From MIT Working Papers in Philosophy and Linguistics, Volume 1 (2000)

Optimality Theory and the Problem of Constraint Aggregation*

Daniel Harbour, MIT and Christian List, Oxford University

Optimality Theory (Prince & Smolensky, 1993) claims that "Universal Grammar consists largely of a set of constraints on representational well-formedness"[1] and that grammaticality is a matter of structural well-formedness. As constraints make "sharply conflicting claims about the well-formedness of most representations", grammars require a means of resolving such conflicts in order to determine a given input's "surface representation", the analysis "most harmonic" with the well-formedness constraints. Prince & Smolensky (1993) claim that conflicts are resolved by "rank[ing] constraints in a *strict dominance hierarchy*" and most optimality-theoretic research has been conducted using constraints in precisely that manner. This paper questions that assumption and aims to present a mathematical answer to the question of whether conflict resolution can be theoretically modelled in other ways. The answer we give casts light on problems in Optimality Theory, in particular, that of opacity.

To examine these issues in a rigorous, mathematical fashion, a certain amount of formalism is requisite. The formal sections of the paper are, however, firmly tied to the linguistics. Before preceding to them, we first sharpen the observation in the opening paragraph, namely, that Optimality Theory is just one implementation of the central intuition of constraint-based phonology. We then ask which idiosyncrasies of Prince & Smolensky's implementation are desirable and which it might be productive to question. In particular, we focus on the nature of constraints, constraint violation, and constraint aggregation. Once one appreciates that there is a difference between constraint-based phonology, generally construed, and Optimality Theory, as a particular example of it, one can ask the following question: Of Optimality Theory's perceived problems, which are indeed problems with Optimality Theory which would not be shared by other constraint-based theories, and which are ineluctable results of the use of constraints? For an answer, we turn to the formalisation.

The formal methods of the paper are drawn from the techniques of mathematical decision theory, in particular, from the area of social choice

* We are grateful to Jason Eisner for detailed comments and suggestions that have helped us to improve and clarify the content of this paper; to Benjamin Bruening for discussion of data; to Bill Idsardi; and to Sylvain Bromberger, John Frampton, Morris Halle and Sam Rosenthall for comments on an early draft. Daniel Harbour gratefully acknowledges the financial support of MIT; Christian List, the financial support of the British Economic and Social Research Council and the German National Scholarship Foundation.

[1] All quotations from Prince & Smolensky (1993) are from pages 1-6, 20-21, and 67-68.

theory. The reasons for this are made apparent below. Some comments on what the mathematics mean for linguistics are interspersed through the formalism. The main discussion of the formal results follows their presentation.

The formalism also permits one to ask what the relationship is between Optimality Theory as presented in Prince & Smolensky (1993) and as it developed in subsequent work (Smolensky 1995; Kirchner, 1996; Flemming, 1997; Itô & Mester, 1997; Burzio 1999; Crowhurst & Hewitt, 1999). That is, we can ask where exactly the difference between Prince & Smolensky (1993) and, say, Kirchner (1996) lies and how that difference affects the generation of phonological surface representations. For further applications of the formal methods presented below, especially in relation to the notion of economy of derivation in syntax (Chomsky, 1995), see Harbour & List (forthcoming).

1. Constraint Conflict Resolution Generally

To appreciate that Optimality Theory is merely one possible implementation of the intuition that constraints on well-formedness form the core of grammar, it is helpful to return to Prince & Smolensky (1993) and to place it in the context of the rise of constraints in phonology generally.

Throughout the seventies and eighties, constraints came to play an increasingly prominent role in phonology. The model of phonology that developed in the years following Chomsky & Halle (1968) comprised four levels: the underlying representation, the concatenation of morphemes, the application of rules, the surface representation. This is shown diagrammatically in (1).

(1) underlying representation
⇓
morpheme concatenation
⇓
application of rules
⇓
surface representations

Archangeli (1997), from whom (1) is taken, observes that this neat picture became more complicated throughout the following decades as constraints began to be introduced at all levels. The result was (2).

(2)		
	underlying representation	constraints on underlying representations apply here
	⇓	
	morpheme concatenation	constraints on morpheme concatenation apply here
	⇓	
	application of rules	constraints on applicability/output of rules apply here
	⇓	
	surface representations	other constraints apply here

In addition, phenomena came to be discussed in which the regularities of interest lay, not in the input structures, nor in the nature of the transformational component, but in the output structures themselves. Examples of this are conspiracies and paradigmatic uniformity.[2] Kisseberth (1970) presents a conspiracy evident in some Amerindian languages, whereby rules of epenthensis and deletion ensure the absence of triconsonantal clusters. Various instances of paradigm uniformity are presented in Kenstowicz (1999).

What unites these cases is that one seems able to ask why the rules of a language are such as they are. The answer is that certain structures are 'more optimal' than others and that the purpose of phonological rules is to force non-optimal structures, resulting, say, from morpheme concatenation, towards more optimal ones. Convergence on a particular output structure (paradigmatic uniformity) or divergence from a given output structure (conspiracies) may thus be achieved by quite heterogeneous rules. Crucially, the motivation for the different rules, a constraint on output structures, is external to the transformational component, and, hence, difficult to express in traditional, derivational phonology.

Faced with this situation, Prince & Smolensky (1993) conclude that any theory of phonology "committed to Universal Grammar" must "rely heavily on well-formedness constraints". What remained in "subformal obscurity" was "the character of the interaction among the posited well-formedness constraints", "the relation between such constraints and whatever derivational rules they are meant to influence", and, of course, the exact nature of the constraints themselves. The most interesting hypothesis under such circumstances is the strongest: that the character of the interaction, whatever it may be, is the whole of phonology and that there is no interaction between constraints and derivations because the derivational component is null. In this sense, Optimality Theory represents the apogee of the trend, current in the seventies and eighties, which gave constraints ever increasing prominence in phonological theories.

Optimality Theory claims that structural well-formedness is the essence of grammaticality. Well-formedness is achieved by satisfying well-formedness constraints. However, the constraints are "sharply conflicting" and so grammars must be able to resolve constraint conflicts in determining how a given input surfaces. Resolution of constraint conflicts requires a specific proposal. Prince & Smolensky recognise that there are two components to the problem of conflict resolution: "first, a means of comparing entire candidates on the basis of a single constraint; then, a means of combining the evaluation of these constraints". They write that "regulating the way these two dimensions ... interact is a key

---

[2] We remain neutral on whether these motivate a shift to constraint-based phonology. They are relevant in their connection to the rise of constraints in generative phonology.

theoretical commitment". Their response is that conflicts are resolved by "rank[ing] constraints in a *strict dominance hierarchy*" and that, internal to any constraint, all that counts is having the minimal number of violations.

We have thus traced a path from the emergence of constraints in phonological theories generally to a specific implementation of them in a pure constraint-based theory. What should be clear is that we are dealing with two separate things: one is an intuition and the other is an implementation, a proposal. The intuition says merely that phonological phenomena are regulated by the resolution of constraint conflicts. The proposal concerns the specifics of conflict resolution. For "comparing ... candidates on the basis of a single constraint", Prince & Smolensky propose simple ordinal measurement: a candidate receives one '*' per violation and the fewer violations a candidate has, the better. For "combining the evaluation of the ... constraints", that is, to aggregate the scores, they propose a hierarchy where each constraint acts in turn to reduce the size of the candidate set. However, ordinal measurement and strict dominance hierarchies are not, conceptually speaking, the only ways to measure and aggregate when selecting the optimal candidate. Therefore, it is natural to ask what the alternative measurement and aggregation procedures are and what effects different choices have on the shape of the resulting theory.

Recognising the availability of numerous solutions to the problem of resolution of constraint conflict leads to another question. Of what are perceived to be problems for Optimality Theory, which problems are genuine problems for any constraint-based phonological theory and which are merely problems for strict dominance hierarchies, Prince & Smolensky's (1993) resolution of constraint conflict? As an example, we ask below whether a different resolution of constraint conflict could (dis)solve the problem of opacity.

The possibility of shedding new light onto a well-known problem adds interest to what is otherwise merely a technical observation. That is, it is one thing simply to observe that Prince & Smolensky (1993) use one of many possible methods for resolving constraint conflicts. The possibility of showing that opacity, a real difficulty for Optimality Theory, is in fact no problem for alternative implementations of constraint-based phonology motivates closer examination of other ways to resolve constraint conflicts. Just as interesting is the alternative possibility, that one might prove opaque phenomena to be unrepresentable in any uniform, tidy fashion within any purely constraint-based phonology.

As part of the historical sketch, it is important to note that Prince and Smolensky were aware of the possibility of alternative to strict dominance hierarchies. The connectonist-theoretic origins of Optimality Theory highlighted in Prince & Smolensky (1997) made weighted constraints one alternative. It happened, however, that the specific problems that Prince & Smolensky (1993) focused on were soluble with weightings in effect equivalent to strict dominance hierachies. The current investigation can be regarded, then, as asking how Prince and Smolensky's theory might have differed, had they chosen to formalise alternative methods of aggregation as part of Optimality Theory. Below, we

present such a formalisation and examine some of its consequences for the theory of phonology.

The remainder of the paper is divided into four parts. In Section 2, we motivate use of social-choice-theoretic methods within constraint-based phonology by illustrating the formal similarities between the problems that the two address. In Section 3, we present our formalism together with various definitions, assumptions, and theorems. In Section 4, we discuss the significance of the results in Section 3 and the appropriateness of the assumptions that underlie those results. We also show that various modifications of Optimality Theory current in the literature can be characterised in terms of our formalism, and discuss the problem of opacity and its relationship to constraint-based phonology generally.

2. Conflict Resolution as an Aggregation Procedure

The central formal intuition pursued in this paper is that the problem of how to resolve constraint conflict is formally very similar to the problems of preference aggregation explored in the branch of mathematical decision theory called social choice theory. Therefore, with suitable modifications, the methods of social choice theory can be applied to the problem within linguistics.

Social choice theory deals with situations such as the following. Suppose a social planner is asked to rank a set of policy options. The planner proceeds in the following way. For each member of society (person) and each policy option, the planner determines a score which measures the strength of that person's (dis)preference for that policy. On the basis of the options' scores across all members of society, the overall *social ordering* of these options is determined. Three questions arise. How much information does each person's set of scores provide about their preferences? To what extent can one compare different people's preferences? How is the social ordering determined on the basis of people's preferences? Call these the problem of measurability, the problem of (interpersonal) comparability, and the problem of aggregation, respectively.

To frame the problems more precisely, suppose that the social planner assigns to every person-policy pair a number measuring that person's (dis)preference for that policy. The problem of measurability asks how much significance we attach to those numbers. Are they significant on some absolute scale? Are only their ratios and/or differences important? Is the way in which they rank-order the different policy options all that is meaningful? The problem of (interpersonal) comparability asks whether one can compare the numbers representing different people's preferences for the same or different policy options. Is it meaningful, for instance, to ask whether option x is more strongly (dis)preferred by person 1 than option y is by person 2? Only once these two questions are answered can one ask how aggregation determines the winner.

To see the relevance of social-choice-theoretic methods to the problem of conflict resolution, recall how Optimality Theory determines the output for any given input. The input is fed into the generator (Gen), which produces the set of possible outputs for that input. These are the policy options in the social-choice-theoretic problem. There are, as mentioned in Section 1, two stages involved in determining the final output from the set of candidate outputs. The first is measurement of the degree to which each candidate violates each constraint. The second is the determination of the optimal candidate on the basis of the first. Clearly, the first of these is the problem of measurability, and the second, the problem of aggregation. Where is the problem of comparability?

In Optimality Theory, as traditionally set out, the problem of (interpersonal) comparability is not addressed and no use is made of correlates of (interpersonal) comparisons. Clearly, the social planner may want to ask whether a homeless person is better off under a policy of wealth redistribution than a millionaire is under a policy of low taxation (whether person 1 is better off under option x than person 2 is under option y). Moreover, this question is obviously distinct from the issue of whether the millionaire has a greater say in the choice of policy than the homeless person has. Who has greater influence is an issue of an individual's 'rank', whereas welfare under distinct policy options is an issue of interpersonal comparability. To the linguist, the issue of ranking is familiar, but interpersonal comparability, not.

The lack of 'cross-constraint comparability', a linguistic counterpart to interpersonal comparability, does not mean that we should dismiss cross-constraint comparability from our considerations. For one thing, closer consideration may furnish an appropriate notion of comparability. For another, if Optimality Theory without cross-constraint comparability should be discovered to be non-viable, then such comparisons might be forced on it. The formal exposition therefore includes, not only measurability and aggregation, but also comparability.

Thus, there is considerable similarity between the social-choice-theoretic problems and the problem of constraint conflict resolution in Optimality Theory. Both attempt to formalise a procedure for selecting the winning or most harmonic or optimal candidate from a given set. In so doing, they face, to varying degrees, three issues, labelled above as the problems of measurability, comparability, and aggregation. Given the formal similarity of the problems, it is justifiable to attempt to use the insights of one field in the other.

3. The Formalism

We have argued that there are good philosophical reasons for examining alternatives to Prince & Smolensky's (1993) procedure for selecting the optimal output for a given input. In what follows, we expound a framework for formalising the problem of constraint aggregation in Optimality Theory. We start with a few definitions. Once strict dominance hierarchies are defined, we relate the definitions back to the conventional optimality-theoretic view.

A *generator* is a function, Gen, which maps each given input, Input, to a set of possible outputs, X. Optimality Theory, in its purest form, should make the generator as simple as possible. This follows from Optimality Theory's commitment to the idea, discussed in the Introduction, that grammaticality is determined by conformity to well-formedness constraints. If X varies from input to input, then part of the work is shifted from the constraints to the generator. This can illustrated using the mock word 'blick'. It should not be the case that, for the input /cats/, it is the *constraints* that rule out the output *blick*, whereas, for the input /dogs/, it is the *generator* that rules out *blick* by not producing it ($blick \notin$ Gen(dogs)). Therefore, to make the generator as neutral as possible in the determination of output forms, we assume that Gen is a constant function.[3] We denote this constant by X (the set of all possible outputs). For the theorems below to hold, X must be assumed to have at least three elements, which is reasonable as every language has at least three words. The outputs contained in X are denoted by $x, y, x_1, x_2, y_1, y_2, \ldots$.

Let $C = \{c_1, c_2, \ldots, c_k\}$ be a set of k *constraints*,[4] where each constraint, c_i, is a mapping

$$c_i : \text{set of all possible inputs} \rightarrow \text{set of all possible } \textit{evaluation functions}.$$

That is, c_i maps each input, Input, to an *evaluation function* (*for* Input), c_i[Input], where c_i[Input] is a mapping

$$c_i[\text{Input}] : X \rightarrow \mathbf{R}.$$

This evaluation function, c_i[Input], maps each generated output in X to a real number[5] representing the 'score', or 'local level of harmony', which constraint c_i assigns to that output given the input, Input.[6] The statement 'c_i[Input](x) ≥

---

[3] While this assumption simplifies the present mathematical exposition, essentially the same results can be proven without it.

[4] We assume that the set of constraints is finite and we apply social-choice-theoretic results for finite sets of 'voters'. However, social-choice-theoretic results can be generalised to infinite sets of 'voters'. See for example Heal (1997) and Efimov & Koshevoy (1994). Infinite constraint sets are suggested, for instance, by Smolensky (1995). Mathematically, the infinite case is no less complex than the finite.

[5] The choice of **R**, rather than, say, **Z**, does not imply any substantive assumption about the accuracy with which constraint violation can be evaluated. Such substantive assumptions are represented by the classes of transformations up to which constraint evaluation functions are taken to be unique (see Table 1, below).

[6] Though a very important question, we do not address how the evaluation function determines the score. The reader is referred to Eisner (1997b) for a discussion of the issues and a proposal using multi-argument mappings. We use single-argument mappings, and it is important to note that it is incorrect to think of 'c_i[Input](output)' as 'c_i(Input,output)'. The two-argument notation suggests that it is meaningful to ask whether, for instance, $c_i(\text{Input}_1,\text{output}_1) \geq c_i(\text{Input}_2,\text{output}_2)$. The one-argument notation,

c_i[Input](y)' is interpreted as "according to constraint c_i, output x is at least as harmonic as output y given the input Input". Below, we present a formal treatment of the question of how much information these 'scores' contain and how they can be interpreted. A *profile (of evaluation functions for a given input)* is a k-tuple, $\{c_i[\text{Input}]\} =_{def} \{c_1[\text{Input}], c_2[\text{Input}], ..., c_k[\text{Input}]\}$, with one c_i[Input] for each constraint.[7] This notation may appear to 'claim' that the constraints are dependent on the inputs. It does not. What depends upon the input is what 'scores' each constraint assigns to different outputs. Hence, the *evaluation function*, $c_i[\text{Input}] : X \rightarrow \mathbf{R}$, assigned to Input by constraint c_i, is input-dependent. The constraint itself is not. Because of the existence of constraints which regulate divergence between output and input, constraints cannot blindly evaluate just the candidate outputs, but must evaluate them in relation to the input. It is this sensitivity which the notation is supposed to reflect. We say more about the typology of constraints in Section 4.

A constraint aggregation function, F, is a function

F : some set of profiles of evaluation functions $\rightarrow$ set of all possible global harmony orderings.

Condition (U) below is the assumption that the domain of F is the set of all logically possible profiles of evaluation functions, but other domain assumptions are possible (see Sen, 1982a). (U) is discussed further in Section 4. F maps each profile of evaluation functions, $\{c_i[\text{Input}]\}$, to a global harmony ordering of X, R $=_{def} F(\{c_i[\text{Input}]\})$. R is assumed to be reflexive, connected and transitive, and xRy is interpreted to mean "output x is at least as globally harmonic as output y". R also induces a strong ordering P and an indifference relation I defined by

xPy	if and only if	xRy and not yRx,
xIy	if and only if	xRy and yRx.

We read xPy as "the output x is strictly more globally harmonic than the output y", and xIy as "the output x and the output y are equally globally harmonic"[8]. Before we can ask what kinds of constraint aggregation functions can be defined,

on the other hand, commits us only to the meaningfulness of comparisons in which the input is held constant (i.e., $\text{Input}_1 = \text{Input}_2$). The difference in the number of arguments our respective mappings take does not make our proposals incompatible as they concern different parts of the grammar.

[7] The use of curly brackets, { }, for profiles (k-tuples) of evaluation functions is standard in social choice theory. They are used to denote *ordered* sets.

[8] It may seem unnecessary to require that the constraint aggregation function determine *orderings* rather than just *winning candidates*. However, an alternative social-choice-theoretic framework uses *choice functions* rather than *orderings* for determining winning candidates. Provided some consistency requirements are imposed on these choice functions, equivalents of the results stated in the present paper hold in the alternative framework too. See, for example, Sen (1982b).

we need to discuss how the 'scores' assigned to the outputs by the constraints can be interpreted.

These issues were discussed in Section 2 as the problems of measurability and comparability. Rephrasing the social-choice-theoretic questions as questions about constraints and outputs, we have the following. The issue of measurability: How much significance do we attach to the numerical values of the 'scores'? Are they significant on an absolute scale? Or are only their ratios and/or differences significant? Or is the way in which they rank-order the possible outputs all that is meaningful? The issue of cross-constraint comparability: Are comparisons between the 'scores' assigned by different constraints meaningful? Is it sensible, for instance, to ask whether output x from the perspective of constraint c_i is at least as harmonic as output y from the perspective of constraint c_j?

Different assumptions about measurability and cross-constraint comparability can be stated formally by reference to classes of transformations. A transformation is simply a function, $\phi : \mathbf{R} \rightarrow \mathbf{R}$, which preserves certain structure. A positive monotonic transformation preserves the ordering of the real numbers: for all real numbers t_1, t_2, if $t_1 < t_2$ then $\phi(t_1) < \phi(t_2)$. A positive affine transformation preserves ratios of differences of real numbers: there exist a, b (b>0) such that, for all real numbers t, $\phi(t) = a + bt$.

Now, the measurement of any profile of evaluation functions will be assumed to be unique up to certain transformations but not up to others. The class of these transformations covaries with how much information the profile contains and how that information is described. The smaller this class of transformations, the more information is contained in a profile. An aggregation function satisfies a given assumption concerning measurability and comparability if the aggregation function is invariant under any member of the class of transformations corresponding to that assumption.

Below, we give a number of possible assumptions about measurability and comparability that have been proposed in the social-choice-theoretic literature.[9] The table will be clearer when the reader is familiar with the formalism and may be disregarded initially.

[9] Numerous other possible assumptions have been studied. The reader is referred to Roberts (1980b), Sen (1982c), Blackorby & Donaldson (1984), and Tsui & Weymark (1997).

(3) Table 1

For any two profiles of evaluation functions, $\{c_i[\text{Input}]\}$ and $\{c'_i[\text{Input}]\}$, in the domain of F, it is required that $F(\{c_i[\text{Input}]\}) = F(\{c'_i[\text{Input}]\})$ if ...
ONC: ordinal measurability, no cross-constraint comparability ... $c'_i[\text{Input}] = \phi_i(c_i[\text{Input}])$, for some k-tuple $\{\phi_i\}$ of positive monotonic transformations
CNC: cardinal measurability, no cross-constraint comparability ... $c'_i[\text{Input}] = a_i + b_i c_i[\text{Input}]$, for some k-tuples $\{a_i\}$, $\{b_i\}$ of real numbers ($b_i > 0$)
OLC: ordinal measurability, cross-constraint comparability of local harmony levels ... $c'_i[\text{Input}] = \phi(c_i[\text{Input}])$, for some positive monotonic transformation ϕ: $\mathbf{R} \rightarrow \mathbf{R}$
CUC: cardinal measurability, cross-constraint comparability of local harmony units ... $c'_i[\text{Input}] = a_i + bc_i[\text{Input}]$, for some k-tuples $\{a_i\}$ of real numbers and a real number $b>0$
CFC: cardinal measurability, cross-constraint comparability of local harmony levels & units ... $c'_i[\text{Input}] = a + bc_i[\text{Input}]$, for some real numbers a and b ($b>0$)

Consider four types of information we might want to extract from a profile of evaluation functions:

(i) 'Is the violation of constraint c_i by candidate x at least as great as its violation by candidate y?'

(ii) 'What is the ratio of score-of-x_1-minus-score-of-x_2 to score-of-y_1-minus-score-of-y_2 according to the same constraint?'

(iii) 'Is the violation of constraint c_i by candidate x at least as great as the violation of constraint c_j by candidate y?'

(iv) 'What is the ratio of score-of-x_1-minus-score-of-x_2-according-to-constaint-c_i to score-of-y_1-minus-score-of-y_2-according-to-constraint-c_2?'

According to (ONC), only questions of type (i) are meaningful.[10] According to (CNC), only questions of types (i) and (ii). According to (OLC), only questions of types (i) and (iii). According to (CUC), only questions of types (i), (ii) and (iv). And according to (CFC), only questions of types (i)–(iv).

[10] To be precise, when we say '(X) is the assumption that *only* questions of type Y are meaningful' we mean 'according to (X), a profile of evaluation functions is unique up to the largest class of transformations under which answers to questions of type Y are invariant'.

We can now ask which constraint aggregation functions are definable. In order to formalise the requirement that global harmony orderings be sensitive to individual constraints, we shall state a number of conditions that can be imposed on constraint aggregation functions. They are versions of the famous minimal conditions in social choice theory proposed by Kenneth Arrow (1951, 1963). Their appropriateness is discussed in Section 4.

Universal Domain (U). The domain of F is the set of all logically possible profiles of evaluation functions.

Strong Pareto Principle (SP). Let $\{c_i[\text{Input}]\}$ be any profile of evaluation functions in the domain of F, let $R = F(\{c_i[\text{Input}]\})$, and let P denote the strong ordering induced by R. For any x_1, x_2 in X, we have x_1Rx_2, whenever $c_i[\text{Input}](x_1) \geq c_i[\text{Input}](x_2)$, for all i; if, in addition, there exists an i such that $c_i[\text{Input}](x_1) > c_i[\text{Input}](x_2)$, we have x_1Px_2.

Independence of Irrelevant Alternatives (I). Let $\{c_i[\text{Input}]\}$ and $\{c'_i[\text{Input}]\}$ be profiles of evaluation functions in the domain of F, and let $R = F(\{c_i[\text{Input}]\})$ and $R' = F(\{c'_i[\text{Input}]\})$. Suppose x_1 and x_2 are elements of X such that, for all i, $c_i[\text{Input}](x_1) = c'_i[\text{Input}](x_1)$ and $c_i[\text{Input}](x_2) = c'_i[\text{Input}](x_2)$. Then x_1Rx_2 if and only if $x_1R'x_2$.

To paraphrase, suppose candidate x_1 is assigned the same scores by two different profiles of evaluation functions and that candidate x_2 is likewise assigned the same scores under each profile. (I) then says that these two candidates must be ranked identically in each case. That is, the aggregation function, focusing exclusively on x_1's and x_2's scores when ranking them with respect to each other, must rank them identically in each case. In particular, how other candidates' scores vary under the two profiles is irrelevant. So, the aggregation function does not take any third candidate into account when ranking x_1 with respect to x_2.

Optimality Theory uses a strict dominance hierarchy of the constraints.

Definition 1. A constraint aggregation function, F, is a *strict dominance hierarchy* of the constraints if there exists a *fixed* permutation σ of $\{1, 2, \ldots, k\}$ such that, for *any* profile $\{c_i[\text{Input}]\}$in the domain of F and *any* x_1 and x_2 in X,

x_1Px_2 if and only if
$c_{\sigma(j)}[\text{Input}](x_1) > c_{\sigma(j)}[\text{Input}](x_2)$ for some $j \in \{1, 2, \ldots, k\}$
and $c_{\sigma(i)}[\text{Input}](x_1) = c_{\sigma(i)}[\text{Input}](x_2)$ for all $i<j$,

where P is the strong ordering induced by $R = F(\{c_i[\text{Input}]\})$.

We can now relate these definitions back to a more typical optimality-theoretic view of phonology. The process of selecting the output corresponding to a given input would run as follows. The input (say, /ıN-put/) would be fed into the generator, which produces X, the list of all possible outputs. Each candidate in X is then evaluated with respect each constraint. That is, there are k constraints, $c_1, c_2, \ldots, c_k$. (The subscripts do not reflect any ranking. The numbering does not change cross-linguistically, though rankings, it is assumed, do. If c_1 denotes NOCODA and c_2 denotes FAITH(high)_{IO}, then they denote them in Universal Grammar. Languages would then choose how to rank them: $c_1 \gg c_2$, or $c_2 \gg c_1$. The ranking is represented by a permutation σ of $\{1, 2, \ldots, k\}$.) Corresponding to each constraint, there is an evaluation function, c_i[ıN-put]. Each evaluation function maps each candidate in X to a real number. So, if c_1 is NOCODA, the candidate *im.pʰut* would be mapped by c_1[ıN-put] to some real number. In the case of Optimality Theory, c_1[ıN-put](*im.pʰut*) = **, where '*' is one violation unit.

For any input, there exists a k-tuple of such evaluation functions, known as a profile. When every candidate in X has been evaluated by the profile of evaluation functions, {c_i[ıN-put]}, the aggregation of the scores begins. Recall that in traditional Optimality Theory, the constraints are arranged in a strict dominance hierarchy and that the optimal candidate is the one that outranks all others in pairwise comparisons. Given Definition 1, the optimality-theoretic claim that cross-linguistic variation is variation in languages' strict dominance hierarchies translates as variation in the specification of σ. To determine the optimal candidate, one proceeds as follows.[11] One takes two candidates, x and y. Suppose that constraint c_{23}, say ONSET, is highest ranked in the language under consideration. Then, in that language, $\sigma(1) = 23$; that is, σ permutes $\{1, 2, \ldots, k\}$ such that constraint c_{23} is first. Thus, in pairwise comparison of x and y, the first constraint we consider is what Universal Grammar labels as constraint c_{23}. More generally, the first constraint we consider is $c_{\sigma(1)}$, the second is $c_{\sigma(2)}$, the third $c_{\sigma(3)}$, and so on. If x outperforms y on $c_{\sigma(1)}$, then x is more optimal. That is, $c_{\sigma(1)}$[ıN-put](x) > $c_{\sigma(1)}$[ıN-put](y) entails xPy. If y outperforms x, then yPx. If neither outperforms the other, that is, if $c_{\sigma(1)}$[ıN-put](x) = $c_{\sigma(1)}$[ıN-put](y), we make the same comparison with respect to $c_{\sigma(2)}$. We continue in this fashion until we find an i, $1 \leq i \leq k$, such that x outperforms y, or vice versa, with respect to $c_{\sigma(i)}$. If there is no such constraint, we conclude xIy.

It is easily checked that a strict dominance hierarchy satisfies (U), (SP) and (I). Moreover, a version of Arrow's famous impossibility theorem (Arrow, 1951, 1963) entails that strict dominance hierarchies are the *only* constraint

[11] The method described is not efficient. Indeed, if the candidate list is infinite, no decision can be reached in this fashion. However, we are keeping to a close reading of Prince & Smolensky (1993). The real question of how optimality-theoretic selection processes could be effective lies beyond the scope of this paper. See Eisner (1997) for a solution.

aggregation functions satisfying (U), (SP), (I) and at least one of (ONC) and (CNC).

Theorem 1.[12] A constraint aggregation function, F, is a strict dominance hierarchy if and only if it satisfies (U), (SP), (I) and at least one of (ONC) and (CNC).

If, on the other hand, we are prepared to introduce relevant forms of cross-constraint comparability, other ways to define constraint aggregation functions emerge.

Definition 2. For each $x \in X$, define a permutation $pos[x] : \{1, 2, \ldots k\} \rightarrow \{1, 2, \ldots k\}$ such that

$$c_{pos[x](1)}[\text{Input}](x) \leq c_{pos[x](2)}[\text{Input}](x) \leq \ldots \leq c_{pos[x](k)}[\text{Input}](x).$$

A constraint aggregation function, F, is a *positional dominance hierarchy*[13] if there exists a *fixed* permutation σ of $\{1, 2, \ldots, k\}$ such that, for *any* profile {c_i[Input]} in the domain of F and for *any* x_1 and x_2 in X,

$x_1 P x_2$ if and only if

$$c_{pos[x_1](\sigma(j))}[\text{Input}](x_1) > c_{pos[x_2](\sigma(j))}[\text{Input}](x_2), \text{ some } j \in \{1,2,\ldots,k\}$$
$$\text{and } c_{pos[x_1](\sigma(i))}[\text{Input}](x_1) = c_{pos[x_2](\sigma(i))}[\text{Input}](x_2) \text{ for all } i<j,$$

where P is the strong ordering induced by R = F({c_i[Input]}).

The following is a particularly interesting version of a positional dominance hierarchy:

Definition 3. A constraint aggregation function, F, is the *leximin rule* if it is the positional dominance hierarchy where σ is the identity permutation.

According to the leximin rule, output x is preferred to output y whenever the smallest local harmony level of x is greater than the smallest local harmony level of y; ties are broken by consecutive comparisons of the second smallest, third smallest, ... local harmony levels. More concretely, the global harmony ordering would be determined as follows. Take a tableau. Erase the column headings (the constraint names, $c_1, c_2, \ldots, c_k$). Within each row, re-order the cells so that there is no cell that contains more violations than any cell to its left. Then label the columns 1 through to k and rank the candidates treating 1 to k as a strict dominance hierarchy ($1 \gg 2 \gg \ldots \gg k$) to determine the optimal candidate. Note, however, that, unlike in a proper strict dominance hierarchy, entries in the same column will not in general correspond to the same constraint. To determine the global harmony ordering in a positional dominance hierarchy,

[12] This result follows immediately from theorem 3.3.6. in d'Aspremont (1985).

[13] The term *positional dominance hierarchy* is derived from the term for its social-choice-theoretic counterpart *positional dictatorship*.

erase headings and re-order each row as for the leximin rule. Relabel the columns 1 to k and permute them according to σ. Rank the candidates treating $\sigma(1), \sigma(2), \ldots, \sigma(k)$ as a strict dominance hierarchy ($\sigma(1) \gg \sigma(2) \gg \ldots \gg \sigma(k)$).

It is easily checked that positional dominance hierarchies, in addition to strict dominance hierarchies, satisfy (U), (SP), (I) and (OLC). However, the range of possible constraint aggregation functions satisfying (U), (SP), (I) and (OLC) is still limited. Roberts (1980a) has shown that the only aggregation functions satisfying these conditions are *single-focus rules*. Informally, such rules lexicographically focus exclusively on single scores amongst $c_1[\text{Input}](x)$, $c_2[\text{Input}](x)$, …, $c_k[\text{Input}](x)$, each 'focus' being either 'fixed' (as in strict dominance hierarchies) or 'positional' (as in positional dominance hierarchies). Combinations of strict dominance hierarchies and positional dominance hierarchies also satisfy (U), (SP), (I) and (OLC). It is, for instance, conceivable that certain constraints might take absolute priority, but ties might be broken by a suitably defined positional dominance hierarchy, or, conversely, that a positional dominance hierarchy for a number of positions less than k takes absolute priority and ties are then broken by a strict dominance hierarchy. If we invoke an additional axiom to rule out the strict-dominance-hierarchy-component of such rules, positional dominance hierachies are the only remaining possibilities. The axiom requires aggregation functions to treat all constraints identically, that is, to be invariant under permutation of the constraints.

SYMMETRY (S). Let $\{c_i[\text{Input}]\}$ be any profile of evaluation functions, and let σ be any permutation of $\{1, 2, \ldots, k\}$. Then $F(\{c_i[\text{Input}]\}) = F(\{c_{\sigma(i)}[\text{Input}]\})$.

Theorem 2.[14] A constraint aggregation function, F, is a positional dominance hierarchy if and only if it satisfies (U), (SP), (I), (S) and (OLC).

If we admit cardinal measurablity, further possibilities emerge.

Definition 4. A constraint aggregation function, F, is a *weak utilitarian rule* if, for every non-empty subset L of $\{1, 2, \ldots, k\}$, there exist *fixed* non-negative real numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$, not all zero, but with $\lambda_i=0$ for all $i \notin L$, such that, for *any* profile $\{c_i[\text{Input}]\}$ in the domain of F and for *any* x_1, x_2 in X with $c_i[\text{Input}](x_1) = c_i[\text{Input}](x_2)$ for all $i \notin L$,

$$\textstyle\sum_{i\in\{1,2,\ldots,k\}} \lambda_i c_i[\text{Input}](x_1) > \sum_{i\in\{1,2,\ldots,k\}} \lambda_i c_i[\text{Input}](x_2) \text{ implies } x_1 P x_2,$$

where P is the strong ordering induced by $R = F(\{c_i[\text{Input}]\})$.

The definition of a *weak utilitarian rule* is quite broad. Even a strict dominance hierarchy can be interpreted as a weak utilitarian rule, with exactly one non-zero λ_i for each non-empty subset L of $\{1, 2, \ldots, k\}$. The rather complex conditions involving L are necessary for breaking ties. For example,

[14] This result follows immediately from theorem 3.4.7. in d'Aspremont (1985).

consider the following weak utilitarian rule F, where the number of constraints $k = 5$. For *any* profile $\{c_i[\text{Input}]\}$ in the domain of F and *any* x_1 and x_2 in X,

$$
\begin{aligned}
x_1 P x_2 \text{ if and only if } \quad & 2c_3[\text{Input}](x_1)+5c_4[\text{Input}](x_1) > \\
& 2c_3[\text{Input}](x_2)+5c_4[\text{Input}](x_2) \\
\text{or} \quad & [2c_3[\text{Input}](x_1)+5c_4[\text{Input}](x_1) \\
& =2c_3[\text{Input}](x_2)+5c_4[\text{Input}](x_2) \\
& \text{and } c_5[\text{Input}](x_1) > c_3[\text{Input}](x_1)] \\
\text{or} \quad & [2c_3[\text{Input}](x_1)+5c_4[\text{Input}](x_1) \\
& =2c_3[\text{Input}](x_2)+5c_4[\text{Input}](x_2) \\
& \text{and } c_5[\text{Input}](x_1) = c_3[\text{Input}](x_1) \\
& \text{and } c_1[\text{Input}](x_1)+3c_2[\text{Input}](x_1) > \\
& c_1[\text{Input}](x_1)+3c_2[\text{Input}](x_1)]
\end{aligned}
$$

where P is the strong ordering induced by $R = F(\{c_i[\text{Input}]\})$. Note that, for example, for $L = \{1, 2, 3, 4, 5\}$, we have $\lambda_1=0$, $\lambda_2=0$, $\lambda_3=2$, $\lambda_4=5$, $\lambda_5=0$. For $L = \{1, 2, 5\}$, $\lambda_1=0$, $\lambda_2=0$, $\lambda_3=0$, $\lambda_4=0$, $\lambda_5=1$. For $L = \{1, 2\}$, $\lambda_1=1$, $\lambda_2=3$, $\lambda_3=0$, $\lambda_4=0$, $\lambda_5=0$. Weak utilitarian rules are the *only* constraint aggregation functions satisfying (U), (SP), (I) and (CUC).

Theorem 3.[15] A constraint aggregation function, F, is a weak utilitarian rule if and only if it satisfies (U), (SP), (I) and (CUC).

If we admit cardinal measurability with cross-constraint comparability of both local harmony levels and units, i.e., if we impose (CFC), the two classes of constraint aggregation functions we have discussed so far, positional dominance hierarchies and weak utilitarian rules, will clearly satisfy (U), (SP) and (I), but further possibilities also emerge. The following classification result holds:

Theorem 4.[16] If a constraint aggregation function, F, satisfies (U), (SP), (I) and (CFC), then there exists a positively linearly homogeneous function $g : \mathbf{R}^k \to \mathbf{R}$ (that is, $g(\lambda t) = \lambda g(t)$ for all $t \in \mathbf{R}^k$ and all $\lambda>0$), such that, for all x_1, x_2 in X,

$$u(\{c_i[\text{Input}](x_1)\}) + g(\{c_i[\text{Input}](x_1) - u(\{c_i[\text{Input}](x_1)\})\}) >$$
$$u(\{c_i[\text{Input}](x_2)\}) + g(\{c_i[\text{Input}](x_2) - u(\{c_i[\text{Input}](x_2)\})\})$$

implies $x_1 P x_2$, where $u(\{c_i[\text{Input}](x)\}) =_{def} {}^1/_k \sum_{i\in\{1,2,\ldots,k\}} c_i[\text{Input}](x)$, and P is the strong ordering induced by $R = F(\{c_i[\text{Input}]\})$.

Roberts (1980b) provides an example of this type of function. Define $g(t) = g(\{t_1, t_2, \ldots, t_k\}) =_{def} \alpha \min_{i\in\{1,2,\ldots,k\}} t_i$, for $t \in \mathbf{R}^k$ and $0 \leq \alpha \leq 1$. For $\alpha = 0$, the corresponding constraint aggregation function is a weak utilitarian rule, with $\lambda_1 =$

[15] This result follows immediately from theorem 3.3.3. in d'Aspremont (1985).

[16] This result follows immediately from theorem 3.5.1. in d'Aspremont (1985).

$\lambda_2 = ... = \lambda_k = {}^1/_k$; for $\alpha = 1$, it is the leximin rule; for $0 < \alpha < 1$, it is a weighted combination of these two rules. Note, however, that Theorem 4 is silent on how ties are broken by constraint aggregation functions satisfying its conditions. This does not entail that these functions do not have tie-breaking procedures. The theorem is simply not sufficiently strong to tell us what they are. Stronger theorems, using stronger conditions, have been proven (for instance, d'Aspremont, 1985).

4. Interpretation of the Formalism

This section examines the consequences of various theoretical implementations of the proposal that phonological phonomena are regulated by constraint interaction. That is, we will now examine what the formal results of the preceding section hold for linguist(ic)s. The obvious starting point, Section 4.1, is to discuss the assumption that constraint aggregation functions must satisfy (U), (SP), and (I). As the ways in which they restrict the possible shape of constraint-based theories of phonology become apparent, we consider the consequences of abandoning them or mollifying their effect. We will also argue that the nature of constraints, ignored above, is crucial to the possibility of mollifying the requirements on aggregation functions. The typology of constraints is addressed in Section 4.3. In Section 4.2, we examine several proposals that slightly modify Optimality Theory as presented by Prince & Smolensky (1993) and show that they are characterisable in terms of the formalism in Section 3. In Section 4.4, we discuss the problem of opacity and how it relates to different constraint typologies and measurement and cross-constraint comparability assumptions.

4.1. Conditions on Constraint Aggregation Functions

We wish to emphasise at the outset that we cannot prove (U), (SP), and (I) indispensible to (aggregation in) constraint-based phonology. Our comments below are intended as broad motivation for linguistic correlates of common social-choice-theoretic assumptions concerning aggregation functions, which further research may call into question.

We imposed three conditions on aggregation functions.

(U) requires that a global harmony ordering be defined for any profile of constraint evaluation functions. Given our set of constraints, C, the issue is whether, for any tableau of scores, there exists an input, Input, such that the constraint evaluation functions in the profile $\{c_i[\text{Input}]\}$, applied to each output in X, produce precisely that tableau. If condition (I) is also invoked (as defended below), the issue simplifies to whether, for any k-tuple of scores, $\{t_i\}$, there exists an input, Input, and an output in X, Output, such that $\{c_i[\text{Input}](\text{Output})\} = \{t_i\}$. It is difficult to give a definitive answer to these questions in the absence of greater information about the constraints.

(SP) imposes a condition on the relationship between candidates' scores and their degree of harmony. It states that, for any candidates, x_1 and x_2, x_1 is at least as globally harmonic as x_2 if there is no constraint with respect to which x_1 is less harmonic than x_2. Moreover, x_1 is more globally harmonic if there is a constraint with respect to which x_1 is more harmonic than x_2.

(I) forces the aggregation function, when ranking two candidates, to consider those candidates' profiles and no others'. If it is "[v]ia pair-wise comparison of alternative analyses" that "the grammar imposes a harmonic order on the entire set of possible analyses of a given underlying form" (Prince & Smolensky, 1993, p. 3), then all that should matter in any pairwise consideration of candidates is their profiles.

The case for (I) and (SP) can be made more compelling by noting a property that they jointly entail (Proposition 6, below). Firstly, observe that if a constraint aggregation function satisfies (SP), it also satisfies (PI).

PARETO INDIFFERENCE (PI). Let $\{c_i[\text{Input}]\}$ be any profile of evaluation functions in the domain of F, and let I denote the indifference relation induced by $R = F(\{c_i[\text{Input}]\})$. For any x_1, x_2 in X, we have $x_1 I x_2$, whenever, for all i, $c_i[\text{Input}](x_1) = c_i[\text{Input}](x_2)$.

Secondly, consider the condition (INV) below. Informally paraphrased, it requires that aggregation functions use nothing but candidates' scores to determine their global harmony ranking.

INDEPENDENCE OF NON-CONSTRAINT-VIOLATION CHARACTERISTICS (INV). Let $\{c_i[\text{Input}]\}$ and $\{c'_i[\text{Input}]\}$ be profiles of evaluation functions in the domain of F, and let x_1, x_2, y_1, y_2 be candidate outputs such that, for all i, $c_i[\text{Input}](x_1) = c'_i[\text{Input}](x_2)$ and $c_i[\text{Input}](y_1) = c'_i[\text{Input}](y_2)$. Then $x_1 R y_1$ if and only if $x_2 R' y_2$, where $R = F(\{c_i[\text{Input}]\})$ and $R' = F(\{c'_i[\text{Input}]\})$.

Proposition 5.[17] If a constraint aggregation function, F, satisfies (U), (PI) and (I), then it also satisfies (INV).

Pure constraint-based phonology is committed to the claim that the evaluation scores of two candidates in X contain sufficient information to determine their position relative to each other in the global harmony ordering. (INV) formalises this claim. We use *evaluationism* to refer to the view that evaluation scores contain all the information necessary for determining global harmony orderings. Pure constraint-based theories are therefore evaluationist. We said that the problems of constraint aggregation in linguistics and preference aggregation in social choice theory are very similar. The similarity continues here. Evaluationism is the linguistic counterpart of *welfarism* in social choice

[17] See Roemer (1996, pp. 28–29).

$\lambda_2 = ... = \lambda_k = {}^1/_k$; for $\alpha = 1$, it is the leximin rule; for $0 < \alpha < 1$, it is a weighted combination of these two rules. Note, however, that Theorem 4 is silent on how ties are broken by constraint aggregation functions satisfying its conditions. This does not entail that these functions do not have tie-breaking procedures. The theorem is simply not sufficiently strong to tell us what they are. Stronger theorems, using stronger conditions, have been proven (for instance, d'Aspremont, 1985).

4. Interpretation of the Formalism

This section examines the consequences of various theoretical implementations of the proposal that phonological phonomena are regulated by constraint interaction. That is, we will now examine what the formal results of the preceding section hold for linguist(ic)s. The obvious starting point, Section 4.1, is to discuss the assumption that constraint aggregation functions must satisfy (U), (SP), and (I). As the ways in which they restrict the possible shape of constraint-based theories of phonology become apparent, we consider the consequences of abandoning them or mollifying their effect. We will also argue that the nature of constraints, ignored above, is crucial to the possibility of mollifying the requirements on aggregation functions. The typology of constraints is addressed in Section 4.3. In Section 4.2, we examine several proposals that slightly modify Optimality Theory as presented by Prince & Smolensky (1993) and show that they are characterisable in terms of the formalism in Section 3. In Section 4.4, we discuss the problem of opacity and how it relates to different constraint typologies and measurement and cross-constraint comparability assumptions.

4.1. Conditions on Constraint Aggregation Functions

We wish to emphasise at the outset that we cannot prove (U), (SP), and (I) indispensible to (aggregation in) constraint-based phonology. Our comments below are intended as broad motivation for linguistic correlates of common social-choice-theoretic assumptions concerning aggregation functions, which further research may call into question.

We imposed three conditions on aggregation functions.

(U) requires that a global harmony ordering be defined for any profile of constraint evaluation functions. Given our set of constraints, C, the issue is whether, for any tableau of scores, there exists an input, Input, such that the constraint evaluation functions in the profile $\{c_i[\text{Input}]\}$, applied to each output in X, produce precisely that tableau. If condition (I) is also invoked (as defended below), the issue simplifies to whether, for any k-tuple of scores, $\{t_i\}$, there exists an input, Input, and an output in X, Output, such that $\{c_i[\text{Input}](\text{Output})\} = \{t_i\}$. It is difficult to give a definitive answer to these questions in the absence of greater information about the constraints.

(SP) imposes a condition on the relationship between candidates' scores and their degree of harmony. It states that, for any candidates, x_1 and x_2, x_1 is at least as globally harmonic as x_2 if there is no constraint with respect to which x_1 is less harmonic than x_2. Moreover, x_1 is more globally harmonic if there is a constraint with respect to which x_1 is more harmonic than x_2.

(I) forces the aggregation function, when ranking two candidates, to consider those candidates' profiles and no others'. If it is "[v]ia pair-wise comparison of alternative analyses" that "the grammar imposes a harmonic order on the entire set of possible analyses of a given underlying form" (Prince & Smolensky, 1993, p. 3), then all that should matter in any pairwise consideration of candidates is their profiles.

The case for (I) and (SP) can be made more compelling by noting a property that they jointly entail (Proposition 6, below). Firstly, observe that if a constraint aggregation function satisfies (SP), it also satisfies (PI).

PARETO INDIFFERENCE (PI). Let $\{c_i[\text{Input}]\}$ be any profile of evaluation functions in the domain of F, and let I denote the indifference relation induced by $R = F(\{c_i[\text{Input}]\})$. For any x_1, x_2 in X, we have $x_1 I x_2$, whenever, for all i, $c_i[\text{Input}](x_1) = c_i[\text{Input}](x_2)$.

Secondly, consider the condition (INV) below. Informally paraphrased, it requires that aggregation functions use nothing but candidates' scores to determine their global harmony ranking.

INDEPENDENCE OF NON-CONSTRAINT-VIOLATION CHARACTERISTICS (INV). Let $\{c_i[\text{Input}]\}$ and $\{c'_i[\text{Input}]\}$ be profiles of evaluation functions in the domain of F, and let x_1, x_2, y_1, y_2 be candidate outputs such that, for all i, $c_i[\text{Input}](x_1) = c'_i[\text{Input}](x_2)$ and $c_i[\text{Input}](y_1) = c'_i[\text{Input}](y_2)$. Then $x_1 R y_1$ if and only if $x_2 R' y_2$, where $R = F(\{c_i[\text{Input}]\})$ and $R' = F(\{c'_i[\text{Input}]\})$.

Proposition 5.[17] If a constraint aggregation function, F, satisfies (U), (PI) and (I), then it also satisfies (INV).

Pure constraint-based phonology is committed to the claim that the evaluation scores of two candidates in X contain sufficient information to determine their position relative to each other in the global harmony ordering. (INV) formalises this claim. We use *evaluationism* to refer to the view that evaluation scores contain all the information necessary for determining global harmony orderings. Pure constraint-based theories are therefore evaluationist. We said that the problems of constraint aggregation in linguistics and preference aggregation in social choice theory are very similar. The similarity continues here. Evaluationism is the linguistic counterpart of *welfarism* in social choice

[17] See Roemer (1996, pp. 28–29).

theory, the view that no information beyond measurements of individuals' welfare is necessary in determining social orderings of policy options.

Moreover, if we accept evaluationism, we are also forced to accept (PI) and (I):

Proposition 6. If a constraint aggregation function, F, satisfies (INV), then it also satisfies (PI) and (I).

Proof. (1) To prove that (INV) entails (PI). Let $\{c_i[\text{Input}]\} = \{c'_i[\text{Input}]\}$, $x_1=y_2$ and $x_2=y_1$. Further, assume that, for all i, $c_i[\text{Input}](x_1) = c_i[\text{Input}](x_2)$, to satisfy the conditions of (PI). Then the conditions of (INV) are met, and x_1Rx_2 if and only if y_1Ry_2, i.e., if and only if x_2Rx_1. But since R is connected, we must have both x_1Rx_2 and x_2Rx_1, i.e., x_1Ix_2, as required. (2) To prove that (INV) entails (I). Let $x_1=x_2$ and $y_1=y_2$. Further, to satisfy the conditions of (I), assume that $c_i[\text{Input}](x_1) = c'_i[\text{Input}](x_1)$ and $c_i[\text{Input}](y_1) = c'_i[\text{Input}](y_1)$, for all i. Then x_1Ry_1 if and only if $x_1R'y_1$, as required.

We can now see that the assumption of (PI) and (I) follows from constraint-based phonology's commitment to evaluationism. In this light, let us reëxamine the three conditions invoked above, (U), (SP) and (I).

(I) is least objectionable, given the evaluationist spirit of constraint-based phonology.

(SP) is justifiable at least to the extent that it is related to (PI). In fact, (SP) is slightly stronger. (PI) asserts only that candidates with identical scores should be ranked indifferently in the global harmony ordering. If we invoke only (PI), then a candidate x could be more globally harmonic than a candidate y even if, for all $i\in\{1,\ldots 2,\ldots,k\}$, x's number of violations on constraint c_i were greater than or equal to y's number of violations on constraint c_i, and, for some i, were *strictly* greater. Therefore, the rejection of (SP) seems tantamount to the rejection of the basic principle behind constraint conflict resolution: the fewer the violations, the more harmonic the candidate. However, there are also weaker conditions that preserve some of this principle behind conflict resolution. The weak Pareto principle, (P), for instance, requires that, if candidate x is *strictly* better than y on *every* constraint, then x must be ranked above y in the global harmony ordering. The condition of monotonicity, (M), requires that, if two profiles are identical except that one constraint has 'raised' the ranking of a candidate x in relation to the other candidates, then x's position in the global harmony ordering cannot fall. If (SP) is replaced with (P), impossibility theorems in essence identical to those in Section 3 still hold (Roberts, 1980b). Similar results still hold if (SP) is replaced with (M) and the requirement that the constraint aggregation function be non-constant, i.e., sensitive to the constraints (Arrow, 1951, 1963).

In light of the evaluationist commitment to (I), (U) is justifiable to the extent that we have no reason to rule out that, for any k-tuple of scores, $\{t_i\}$, there exists an input, Input, and an output in X, Output, such that $\{c_i[\text{Input}](\text{Output})\} = \{t_i\}$. However, there is some indication that not every logically possible assignment of scores to a candidate is possible for real constraints. The assessment of a candidate by different constraints may not be independent: for instance, a candidate cannot violate DEP(F) more than it violates *F, where F is any feature; and exactly one of the constraints "X" and "NOT X" will be violated, where X is a property of whole candidates, not of substrings of candidates. Moreover, a constraint's behaviour across different candidates may be constrained too: for constraints that can be computationally realized by finite-state machines (Ellison, 1994; Eisner 1997a) or weighted context-free machines (Tesar, 1996), the number of violations cannot grow more than linearly with the length of a candidate. But at present it is an open question whether or not such interdependencies across rows and/or columns of tableaus are sufficiently systematic to rule out the possible occurrence of certain profiles and thus to warrant domain restrictions of constraint aggregation functions.

(It is important to note that systematic domain restrictions would provide an escape-route from the impossibility result under ordinal or cardinal measurability without cross-constraint comparability (Theorem 1). As an example, consider the condition of *triple-wise value-restriction*. A profile of constraint evaluation functions, $\{c_i[\text{Input}]\}$, is *triple-wise value-restricted* if, for every triple of candidate outputs in X, there exists at least one alternative amongst this triple, x, and at least one position, j^{th} (j=1: 'most harmonic'; j=2: 'medium harmonic'; j=3: 'least harmonic'), such that *all* the constraint evaluation functions $c_i[\text{Input}]$ in $\{c_i[\text{Input}]\}$ 'agree' that candidate x is *not* j^{th} amongst the given triple.[18] Sen (1982a) has established the following result:

Theorem 7.[19] There exist constraint aggregation functions on the domain of triple-wise value-restricted profiles of constraint evaluation functions satisfying (U), (SP), (I) and (ONC).

An example is the following constraint aggregation function F: for any triple-wise value-restricted profile $\{c_i[\text{Input}]\}$, let $R = F(\{c_i[\text{Input}]\})$ be the ordering such that, for any x_1, x_2 in X,

$$x_1Rx_2 \quad \text{if and only if} \quad |\{i\in\{1, 2, \ldots, k\} : c_i[\text{Input}](x_1) \geq c_i[\text{Input}](x_2)\}| \geq |\{i\in\{1, 2, \ldots, k\} : c_i[\text{Input}](x_2) \geq c_i[\text{Input}](x_1)\}|.$$

Informally, x_1 is at least as globally harmonic as x_2 if and only if x_1 beats, or at least ties with, x_2 in a simply majority ballot across all constraints.

---

18 We adopt the convention that, if different candidates amongst a triple are assigned identical scores by $c_i[\text{Input}]$, then each candidate has more than one position (i.e., 1^{st}, 2^{nd} or 3^{rd}) amongst this triple according to $c_i[\text{Input}]$.

19 This result is entailed by (though not identical to) theorem 1 in Sen (1982a).

However, in the absence of a clearer empirical justification for systematic domain restrictions, these considerations must remain hypothetical.)

In light of the preceding, we conclude that (I) and (SP) are very strongly in the spirit of constraint-based phonology, and that (U) is a reasonable null hypothesis concerning the domain of aggregation functions. We can now proceed to ask what the significance of the theorems of Section 3 is. Of those theorems, the first makes the most demanding assumption concerning invariance and thus the least demanding assumption concerning measurability and cross-constraint comparability, and the last, the least demanding assumption concerning invariance and thus the most demanding assumption concerning measurability and cross-constraint comparability. The intermediary theorems represent assumptions of intermediary strength.

We consider two types of measurement: ordinal and cardinal. Ordinal measurability assumes that only the ordering between scores of different candidates is important. By contrast, cardinal measurability assumes that the scores – in particular, the units of the scores – have additional numerical significance. The most restrictive assumptions concerning measurability and cross-constraint comparability are that there is no comparability and that measurement is merely ordinal (ONC). Strict dominance hierarchies are the only aggregation functions satisfying (U), (SP), (I) and (ONC). They are also the only functions satisfying (U), (SP), (I) and (CNC), the result of admitting cardinal measurability. This yields the situation with which we are familiar in Optimality Theory. Strict dominance hierarchies are the only logical possibility for an aggregation function given the basic assumptions. However, alternatives are current in the literature. We now examine some of these.

4.2. Variants on Optimality Theory

As optimality-theoretic research has expanded and encountered empirical challenges, researchers have suggested modifying the version of Optimality Theory presented in Prince & Smolensky (1993). These modified theories diverge also from the system outlined above: a strict dominance hierarchy of output and input-output constraints. Some of these modifications concern the typology of constraints and are discussed in Section 4.3, particularly in relation to the problem of *opacity* (to be defined). Others concern measurability and cross-constraint comparability assumptions and are discussed immediately below. With regard to these assumptions, it is striking that the apparently quite different proposals discussed below actually have a single unifying feature: that they diverge from standard Optimality Theory in precisely what they assume about the nature of measurability and cross-constraint comparability. Our ability to characterise the nature of this variation in terms of the concepts and formalism presented in Section 3 constitutes, we believe, strong support for one of the general claims of this paper, namely, that the nature of measurability and cross-constraint comparability is an interesting and important area for research in constraint-based phonology.

4.2.1. Different Measurability and Cross-Constraint Comparability Assumptions

A not uncommon situation in optimality-theoretic reasearch is that part of a data set cannot be accounted for under any strict dominance hierarchy of an otherwise successful set of constraints. In this situation, researchers have sometimes suggested, not a revision of the constraint set, but a revision of how global harmony is determined given the number of times each candidate violates each constraint. Below, we discuss six such proposals and argue that what is really being modified in each case is the assumed nature of measurability and cross-constraint comparability. The examples we discuss are Burzio (1999), Crowhurst & Hewitt (1999), Flemming (1997), Itô & Mester (1997), Kirchner (1996) and Smolensy (1995). Since the discussion focuses on constraint aggregation after the number of violations has been calculated, the meaning of each constraint is irrelevant. We therefore omit their definitions and refer the reader to the relevant article in each case.

In the Tokyo dialect of Japanese, a compound of two words can sometimes have two acceptable surface realisations. Itô & Mester (1997) suggest the use of "tied rankings" to account for this "surface indeterminacy". "Tied ranking means that violations of the two constraints IdentSS and *g count as equivalent: It is just as bad to violate IdentSS as it is to violate *g" (p. 433). More specifically, consider (4), adapted from Itô & Mester's (26c) (p. 434).

(4) TOKYO JAPANESE TABLEAU

Lex = /niwa-Geta/ S = [*geta*]	*[ŋ	IdentSS	*g	IdentLexS
☞ [niwa *geta*]			*	*
☞ [niwa *ŋeta*]		*		*

Clearly, if IdentSS and *g are in a strict dominance relation, then only one of *niwa geta* and *niwa ŋeta* will be considered optimal. If the constraints are equally ranked, however, the correct result, surface indeterminacy, follows.

Precisely the required ranking can be achieved under Definition 4, using a (version of a) weak utilitarian rule. Partition {1, 2, ..., k} (corresponding to the k constraints in C) into $S_1, S_2, ..., S_n$ (i.e., $S = S_1 \cup S_2 ... \cup S_n$ where $S_i \cap S_j = \varnothing$ whenever $i \neq j$) and define $F(\{c_i[\text{Input}]\})$ to be the ordering R such that

$$
\begin{aligned}
x_1 P x_2 \quad & \text{if and only if} \\
& \textstyle\sum_{i \in S} \lambda_i c_i[\text{Input}](x_1) > \sum_{i \in S} \lambda_i c_i[\text{Input}](x_2) \\
\text{or } & [\textstyle\sum_{i \in S} \lambda_i c_i[\text{Input}](x_1) = \sum_{i \in S} \lambda_i c_i[\text{Input}](x_2) \\
& \quad \text{and } \textstyle\sum_{i \in S} \lambda_i c_i[\text{Input}](x_1) > \sum_{i \in S} \lambda_i c_i[\text{Input}](x_2)\ ] \\
\text{or } & \ldots \\
\text{or } & [\ (\forall j : j<n)\ (\textstyle\sum_{i \in S} \lambda_i c_i[\text{Input}](x_1) = \sum_{i \in S} \lambda_i c_i[\text{Input}](x_2)\ ) \\
& \quad \text{and } \textstyle\sum_{i \in S} \lambda_i c_i[\text{Input}](x_1) > \sum_{i \in S} \lambda_i c_i[\text{Input}](x_2)\ ],
\end{aligned}
$$

where P is the strong ordering induced by R and, for each i, $\lambda_i \geq 0$. So, the constraint aggregation function 'looks' at the highest ranked partitioning set and, by summing constraint violations for that set, determines which candidates must be suboptimal. The same procedure applies to the next highest ranked partitioning set, and then to the next, and so on.

For Itô & Mester, the following ranking is sufficient.

c_1 = *[ŋ	» {c_2 = IdentSS, c_3 = *g}	»	c_4 = IdentLexS
$S_1 = \{1\}$	$S_2 = \{2, 3\}$		$S_3 = \{4\}$
$\lambda_1 = 1$	$\lambda_2 = \lambda_3 = 1$		$\lambda_4 = 1$

In fact, the numbers themselves are irrelevant. What matters is the equality $\lambda_2 = \lambda_3$, which guarantees "that violations of the two constraints IdentSS and **g* count as equivalent" (page 433).

Weak utilitarian rules that mimick strict dominance hierarchies (with some constraints evaluated simultaneously) also present a way to 'turn off' constraints. In general, it is assumed that, in strict dominance hierarchies, very low ranked constraints cannot affect the outcome of an aggregation procedure, because the optimal candidate has essentially already been determined by the stage at which they are 'consulted'. The reasoning is that, if only one candidate is in contention on a given constraint, then it is optimal with respect to that constraint, no matter how many times it violates it. This reasoning is no longer valid if, for some inputs, there is never a stage where only one candidate is in contention. So, constraints that are provided by Universal Grammar but that are inactive in a particular language may need to be 'turned off' to permit surface indeterminacy.

As a concrete example, consider the spread of the feature [+nasal] from consonants to (adjacent) vowels, which is not uncommon cross-linguistically. However, onset nasals do not spread nasality onto adjacent vowels in Tokyo Japanese (Shogo Suzuki, personal communication). The problem for Itô & Mester is that the 'nasal spreading' constraint would be violated twice by *niwa ŋeta* but only once by *niwa geta*. This would make *niwa geta* optimal, destroying surface indeterminacy.

The problem can be fixed, at least with respect to constraints entirely inactive in Japanese, given a slight modification of definition 4. Suppose that c_{23} is the lowest ranked constraint active in Japanese, and that the inactive constraints are $c_{24}, c_{25}, \ldots, c_{30}$. Then we partition {1, 2, ..., k} so that the lowest ranked partitioning set is {23, 24, 25, ..., 30} and define the weak utilitarian rule with weights $\lambda_{23} = 1$ and $\lambda_{24} = \lambda_{25} = \ldots = \lambda_{30} = 0$. By setting $\lambda_{24}, \lambda_{25}, \ldots, \lambda_{30}$ to zero, violations of constraints $c_{24}, c_{25}, \ldots, c_{30}$ are made irrelevant to a candidate's global harmony level. However, we require a modification of definition 4, because, since $\lambda_{24}, \lambda_{25}, \ldots, \lambda_{30}$ are always 0, the condition that $c_{24}, c_{25}, \ldots, c_{30}$ should sometimes act as tie-breakers is violated. The modification required, then, is the removal of the phrase 'not all zero' in definition 4. Note that this proposal violates (SP).

Flemming's (1997) proposal is similar to Itô & Mester's (1997). He writes that "[t]he resolution of constraint conflict must employ numerically weighted constraints rather than the strict constraint domination assumed in Optimality Theory" (p. 1). Clearly, he is suggesting that measurability and cross-constraint comparability assumptions be changed so as to permit weak utilitarian rules. Note, however, that Flemming does not use weak utilitarian rules to permit surface indeterminacy but to select uniquely optimal outputs and, therefore, does not need to 'turn off' constraints as Itô & Mester do. Another proposal involving weak utilitarian rules is found in Burzio (1999) who claims that "the regularities of language can be accounted for by constraints that apply simultaneously, in parallel" (p. 15).

Further modifications to Prince & Smolensky's (1993) proposal are found in Crowhurst & Hewitt (1999), Kirchner (1996) and Smolensy (1995). Our remarks on these proposals are tentative as we are not sure of having understood their finer aspects. Modulo that caveat, these proposals diverge from Prince & Smolensky's version of Optimality Theory with regard to their measurability and cross-constraint comparability assumptions.

All these proposals involve 'constraint conjunction' but use that term in quite different ways. In Crowhurst & Hewitt (1999), a candidate *passes* the conjunction of constraints c_1 and c_2 if and only if it *passes* c_1 and it *passes* c_2. In Kirchner (1996), a candidate *violates* the conjunction of constraints c_1 and c_2 if and only if it *violates* c_1 and it *violates* c_2. In both cases, constraint conjunction is an idiosyncrasy of an individual language: two constraints' being conjoined in Japanese implies nothing concerning their conjunction in English. So, we can assume that all selection processes in all languages proceed identically up to the point where constraint aggregation begins and that the aggregation function is responsible for conjoining the constraints. That is, whether constraints are conjoined or not does not affect how the number of violations of a single constraint by a given candidate is calculated. We can therefore ask what measurability and cross-constraint comparability assumptions must be made in order to permit a constraint aggregation function to conjoin constraints (in either of the ways described). To answer this question, it is important to note that, for both Crowhurst & Hewitt and Kirchner, what matters is not (primarily) which of two candidates violates two conjoined constraints the fewer number of times, but whether the candidates violate the constraints at all. To take account of this significance of *zero*, the constraint aggregation function *cannot* be invariant under transformations of the constraint evaluation functions that shift the origin (as can occur under positive monotonic transformations). Rather, we must assume (ONC+0): for any two profiles $\{c_i[\text{Input}]\}$ and $\{c'_i[\text{Input}]\}$, in the domain of F, $F(\{c_i[\text{Input}]\}) = F(\{c'_i[\text{Input}]\})$ if $c'_i[\text{Input}] = \phi_i(c_i[\text{Input}])$ for some k-tuple $\{\phi_i\}$ of positive monotonic and sign-preserving transformations[20] (List, in press).

Smolensky (1995) proposes 'self-conjunction' of constraints: if candidate x violates constraint c_i n times, then it violates c_i-self-conjoined-m-times once if n is greater than or equal to m, and no times otherwise. This also represents a deviation from (ONC), the standard measurability and cross-constraint comparability assumptions underlying Optimality Theory. Depending on whether the concept of 'self-conjunction' is incorporated into the formalism exogenously or endogenously – that is, depending on whether the 'work' of self-conjunction is assumed to be done by the constraints or by the constraint aggregation functions – special significance must be attached at least to the number 0, if not to a constraint-specific number m, or even to the units of evaluation. Again, the constraint aggregation function *cannot* be invariant under transformations of the constraint evaluation functions that do *not* leave 0 (or m, if necessary) invariant. (ONC+0) would thus seem to be the minimal requirement for Smolensky's proposal to be viable.

The proposals reviewed in this section are all modifications of Optimality Theory as it is presented by Prince & Smolensky (1993). We have argued that, in each case, the locus of modification is the assumed nature of measurability and cross-constraint comparability. Moreover, we proposed that a problem facing one of the accounts (i.e., the problem of optionality or 'surface indeterminacy') can be solved by a modification of measurability and comparability assumptions. This underlines both the utility and the centrality of the issues that our formalism addresses.

4.3. The Typology of Constraints

Two types of constraints seem indispensable to constraint-based phonology[21]. The very notion of a well-formedness constraint suggests restrictions on output structures. So, we require output constraints. Furthermore, if the phonology is not merely to produce the least marked syllable of the language (/stop/ emerging as *ta-ta-ta*), then input-output faithfulness constraints are necessary to minimise divergence between input and output. So, constraints can refer to inputs and to outputs. This opens the way for other combinations: input constraints, output-output constraints.

Input constraints are to be rejected because they do no work. We are concerned with the problem of aggregation. That is, how the optimal candidate is selected from X. Input constraints are irrelevant. Besides this, they are conceptually incompatible with Optimality Theory. Recall that in the seventies and eighties, constraints had entered all levels of the grammar, including the input level (see (1)). Given the historical context, the claim that all constraints concern the well-formedness of outputs entails that there are no input constraints. Of course, it may emerge later that such constraints are required. This would have consequences for Optimality Theory and its conceptual relatives. However, in the absence of concrete arguments, this possibility should be discounted.[22]

[20] A transformation $\phi : \mathbf{R} \longrightarrow \mathbf{R}$ is sign-preserving if, for all $t \in \mathbf{R}$, $t < 0$ implies $\phi(t) < 0$, $t = 0$ implies $\phi(t) = 0$, and $t > 0$ implies $\phi(t) > 0$.

[21] C.p. Burzio (1999), who argues for the rejection of underlying representations and, *a fortiori*, of input-output constraints. However, in the absence of a specific model for his theory, it is hard to assess this proposal.

Output-output constraints are more interesting. Conceptually, they could come in two varieties. They could induce faithfulness of the output of a given selection process to the *optimal* output *of another selection process*. This type of output-output constraint is used by Kenstowicz (1999) to achieve paradigm uniformity. Alternatively, such constraints could induce faithfulness of the actual output of a given selection process to some *suboptimal* output of the Generator *within the same selection process*. This type of output-output constraint is utilised by McCarthy (1997). Call these two types of output-output constraints *OOO* and *OOS* constraints, respectively ('output-output: optimal', 'output-output: suboptimal'). (By 'selection process', we intend the process of selecting the optimal output for a given input. OOO constraints, therefore, involve two different inputs; OOS, one.)

So far, we have been assuming an *exogenous incorporation* of the constraints. That is, the constraints are treated, formally, in an identical fashion and do their work 'semantically'. All constraints are included in the constraint set, C, and to each constraint there corresponds an evaluation function, as defined above. The differences between constraints arise from differing *interpretations* of '$c_i[\text{Input}](x) > c_i[\text{Input}](y)$':

(i) "candidate x is more well-formed than candidate y", for c_i, an output constraint.

(ii) "candidate x is more faithful to Input than candidate y is" for c_i, an input-output constraint.

(iii) "candidate x is more faithful to the optimal output of the other relevant selection process than candidate y is", for c_i, an OOO constraint.

(iv) "candidate x is more faithful to candidate z than candidate y is", for c_i, an OOS constraint.

In all cases, the constraint is treated identically for the purposes of measurability, comparability, and aggregation.

In the case of OOS constraints, an alternative incorporation is possible. An *endogenous incorporation* would have the effect of introducing OOS constraints without actually introducing any. Instead of having constraints in the constraint set which, in order to assign a score to a given candidate, require

[22] Archangeli & Suzuki (1997) propose input constraints. However, it is not clear that these are constraints are entirely warranted by their data, nor is it clear that some of the same work could not be done by an algorithm for determining underlying representations. Roca (1997) observes that Archangeli & Suzuki "stretch [Optimality Theory] in ways which, if nothing else, are not obviously congenial to its spirit" (p. 27). Constraints on input are the topic of Harbour (forthcoming).

information about other candidates, we deploy an aggregation function which, when assessing the relative global harmony of two candidates, considers candidates other than the two under comparison. That is, the work is shifted from the evaluation function, which determines a candidate's score on a given constraint, to the aggregation function. The constraint set itself contains only output and input-output constraints (and OOO constraints, should any be introduced).

Exogenous incorporations are compatible with (I), endogenous not. Endogenous incorporation explicitly makes use of constraint aggregation functions that violate (I). But violation of (I) permits a logical interdependence between candidates which is not easily tractable, thus making the job of the linguist very difficult. Also, violation of (I) is strongly at odds with the evaluationist spirit of Optimality Theory. And, finally, endogenous incorporation of OOS constraints would break the *formal* symmetry between the different types of constraints. For all these reasons, we advocate the exogenous formalisation of all constraints, including the OOS type.

The theoretical considerations above force output constraints and input-output constraints on us, but leave open the question of output-output constraints, of either type. To begin with, we should attempt to use only the first two. Several proposals involving other constraint types are intended as solutions to the problem of opacity, in particular McCarthy (1997). We therefore turn now to opacity. The discussion of the constraint typology with regard to opacity is resumed in Section 4.4.1.

4.4. Opacity

The theory resulting from (U), (SP), (I), output and input-output constraints, and (ONC) or (CNC) is empirically inadequate, at least in its current versions. The most well-known difficulty is the problem of opacity. We can state the problem as follows. Let x and y be two segments and let E be an environment. Under normal, non-opaque conditions, an underlying x may emerge as surface y if y is less marked in E than x is. So, if E is 't__#' and x is the English present tense morpheme /z/, then x will surface unfaithfully as *s* in the context E: we say *eat*[*s*] not *eat*[*z*]. Opacity arises in one of two cases: (a) x surfaces unfaithfully as y even though E is not present; (b) a third segment z surfaces unfaithfully in E as x, not as the less marked y. A pseudo-example of (b) would be a morpheme /v/ such that [eat-v] surfaces as *eat*[*z*] rather than *eat*[*s*], even though [eat-z] surfaces as *eat*[*s*]. Real examples include synchronic chain shifts (see, for example, Kirchner, 1996).

As an example, we can consider the (in)famous opaque epenthesis in Tiberian Hebrew, (McCarthy, 1997). In derivational terms, the segholate /dašʔ/, 'grass', is targeted for epenthensis and the resulting /dɛšɛʔ/ undergoes deletion of the glottal stop, yielding *dɛšɛ*. This process is opaque, by clause (a) above: E, the environment for epenthesis, is 'C__C', x is Ø and y is *ɛ*. In *dɛšɛ*, Ø surfaces unfaithfully as *ɛ* even though E is not present.

The normal cases of segholate epenthesis (e.g., input /pasl/ surfacing as *pɛsɛl*) can be treated by Optimality Theory using only output and input-output constraints under a strict dominance hierarchy. McCarthy (1997), for instance, uses five constraints: *COMPLEX, a ban on complex codas; ANCHOR$_{IO}$, the requirement that, if the final segment of the input is present in the output, then it be final in the output; CODA-CONDITION, a ban on codic glottal stops; MAX-C$_{IO}$, a ban on deletion on input consonants; and DEP-V$_{IO}$, a ban on vowel insertion. However, these constraints fail to select the correct output when opaque epenthesis occurs. The following is a reduced version of McCarthy's tableau (page 22).

(5) TIBERIAN HEBREW TABLEAU: 1

/dɛšʔ/	*COMPLEX	ANCHOR$_{IO}$	CODA-COND	MAX-C$_{IO}$	DEP-V$_{IO}$
☞ dɛšɛ				*	*!
☜ dɛš				*	

Clearly, *dɛš* will be selected as optimal in this tableau. Moreover, *dɛšɛ* does not outperform *dɛš* on any constraint. However, (5) contains a very limited number of constraints. In reality, Optimality Theory posits many more constraints, including a large number which, in any given language, are very low ranked. How one might solve opacity depends on the behaviour of these lower ranked constraints. Specifically, it depends on whether *dɛš* is (weakly) Pareto superior to *dɛšɛ*, that is, whether, for every member of the constraint set, *dɛš* is at least as good as *dɛšɛ*, and, for at least one constraint, strictly better than it. We examine two cases: when Pareto superioty obtains (4.4.1) and when it does not (4.4.2). In cases of opacity, we refer to the candidate that ought to win (but does not) as the *opaque candidate*, and to the candidate that does win (but should not) as the *transparent candidate*.

4.4.1. Pareto Opacity

By *Pareto Opacity*, we intend cases in which the transparent candidate is (weakly) Pareto superior to the opaque candidate. Pareto opacity is an insoluble problem for constraint-based systems observing (SP). This follows because any constraint aggregation function satisfying (SP) will rank output x above output y if x is (weakly) Pareto superior to y. But Pareto opacity informs us that the opaque candidate is always (weakly) Pareto inferior to some other candidate, the transparent candidate, and so will never be maximally harmonic.

Unless we are prepared to relax (SP), we are forced to introduce new constraints which rank the opaque candidate above the transparent one, thereby breaking the transparent candidate's (weak) Pareto superiority. From the discussion of constraints in Section 4 above, the relevant constraints seem to be output-output constraints. Recall that these come in two types: OOO, inducing faithfulness to the output of another selection process; and OOS, inducing faithfulness to a suboptimal output of the generator within the same selection process. Which might solve opacity?

For faithfulness to the optimal output of another selection process to work, we require the following correspondence: for every opaque, *o*, and every transparent candidate, *t*, corresponding to a given input, there exists a second input with surface form, *s*, such that *o* is more faithful to *s* than *t* is. Whether there is such a correspondence is an empirical matter, as is whether one could identify the relevant input. We are aware of no general treatment of opacity which takes this as its starting assumption, however, perhaps reflecting a difficulty with such an approach. The question is left open.

Famously, there is a proposed solution to opacity using the other type of output-output constraint: McCarthy's (1997) Sympathy Theory.[23] McCarthy proposes introduction of an OOS constraint on which the opaque candidate outperforms the transparent.

How could the OOS constraint of Sympathy Theory be formalised? Recall that the effect of OOS constraints can be achieved in two ways, in an exogenous way and in an endogenous way. Under an exogenous formalisation, the effect of OOS constraints is achieved 'semantically', whereas, under an endogenous formalisation, the effect of the constraints is built into the constraint aggregation function by allowing this function to violate (I) (in a suitably defined systematic way). We should therefore note that an endogenous formalisation of OOS constraints could never solve the problem of Pareto opacity, unless, of course, we are willing to abandon (SP). For, if no new constraints are added to the constraint set and the transparent winner (weakly) outperforms the opaque one on every constraint included in that set, the opaque winner will never be ranked above the transparent one by a constraint aggregation function satisfying (SP), no matter how much logical interdependence between the rankings of (more than two) candidates we admit.

In order to tackle Pareto opacity, sympathy theory must therefore formalise OOS constraints exogenously. A new constraint on which the opaque winner scores better than the transparent one is introduced into the dominance hierarchy higher than the constraint which would otherwise select the transparent candidate over the opaque. This constraint breaks the Pareto superiority of the transparent candidate.

A suitable aggregation procedure can then rank the opaque candidate above the transparent one without violating (SP). Furthermore, as Gen is assumed to include all linguistically coherent representations, Sympathy Theory does not suffer from the difficulty noted above, that there may not exist a structure such that the opaque candidate is more faithful to it than is the transparent candidate.

[23] We do not claim any of the cases considered by McCarthy to be cases of Pareto opacity. Rather, we aim to show how his proposal can be implemented against cases of Pareto opacity, should any arise.

Sympathy Theory is, however, in friction with the spirit of Optimality Theory, and of constraint-based phonology generally. OOS constraints introduce, in effect, an intermediary level of representation into phonology. The target of OOO-faithfulness is the output of a selection process. However, there is no process which selects as optimal the target of OOS-faithfulness (unless the target of sympathy is the optimal output itself). Yet, this is the only significant difference. Formally, OOS and OOO constraints are very similar, both inducing faithfulness to something other than the input. This leads one to say that the target of faithfulness in OOS constraints is just as much a part of the language as the target of faithfulness in OOO constraints. Moreover, the process which selects the sympathetic candidate, a strict dominance hierarchy on a proper subset of X, is very similar to that which selects the target of OOO-faithfulness. If one wished to introduce intermediary levels of representation to Optimality Theory, one would be hard pressed to find a better way to it than McCarthy proposes. However, intermediary levels are contrary to the spirit of Optimality Theory, which aims to make grammaticality purely a matter of conformity of the output to structural well-formedness constraints.

Thus, if there is Pareto opacity in natural language, it presents a real problem for Optimality Theory. Unless we are prepared to abandon (SP), it cannot be accommodated with the most sparse assumptions concerning the nature of constraints. If OOO constraints are adequate, then there must exist a correspondence between opaque and transparent candidates of one selection process and optimal candidates of another, as well as a means of identifying the relevant selection process. On the other hand, OOS constraints seem substantially to be in friction with the spirit of constraint-based theories of phonology in that they tacitly introduce intermediary levels of representation, the need for which constraints are meant to obviate.

4.4.2. Non Pareto Opacity

Suppose that there is at least one output or input-output constraint on which the opaque candidate outperforms the transparent. Then Pareto opacity does not obtain. Given the enormous number of conflicting constraints, this opacity seems likely to be the more common. In what follows, we discuss whether the problem non-Pareto opacity can be solved by changing measurability and cross-constraint comparability assumptions. We should note that Kirchner's (1996) proposal, discussed in Section 4.2.1, is an example of a change in measurability and cross-constraint comparability assumptions that is intended to account for an instance of opacity, namely, synchronic chain shifts.

In the case of opaque epenthesis in Tiberian Hebrew segholates, the transparent candidate is not Pareto superior to the opaque candidate, owing to NOCODA, a well-motivated constraint banning codas: the transparent candidate, though not the opaque, violates it. The realisation that *dɛš* is not Pareto superior to *dɛšɛ* presents us with a very simple solution to the problem of epenthesis in

/daš'/.[24] We can take the five constraints, ranked as (6), and simply insert the constraint that breaks Pareto superiority between the constraints that are second lowest and lowest ranked.

(6) *COMPLEX, ANCHOR$_{IO}$, CODA-CONDITION » MAX-C$_{IO}$ » DEP-V$_{IO}$

(7) *COMPLEX, ANCHOR$_{IO}$, CODA-CONDITION » MAX-C$_{IO}$ » NOCODA » DEP-V$_{IO}$[25]

This yields a set of constraints, which, under a strict dominance hierarchy, selects *dɛšɛ* as transparently optimal. This is shown explicitly in (8). The tableau presents three inputs: /daš'/; /pasl/, a normal segholate; and /lo'/, a normal glottal-stop-final word. (We choose /pasl/, 'idol', over /malk/, 'king', used by *inter alia* McCarthy (1997), Idsardi (1998), to avoid the complication of spirantisation.) All the correct selections are made using the constraints in the dominance hierarchy shown.

(8) TIBERIAN HEBREW TABLEAUX: 2

/daš'/	*COMPLEX	ANCHOR$_{IO}$	CODA-COND	MAX-C	NOCODA	DEP-V
☞ dɛšɛ				*		*
dɛš				*	*	
dɛšɛ'			*		*	*
dɛš'ɛ		*			*	*
dɛš'	*		*		*	
/pasl/						
☞ pɛsɛl					*	*
pɛsɛ				*		*
pɛs				*	*	
pɛsl	*				*	
pɛslɛ		*			*	*
/lo'/						
☞ lo				*		
lo'			*		*	
lo'ɛ		*				*

[24] We thank John Frampton for suggesting that we look for a constraint that breaks the Pareto superiority of *dɛš*, Jason Eisner for not letting us miss the significance of NOCODA, and Benjamin Bruening for reminding us of 'sin'.

[25] Bill Idsardi (p.c.) points out that this ranking would cause epenthesis in word-internal consonants clusters, a problem that can be fixed if NoCoda is interpreted as referring to word final-codas only in the current case.

This solution is quite obvious and natural and underlines the usefulness of the distinction we have drawn between Pareto and non-Pareto opacity (and, consequently of the formalism in Section 3). However, the phonology of Tiberian segholates presents a variety of opacities: epenthesis, stress, tonic lengthening, spirantisation, and vowel harmony (Idsardi, 1998). The introduction of NOCODA solves only a part of one of these problems.

To see that the strict dominance hierarchy in (7) does not select the correct results for all cases of epenthesis consider another segholate, /ħe:ṫ'/, 'sin'.[26] This is another glottal-stop-final segholate. However, it does not undergo epenthesis: its singular is *ħe:ṫ* not *ħe:ṫɛ*. The wrong result is predicted by (7), as shown below.

(9) TIBERIAN HEBREW TABLEAU: 3

/daš'/	*COMPLEX	ANCHOR$_{IO}$	CODA-COND	MAX-C	NOCODA	DEP-V
☞ dɛšɛ				*		*
dɛš				*	*	
/ħe:ṫ'/						
☞ (wrong) ħe:ṫɛ				*		*
☞ ħe:ṫ				*	*	

The problem is that, so far as the constraints in the tableau are concerned, /ħe:ṫ'/ and /daš'/ are structurally identical: /CVC'/. So, either epenthesis will be optimal in both cases or it will be suboptimal in both cases. Either way, the wrong surface form will be selected for one of the inputs.

This new problem cannot be solved as the problem of *dɛšɛ* versus *dɛš* was, by the introduction of a new constraint such as NOCODA, as becomes apparent when one tries to imagine what the new constraint would need to do. According to (9), *dɛšɛ* P *dɛš* if and only if *ħe:ṫɛ* P *ħe:ṫ*. The new constraint must ensure that *dɛšɛ* P *dɛš* and *ħe:ṫ* P *ħe:ṫɛ*. To do this, the constraint must be violated more by *ħe:ṫɛ* than by *ħe:ṫ*.

The constraint could not be a feature markedness constraint, because all the features of *ħe:ṫɛ* are also features of *ħe:ṫ* and *dɛšɛ*. If the new constraint penalises a feature of common to *ħe:ṫɛ* and *ħe:ṫ*, then the extra violations will not affect the relative global harmony of the two outputs. And if the new constraint penalises a feature common to *ħe:ṫɛ* and *dɛšɛ*, then it will not alter the current situation, that *dɛšɛ* P *dɛš* if and only if *ħe:ṫɛ* P *ħe:ṫ*. Similarly, the new constraint could not be a faithfulness constraint. So, the constraint would have to be a structural markedness constraint. That is, [$_{PrWd}$...*ṫɛ*] must be more marked than [$_{PrWd}$...*ṫ*]. However, there is simply no motivation for such a constraint. An *ad hoc* prohibition on, say, the string 'C̣ɛ' ('pharyngeal+ɛ') could be invoked, which only *ħe:ṫɛ* would violate, and this would produce the correct output if the

[26] [ṫ] denotes a pharyngealised [t].

new constraint were ranked between MAX-C_{IO} and NOCODA. But this seems more formal trick than insight.

To clarify, the problem is that there are two different vectors of scores, v_1 and v_2, such that, with respect to $Input_1$, v_1 corresponds to the optimal candidate and v_2 to the suboptimal, whereas with respect to $Input_2$, v_1 corresponds to the suboptimal candidate and v_2 to the optimal. However, any constraint aggregation function satisfying (INV) will rank the two pairs of candidates with the same pair of score-vectors, v_1 and v_2, identically. Moreover, the four candidates appear similar enough in the relevant respects for there to be no constraint that differentiates between them in such a way as to permit a new strict dominance hierarchy under which only the correct candidates are selected.

Two observations are needed at this point, one concerning what this problem means for Optimality Theory and constraint-based phonology in general and the other concerning what it means for the analysis of Tiberian Hebrew in particular. In brief, we will argue that situations such as the current one are very real problems for any constraint-based phonology; and that is so, even though the tableau reveals that the analysis of Hebrew using the constraints in (6) is fundamentally misconceived.

There is a very simple reason to suspect that the analysis of Tiberian Hebrew based on the constraints in (6) is fundamentally misguided: it represents *dɛš* as a reasonable contender for the singular *absolute* form of /daš'/. Hebrew nouns have two forms, *absolute* or *construct*: the absolute of, say, *pɛsɛl* means simply '(an) idol', whereas the construct means '(an) idol of'. Construct forms are dependent forms in that they must be followed by another nominal and in that they often have a special, 'dependent' phonological forms. Now, all Hebrew words of the form *CɛC* are phonological dependents, such as construct forms or closed class items.[27] In Masoretic texts, dependents are hyphenated with what supports them. Some are shown in (11), with '–' for the Masoretic hyphen.

(11)	*pɛn–*	'lest'	*šɛl–*	'of'
	bɛn–	'son-CONSTR.'	*šɛm–*	'name-CONSTR.'
	tɛn–	'give-IMPERATIVE-M.SG'		

Several of these words have phonologically independent forms, but these forms crucially have a different vowel: long *e:* not short ɛ. These is shown in (12).

(12)	*be:n*	'son-ABS.'	*še:m*	'name-CONSTR./ABS.'
	te:n	'give-IMPERATIVE-M.SG'		

These facts reveal that the constraints in (7) are substantially wrong: they represent *dɛšɛ* as being in crucial competition with *dɛš* for a role (the singular absolute of /daš'/) that *dɛš*, as an improper word of Tiberian Hebrew, simply cannot fill.

A more reasonable competitor for *dɛšɛ* is *de:š/då:š*, or perhaps even *daš*, given the restrictions on the structure of monosyllabic words in Tiberian Hebrew. These candidates would break the problematic symmetry, *dɛšɛ~ħe:tɛ* and *dɛš~ħe:t*, and could, therefore, permit a solution. In particular, the reader is referred to Bruening (1999) whose proposal involving morphological templates and metrical feet suggests we are correct to consider *de:š/då:š* a more appropriate competitor for *dɛšɛ* than *dɛš*, and to Harbour (forthcoming) for an analysis of monosyllabic nouns and adjectives in Tiberian Hebrew.

The possibility of solving the Tiberian Hebrew instance of the problem formulated three paragraphs higher does not answer the more general issue of whether Optimality Theory, or other constraint-based phonologies, can cope with such cases. In the absence of a concrete example, the question may seem somewhat pointless. So, we sharpen it by showing that the problem remains even if the Hebrew instance of it does not. We then argue that such cases are a problem for any constraint-based grammar assuming (U), (I), (SP) and any of the options in Table 1.

The new example we consider comes from Bedouin Arabic. This case of opacity is considered by McCarthy (1997), whose exposition we follow. The problem concerns vowel raising in open syllables. When it occurs finally in open syllables in Bedouin Arabic, *a* raises to *i*. Thus, /katab/ surfaces as *ki.tab*. Furthermore, underlying glides vocalise under certain circumstances. Interestingly, if an underlying *a* occurs in an open syllable created by glide vocalisation, then it does not undergo raising. Thus, /badw/ surfaces as *ba.du*, rather than *bi.du*.

These facts are hard to account for in Optimality Theory. The relevant constraints are as follows. Since *a* becomes *i*, the markedness constraint concerning *a* in open syllables must outrank the relevant faithfulness constraint. I.e., $*a]_\sigma$ » IDENT(high)$_{IO}$. Glide vocalisation shows the relevant markedness constraint to outrank the relevant faithfulness constraint: *COMPLEX » DEP-μ_{IO}. Lastly, *COMPLEX must outrank $*a]_\sigma$ as *ba.du* violates $*a]_\sigma$ whilst satisfying *COMPLEX. So, we have *COMPLEX » $*a]_\sigma$ » IDENT(high)$_{IO}$, DEP-μ_{IO}. However, this predicts the wrong result. The tableau shows the problem case and /badu/ and the simple case /katab/. (This is not a case of Pareto opacity, since the opaque candidate outperforms the transparent candidate on the constraint IDENT(high)$_{IO}$.)

[27] The two apparent exceptions to this generalisation (*pɛh* 'mouth'; *śɛh* 'sheep') are in reality simply *Cɛ* words. The *h* is silent.

(13) BEDOUIN ARABIC TABLEAU

/badw/	*COMPLEX	$*\,a]_\sigma$	IDENT(high)$_{IO}$	DEP-μ_{IO}
☞ ba.du		*		*
bi.du			*	*
badw	*			
/katab/				
ka.tab		*		
☞ ki.tab			*	

If the tableau contains all relevant constraints, then the situation cannot be handled by a strict dominance hierarchy. However, it also cannot be handled by any change in assumptions concerning measurability and cross-constraint comparability. The difficulty, is essentially the same as the case of *dεšε~ħe:t́ε* and *dεš~ħe:t́*. With regard to output structures, *ba.du/bi.du* is a practically minimal pair with *ka.tab/ki.tab*: both differ only with respect to *i* versus *a*. But the candidate containing *i* wins in the lower part of the tableau, and the candidate with *a* in the upper. There appears to be no constraint that sorts *i* from *a* in one case and *a* from *i* in the other, nor is one easily imagined.

To see that a change in assumptions concerning measurability and cross-constraint comparability does not help, suppose that we permit weak utilitarian rules by assuming cardinal measurability and cross-constraint comparability of local harmony units (CUC). Then, we need to find non-negative coefficients λ_i, not all zero, such that the required constraint aggregation function F can be defined as follows: for all x_1, x_2 in X,

$$x_1 R x_2 \text{ if and only if } \sum_{i\in\{1,2,\ldots,k\}} \lambda_i c_i[\text{Input}](x_1) \geq \sum_{i\in\{1,2,\ldots,k\}} \lambda_i c_i[\text{Input}](x_2),$$

where R = F({c_i[Input]}). We take each violation to measure minus-one unit: that is, c_i[Input](x) equals the negative of the number of violations of candidate x on the ith constraint. In (13), therefore, every cell is filled by '0' or '−1' so far as the aggregation function is concerned. (The use of negatives may seem somewhat awkward, but we should bear in mind that, if we impose (CUC), significance is attached only to the units of the scores and not to the position of the origin.)

Let the set of constraints be C = {c_1 = *COMPLEX, c_2 = $*a]_\sigma$, c_3 = IDENT(high)$_{IO}$, c_4 = DEP-μ_{IO}}. We must assign values to λ_i, $i\in\{1, 2, 3, 4\}$, in such a way that

$$\sum_{i\in\{1,2,3,4\}} \lambda_i c_i[\text{badw}](\text{ba.du}) > \sum_{i\in\{1,2,3,4\}} \lambda_i c_i[\text{badw}](\text{bi.du}).$$

I.e., $\lambda_2(-1) + \lambda_4(-1) > \lambda_3(-1) + \lambda_4(-1)$. This immediately implies that $\lambda_2(-1) > \lambda_3(-1)$ (subtracting $\lambda_4(-1)$ on both sides of the inequality). So,

$$\sum_{i\in\{1,2,3,4\}} \lambda_i c_i[\text{katab}](\text{ka.tab}) > \sum_{i\in\{1,2,3,4\}} \lambda_i c_i[\text{katab}](\text{ki.tab}),$$

contrary to the desired result. Therefore, any weak utilitarian rule that ranks *ba.du*, the optimal candidate, above *bi.du*, the suboptimal one, will also rank *ka.tab*, the suboptimal candidate, above *ki.tab*, the optimal one. Clearly, this situation is not improved by altering (CUC) to (CFC).

It may be that we are not considering the correct constraints above, despite the motivation given. Or some fifth constraint may assign sufficiently different scores to *ba.du*, *bi.du*, *ka.tab*, *ki.tab* to make a weak utilitarian rule or an alternative viable. In the absence of such proposals, we must conclude that some cases of non-Pareto opacity are problematic for constraint-based phonology under a variety of assumptions concerning measurability and cross-constraint comparability.

5. Conclusion

We began this paper with a simple observation. To do constraint-based phonology whilst using constraints that make incompatible demands, we require a means of resolving constraint conflicts. There are many ways of resolving such conflicts, and which resolution method one chooses is related to one's assumptions about measurement of violations, cross-constraint comparability of scores, and the way in which the aggregation function works. Optimality Theory uses a strict dominance hierarchy, the only option if we impose cardinal/ordinal measurability, no comparability, universal domain, the strong Pareto principle, and independence of irrelevant alternatives. Recognising there to be alternatives to strict dominance hierarchies, one is led to ask which shortcomings of Optimality Theory are simply shortcomings of strict dominance hierarchies and which reflect a deeper inadequacy, perhaps inevitable in constraint-based theories of any description.

Our aim in this paper has been to address these issues rigorously, employing methods from social choice theory, which deals with problems formally similar to those in phonology (and in aspects of syntax, particularly the economy of derivation; Harbour & List, forthcoming). Invoking these formal methods, we have examined different proposals on how Optimality Theory could be modified, particularly in view of the problem of *opacity*. The modifications we have discussed concern Optimality Theory's assumptions on measurability, comparability and aggregation, or Optimality Theory's typology of constraints. We underlined the centrality of these issues to constraint-based phonology by showing that several variants of Optimality Theory are in fact what would result from changing Prince & Smolensky's assumptions concerning measurability and cross-constraint comparability. And we have argued that certain cases of opacity are simply insoluble within a constraint-based system which obeys (SP) and which uses only output and input-output constraints together with ordinal or cardinal measurement and no cross-constraint comparability. If these cases are not misanalysed then a solution to opacity must come either from a richer constraint typology or from a change in measurability and comparability assumptions.

A richer constraint typology would include OOO or OOS constraints. OOS constraints, used in Sympathy Theory (McCarthy, 1997), suffer from the theoretical problem (independent of questions concerning their empirical adequacy) that they introduce an intermediary level of representation which constraint-based theories are designed to obviate. Whether OOO constraints offer a solution is left an open question.

Different measurability and cross-constraint comparability assumptions may solve some instances of opacity, as appears to be the case with, for instance, Kirchner's (1996) proposal. However, some cases of opacity are impervious to changes in assumptions. One example is Bedouin Arabic *a*-raising, where the set of scores for the optimal candidate of one selection process can be too similar to the set of scores for a suboptimal candidate of another selection process, thereby making it impossible for, e.g., a weak utilitarian rule to determine the right outcome in both cases.

The use of more powerful constraint aggregation functions raises the issue of what cross-constraint comparability signifies in phonology. It is not clear what it means to ask whether a violation of constraint c_1 by candidate x is worse than a violation of constraint c_2 by candidate y in some sense above and beyond a question of constraint ranking. Until clarified, the use of utilitarian rules, or more sophisticated methods, raises questions that are hard to answer. This question is, therefore, raised also by the authors discussed in Section 4.2.1.

We have not exhausted all possible solutions here. In particular, more constraints may be relevant to the Bedouin Arabic case than we have considered. However, if correct, the preceding cases greatly restrict the possible structure of a constraint-based theory of phonology. Indeed, we have been able to argue the most natural version of constraint-based phonology, one using only output and input-output constraints together with minimal assumptions about measurability, comparability, and aggregation, to be non-viable. Furthermore, we have argued that the most appealing alterations to that theory are problematic in some cases.

We have seen that, in order for pure constraint-based phonology to work, we need to accept evaluationism, the claim that the evaluation scores of two candidates in X must contain sufficient information to determine their position relative to each other in the global harmony ordering. Is this

information sufficient? Let us briefly return to social choice theory. Issues such as morality, context, culture, and the 'history' of each policy option (its 'path' or 'derivation') may be relevant to social orderings and, hence, may have to be represented in the choice procedure over and above individuals' welfare scores. This is inconsistent with welfarism. Our failure to construct a constraint aggregation function which generates global harmony orderings that deal adequately with transparent and opaque candidates raises a corresponding question in phonology. Is it conceivable that the pair of score-vectors a constraint set, C, assigns to the pair of candidates x_1, y_1 is identical to the pair of score-vectors C assigns to the pair of candidates x_2, y_2 (possibly for a different input), and yet x_1 is more globally harmonic than y_1 whilst x_2 is *not* more globally harmonic that y_2 given the relevant input(s)? Presumably the answer to this question depends strongly on what constraints (and how many of them) we include in the set C. If the set of (relevant) constraints is small, this is clearly conceivable, as our example of the two pairs of candidates (*dɛšɛ*, *dɛš*) and (*ħe:ṭɛ*, *ħe:ṭ*) in Section 4.4.2 shows. If a very large number of constraints is included, on the other hand, then maybe it is unlikely that different candidates will be assigned *perfectly identical* vectors of scores. If identical vectors of scores never occur, then it *might* be logically possible to define a descriptively adequate constraint aggregation function *extensionally*, by matching up tableaux of scores with the desired rankings in an *ad hoc* manner. There would, however, be no guarantee that a reasonably parsimonious and systematic *intensionsal* definition of this function could be given, let alone that the function would satisfy any of the 'minimal conditions' from social choice theory discussed above. In particular, there would be no guarantee that such a constraint aggregation function would represent a learnable method of resolving constraint conflict.

Might there be more to grammaticality than structural well-formedness? If the answer is 'yes', then, whatever role constraints play in phonology, it is not captured by Optimality Theory, nor can it be by any other pure, constraint-based theory. In seeking an answer, there will be many issues to consider. We have argued that the nature of constraint aggregation is one those issues, and a fundamental one at that.

References

Archangeli, Diana (1997). Optimality theory: An introduction to linguistics in the 1990's. In Diana Archangeli and D. Terence Langendoen (eds.), *Optimality Theory: An Overview*, Oxford: Blackwell, pp. 1-32.

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

Arrow, Kenneth (1950). *Social Choice and Individual Values*. New York: Wiley.

Arrow, Kenneth (1963). *Social Choice and Individual Values*, 2nd edition. New York: Wiley.

d'Aspremont, Claude (1985). Axioms for social welfare orderings. In L. Hurwicz, D. Schmeidler, and H. Sonnenschein (eds.), *Social Goods and Social Organisation*, Cambridge: Cambridge University Press, pp. 19-76.

Blackorby, Charles & David Donaldson (1984). Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: Admissible social evaluation functions. *International Economic Review* 25: 327-356.

Bruening, Benjamin (1999). Morphological templates and phonological opacity. Handout at Mediterranean Morphology Meeting 2, Malta.

Burzio, Luigi (1999). Missing players: Phonology and the past-tense debate. Ms., Johns Hopkins University.

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

Chomsky, Noam and Morris Halle (1991). *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

Crowhurst, Megan and Mark Hewitt (1999). Boolean operations and constraint interactions in optimality theory. Ms., Rutgers Optimality Archive No. 229-1197.

Efimov, Boris-A. and Gleb-A. Koshevoy (1994). A topological approach to social choice with infinite populations. In *Mathematical Social Sciences* 27(2): pp. 145-157.

Eisner, Jason (1997a). Efficient generation in primitive optimality theory. Handout at ACL, Madrid.

Eisner, Jason (1997b). Constraining OT: Primitive optimality theory. Handout at Ling-Lunch, MIT.

Ellison, Mark T. (1994). Phonological derivation in optimality theory. In COLING'94, vol. II, Kyoto, Japan, pp. 1007–1013.

Even-Shoshan, Avraham (1977). *קונקורדנציה חדשה לתורה נביאים וכתובים*. Kiryath Sefer, Jerusalem.

Flemming, Edward (1997). Phonetic Optimization: Compromise in Speech Production. In *Univerity of Maryland Working Papers in Linguistics*, vol. 5.

Harbour, Daniel (forthcoming). A Constraint on Underlying Representations (working title).

Harbour, Daniel & Christian List (forthcoming). Aggregation and the Economy of Derivation (working title).

Heal, Geoffrey-M. (ed.) (1997). Topological Social Choice. In special issue of *Social Choice and Welfare*, vol. 14, no. 2.

Idsardi, William (1998). Segholate Opacities. Handout at Fall Colloquium, MIT.

Itô, Junko & Armin Mester (1997). Correspondence and Compositionality: The Ga-gyō Alternation in Japanese Phonology. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*, Oxford: O.U.P.

Kenstowicz, Michael (1999). Paradigmatic Uniformity and Contrast. Handout at Phonology 2000 Conference, MIT and Harvard; forthcoming in MITWPL.

Kirchner, Robert (1996). Synchronic Chain Shifts in Optimality Theory. In *Linguistic Inquiry*, vol. 27, no. 2.

Kisseberth, Charles (1970). On the Functional Unity of Phonological Rules. In *Linguistic Inquiry*, vol. 1, pp. 291–306.

List, Christian (in press). A Note on Introducing a "Zero-line" of Welfare as an Escape Route from Arrow's Theorem. In *Social Choice and Public Policy*, special issue of *Pacific Economic Review* in honour of Amartya Sen.

McCarthy, John (1997). Sympathy & Phonological Opacity. Ms. University of Amherst

Prince, Alan & Paul Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms. Rutgers University and University of Colorado at Boulder.

Prince, Alan & Paul Smolensky (1997). Optimality: From Neural Networks to Universal Grammar. In *Science*, vol. 275, pp. 1604-1610.

Roberts, Kevin (1980a). Possibility Theorems with Interpersonally Comparable Welfare Levels. In *Review of Economic Studies*, XLVII, pp. 409-420.

Roberts, Kevin (1980b). Interpersonal Comparability and Social Choice Theory. In *Review of Economic Studies*, XLVII, pp. 421-439.

Roca, Iggy (1997). Derivations or Constraints, or Derivations and Constraints. In Iggy Roca (ed.) *Derivations and Constraints in Phonology*, Oxford: O.U.P.

Roemer, John E. (1996). *Theories of Distributive Justice*. Cambridge, Mass.: Harvard University Press.

Sen, Amartya (1982a). A Possibility Theorem on Majority Decisions. In Sen *Choice, Welfare and Measurement*, Oxford: Basil Blackwell.

Sen, Amartya (1982b). Social Choice Theory: A Re-examination. In Sen *Choice, Welfare and Measurement*, Oxford: Basil Blackwell.

Sen, Amartya (1982c). On Weights and Measures. In Sen *Choice, Welfare and Measurement*, Oxford: Basil Blackwell.

Smolensky, Paul (1995). On the Internal Structure of the Constraint Component 'Con' of UG. Handout at UCLA, April, 1995.

Tesar, Bruce (1996). *Computational Optimality Theory*. Ph.D. thesis, University of Colorado, Boulder, CO.

Tsui, Kai-yuen & John Weymark (1997). Social Welfare Orderings for Ratio-Scale Measurable Utilities. In *Economic Theory*, 10, pp. 241-256.

Daniel Harbour
Department of Linguistics & Philosophy
E39-245 MIT
77 Massachusetts Avenue
Cambridge, MA 02139
USA

harbour@mit.edu

Christian List
Nuffield College
Oxford OX1 1NF
UK

christian.list@nuffield.oxford.ac.uk

MIT Working Papers in Philosophy and Linguistics
Volume 1 (2000)

The Linguistics Philosophy Interface

Edited by
Rajesh Bhatt Patrick Hawley