullum, Geoffrey (1976). The Duke of York gambit. *Journal of Linguistics* 12: 83-102.

Rice, Keren (2007). Markedness in phonology. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.79-98.

Riggle, Jason (2004). Generation, Recognition, and Learning in Finite State Optimality Theory. Ph.D. dissertation, UCLA. http://hum.uchicago.edu/~jriggle/riggleDiss.html

Rose, Sharon and Rachel Walker (2004). A typology of consonant agreement as Correspondence. *Language* 80: 475-531.

Sagey, Elizabeth (1986). The representation of features and relations in nonlinear phonology. Doctoral dissertation, MIT.

Samek-Lodovici, Vieri (1996). Constraints on subjects: An Optimality Theoretic analysis. Doctoral Dissertation, Rutgers University.

Samek-Lodovici, Vieri (1998a). "OT-interactions between focus and canonical word order: Deriving the cross-linguistic typology of structural contrastive focus." Rutgers Optimality Archive #257.

Samek-Lodovici, Vieri (1998b). "Opposite constraints: Left and right focus alignment in Kanakuru." Lingua 104: 111-130.

Schein, Barry and Donca Steriade (1986). On geminates. *Linguistic Inquiry* 17:691-744.

Selkirk, Elisabeth (1991). A two-root theory of length. In E. Dunlap and Jaye Padgett (eds.) *University of Massachusetts Occasional Papers in Linguistics 14: Papers in phonology*. Amherst, MA: GLSA, pp. 123-172.

Smolensky, Paul (1993). Harmony, markedness, and phonological activity. Rutgers Optimality Archive 87.

Smolensky, Paul & Géraldine Legendre (2006). *The harmonic mind: From neural computation to Optimality-Theoretic Grammar*. *Volume 2: Linguistic and Philosophical Implications*. Cambridge, MA.: MIT Press.

Struijke, Caro (2000). Existential faithfulness: A study of reduplicative TETU, feature movement, and dissimilation. Doctoral dissertation, University of Maryland, College Park.

Suzuki, Keiichiro (1998). A typological investigation of dissimilation. Doctoral dissertation, University of Arizona.

Tesar, Bruce (2007). Learnability. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.555-574.

Tesar, Bruce and Paul Smolensky (1998). Learnability in Optimality Theory. Linguistic Inquiry 29.2: 229-268.

Truckenbrodt, Hubert (1999). On the relation between syntactic phrases and phonological phrases. Linguistic Inquiry 30.2: 219-255.

Truckenbrodt, Hubert (2007). The syntax–phonology interface. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.435-456.

Urbanczyk, Suzanne (2007). Reduplication. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.473-494.

Ussishkin, Adam (2007). Morpheme position. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.457-472.

van Oostendorp, Marc (2007). Restricting repairs. ms. Meertens Institute. http://www.vanoostendorp.nl/pdf/toomanyrepairs.pdf

Wilson, Colin (2001). Consonant cluster neutralization and targeted constraints. *Phonology* 18: 147-197.

Wolf, Matthew (2005). For an autosegmental theory of mutation. ROA 754.

Wolf, Matthew (2007). What constraint connectives should be permitted in OT? ROA 926.

Wolf, Matthew (2008). Optimal Interleaving: Serial phonology-morphology interaction in a constraint-based model. PhD Dissertation, University of Massachusetts Amherst. Rutgers Optimality Archive 2008.

Yip, Moira (2001). The complex interaction of tone and prominence. In M. Kim and U. Strauss (eds.) Proceedings of the North-Eastern Linguistics Society 31. Amherst, MA, GLSA, pp.531-545.

Yip, Moira (2007). Tone. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.195-227.

Zec, Draga (2007). The syllable. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press, pp.161-193.

Zhang, Jie (2000). Non-contrastive features and categorical patterning in Chinese diminutive suffixation: MAX[F] or IDENT[F]? *Phonology* 17: 427-478.

PAUL DE LACY
Associate Professor
Rutgers the State University of New Jersey

E-mail: delacy@rutgers.edu

Biographic information:
Paul de Lacy is an associate professor in the Department of Linguistics at Rutgers University, an Associate of the Rutgers Center for Cognitive Science, co-director of the Rutgers Phonetics and Fieldwork Laboratory, and editor of the *Cambridge Handbook of Phonology*. He works on phonology and its interfaces with syntax, morphology, and phonetics.