Comments on Luigi Burzio’s
The rise of optimality theory (Glot 1, 6)

I write to comment on a number of issues raised in the recent paper by Luigi Burzio “The Rise of Optimality Theory” (Glot International 6).

Burzio begins his paper by recalling that in his classes of ca. 1976 Chomsky “would argue that writing a rule does not constitute a solution to a problem, but merely a statement of it.” He finds this position incompatible with Chomsky’s more recent statement that phonology is rule-based and that “the rules deriving the alternants decided on have decision (...) are straightforward and natural at each step.” In fact, there is no incompatibility between the two statements. Chomsky’s point has always been that rules differ from mere statements of fact in that, unlike the latter, rules are subject to specific independently motivated constraints as to their form and to their interaction with one another. Thus, the validity of a rule is not established unless and until its manner of interaction with other rules has also been established. Because of this interaction, a rule-based account of a given state of affairs is always more than just the sum of the separate rules, but this can only be seen if more than one rule is involved.

One of the basic properties of phonological rules is that they are ordered. Rule ordering is a central property of the framework in which rules interact: when, for example, a given string satisfies the structural description of more than one rule, implicit in this proposition is that some rule interactions will be feeding, others will be bleeding, and yet still others will be counter-feeding. And there are even more complicated effects that a given rule may have on the application of a later rule in the order, although no one has invented special names for the latter relationships among rules. It is therefore not clear why counter-feeding and counter-feeding effects have been singled out by Burzio (see, e.g. 5) to be listed among the outstanding empirical issues for OT, especially since the evidence that such rule interactions play a role in phonology is meager.

On the other hand, Burzio fails to mention the one type of rule interaction that is known to play a special role in phonology, namely the interaction ruled out by the Elsewhere Condition of Kiparsky (1973). This omission is unfortunate, since the Elsewhere Condition effects cannot be readily captured in the OT framework. As Kiparsky noted, the Elsewhere Condition was an important convention for the application of the rules already in Pāṇini’s grammar. Formally, the Elsewhere Condition is a special proviso on rule interaction that goes beyond the interaction expressed by linear order. The Elsewhere Condition states that a) if two rules resemble each other formally so that the structural description of the more restrictive rule includes (i.e., entails, but is not entailed by) the structural description of the less restrictive rule, and b) if their structural changes are either identical or incomparable (e.g., one inserts a segment S and the other deletes S), then c) the more restrictive rule must be ordered before the less restrictive, and d) the order is disjunctive, where “disjunctive” means that the less restrictive rule may not apply to a string that has the form of the output of the more restrictive rule. (The definition of “disjunctive” here is somewhat more general than the one given in Kiparsky (1973), where “disjunctivity” was limited to strings to which the more “restrictive” rule had applied. Evidence supporting the modified definition is to be found in Halle & Idsardi (1995).

The Elsewhere Condition was discussed in Prince & Smolensky (1993), and it has been widely assumed that this discussion has shown that much of the Elsewhere Condition dissolves into logic, and that what remains is unimportant or plainly incorrect. I do not believe that this is a justified inference. Prince & Smolensky’s discussion focuses on what they call Pāṇini’s Theorem on Constraint-Ranking (PTC). PTC concerns the ranking of constraints, and not the ordering of rules. Prince & Smolensky remark that although PTC has obvious affinities with the Elsewhere Condition, “[t]here is an important difference: PTC is merely a point of logic, but the Elsewhere Condition has the right of as a principled specific to UG, responsible for empirical results that could well be otherwise.” It follows from this that if the Elsewhere Condition is correct, it is an empirical result of some importance that should not be set aside on the basis of a few putative counter-examples. But that is precisely what Prince & Smolensky do, and their conclusions, as noted above, have been widely accepted. So let us look at Prince & Smolensky’s discussion.

Prince & Smolensky consider only two examples: Kiparsky’s stress examples and the choice among the plural allomorphs in English. The latter is clearly irrelevant to the Elsewhere Condition: it is a problem in morphology rather than in phonology, for it has been known since Aronoff’s (1976) discussion of Blocking effects in the morphology that these are fundamentally different from the Elsewhere effects in phonology. Kiparsky’s stress examples are indeed problems in phonology, but they have a solution that requires no reliance on the Elsewhere constraint, as Prince & Smolensky point out. Thus, neither of these examples provides support for the Elsewhere Condition.

From the irrelevance of these two examples, however, one cannot conclude that the Elsewhere Condition is invalid, for there are many examples that support the Condition. A complex and valid example of the Elsewhere Condition is provided by the interaction of the Lengthening and Shortening rules of English, as discussed in interesting detail by Myers (1987), see also Halle & Idsardi (1992).

The phenomena captured by these rules are central to the phonology of English, perhaps the most intensely studied language we have. Since the current literature rarely does them full justice, I mention a few of their most salient characteristics below.

On the one hand, in words such as divinity, natur-al and ton-ic, athlete-ic, as well as in Pāṇini the stem vowel is shortened, but there is no shortening in ton-al or atone-mot. On the other hand, there is lengthening in Caucas-ian, remedi-al, but not in remedy-ing, buri-al, or Casp-ian. These two length alternations share the important property that the alternating stem vowel is stressed in all cases. In current metrical theories, stress is assigned exclusively to heads of feet. The alternations of interest affect therefore heads of feet.

As pointed out by Myers, Shortening takes place only if the foot in question is poly-syllabic.

Particularly instructive here is the comparison of the adjectival suffixes -ic and -al.

The stress placement in forms such as original, parent-al, suicid-al shows that -al is extra-metrical; we express this formally by placing a Right parenthesis before -al. <utilize here the formalism of the metrical theory of Idsardi (1992), because it allows me to bring out the points with maximum clarity.> In addition, English is also subject to binary foot construction proceeding from Right to Left, and to a rule accenting heavy syllables, which is formally implemented by inserting a Left parenthesis to the Left of such syllables. These rules assign to original- al pal(at-ic)al, (suicid)-al the foot structure shown. By contrast, -ic always places the stress in the last syllable of the word, and this is reflected formally by representing -ic without parentheses of any kind. The binary foot construction rule will therefore assign stress always to the penultimate syllable by placing a Left Parenthesis before it; e.g., (ton)-ic, athl(e -tic).

Following Myers (1987) we state the main Shortening Rule of English (the counterpart of SPE’s Trisyllabic Laxing rule) as follows: “Shorten the head vowel of a branching foot”. Shortening therefore applies in (natur)-al and (di)v(in-i)ty as well as in (ton-ic) al and athl(e -tic). These two length alternations are therefore as pointed out by Myers — in the latter two forms the stressed syllable is the head of a non-branching foot.

Like Shortening, the Lengthening rule also affects the head of a branching foot, but here a host of additional conditions must be satisfied. As was noted above, the vowel must be [-high] and must moreover be followed by an open syllable ending in /i/, which in turn must be followed by a vowel in hiatus. Lengthening therefore applies in (remedi)-al and (co(n)i)-al, but it does not apply in (natur)-al, (athl(e -tic)).

Since the Lengthening rule must meet several conditions in addition to those also met by Shortening, the structural description of Lengthening includes the structural description of Shortening. It is by virtue of this fact that the two rules stand in the Elsewhere
relation; i.e. they constitute a disjunctive block, where the more restrictive Lengthening rule applies before the less restrictive Shortening rule second. Now, as noted above, strings having the same form as the output of the more restrictive \(<\text{Lengthening}>\) rule are prohibited from undergoing the less restrictive \(<\text{Shortening}>\). As a consequence, neither Shakespearean nor jovial are subject to Shortening, whereas stems formed with \(-\text{ing}, \text{-ION}, \text{-al}\) are.

Since the English stress rules are part of the cyclic stratum II, noncyclic suffixes such as \(-\text{ing}\) and the noun-forming \(-\text{al}\) are stress-neutral. Since the

Lengthening and Shortening rules are also in the cyclic stratum I, these rules are not triggered by stress change. In this discussion as noted, these suffixes form non-cyclic constituents. Once again, there is a lot of structure in words to which the rules are sensitive, and this structure accounts for all kinds of subtle phonological effects. It is not obvious how these effects can be captured in an OT account, and Burzio's present answer is wrong.

These effects also have special bearing on the OT Faithfulness constraints. As illustrated in the discussion above, in the rule-based account there is no global Faithfulness requirement—however that is to be expressed for representations trigger derivations that violate Faithfulness in part, or entirely, whereas other underlying representations trigger derivations with a bottom line that satisfies Faithfulness. It is to be noted that Faithfulness is a formal condition on the relation between input and output; a claim put forward only by OT, and support for this claim has been largely theoretically internal.

The independent evidence for Faithfulness understood as the claim that underlying and surface representations preserve certain underlying properties is not compelling. The fact that the analysis of every language that has been seriously studied has invariably revealed crucial differences between underlying and surface representations suggests that Faithfulness is of no more than marginal importance. Since there is nothing for phonology to do where Faithfulness successfully suppresses differences between underlying and surface representations, the existence of phonology in every language shows that Faithfulness is at best an ineffective principle that might well be done without.

As noted in Brumberger & Hale (1988) the fundamental fact that makes phonology different from syntax is that the words of a language are not part of our innate linguistic knowledge, but rather are learned; i.e., stored in memory. When we speak we take something out of memory and use it to produce a linguistic signal by appropriate gymnastics of our articulators. When spoken language is parsed, the phonology in the human language system is a part of the Morphology module, which also contains the Readjustment rules. Once the Morphology has done its work the terminal string is operated on by the rules of the phonology. It is suggested further in Halle-Vergnaud (1987) that the rules of the phonology are assigned to four strata or blocks: I. word-internal (cyclic); II. word-internal (noncyclic); III. word sequence (cyclic); IV. word sequence (noncyclic). As explained in Halle-Mohan (1985) the rules of the cyclic stratum apply to each cyclic constituent of the word phrase in turn beginning with the innermost constituent and proceeding outward. The rules of the noncyclic stratum II apply only once to the entire word. Whether a constituent is or is not cyclic is an idiosyncratic property of its components. Thus, English words formed with the verb suffix \(-\text{ing}, \text{-ED}\) and the modal auxiliary \(\text{be}\) are non-cyclic constituents, whereas stems formed with \(-\text{ous}, \text{ic}, \text{-tion}, \text{-al}\) are cyclic.

Since the English stress rules are part of the cyclic stratum I and do not figure in stratum II, noncyclic suffixes such as \(-\text{ing}\) and the noun-forming \(-\text{al}\) are stress-neutral. Since the