

- Stepanov, Artur. 1998. On *wh*-fronting in Russian. In *NELS 28*, 453–467. GLSA, University of Massachusetts, Amherst.
- Stjepanović, Sandra. 1998. *Scrambling in Serbo-Croatian*. Ms., University of Connecticut, Storrs.
- Strakhov, Natalia. 2001. A scrambling analysis of *wh*-questions in Russian. In *Formal Approaches to Slavic Linguistics: The Bloomington Meeting 2000*. Michigan Slavic Publications, University of Michigan, Ann Arbor.
- Švedova, Natalia, ed. 1980. *Russkaja grammatika*. Vol. 2. Moscow: Nauka.
- Vallduví, Enric. 1992. *The informational component*. New York: Garland.
- Webelhuth, Gert. 1989. Syntactic saturation phenomena and the modern Germanic languages. Doctoral dissertation, University of Massachusetts, Amherst.
- Yadroff, Mikhail. 1992. The syntactic properties of adjunction (scrambling in Russian). Ms., Indiana University, Bloomington.
- Yokoyama, Olga. 1986. *Discourse and word order*. Philadelphia: John Benjamins.
- Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Mass.: MIT Press.

Department of Linguistics
 Social and Behavioral Sciences
 SUNY at Stony Brook
 Stony Brook, New York 11794-4376
 jbailyn@notes.cc.sunysb.edu

Zero Derivations

Luigi Burzio

Rubach (2000) proposes a modified version of Optimality Theory (OT) that features derivations. While Prince and Smolensky's (1993) original formulation requires some modification, I argue here that, rather than reintroducing derivations, the correct approach is to take fuller advantage of OT's inherent parallelism. I propose that outputs must be related not only to inputs, but to other, "neighboring" representations as well—a feature that is shared by both the output-to-output faithfulness approach and the theory of targeted constraints developed by Wilson (2000, to appear). I show that all the cases cited by Rubach that seem to support derivations are in fact handled by the latter two related theories, and that both of these have significant advantages over derivations.

Keywords: derivations, opacity, output-to-output faithfulness, targeted constraints

1 Introduction

Rubach (2000) proposes a version of Optimality Theory (OT) that features derivational levels or stages: Derivational OT (DOT), a development also advocated by Kiparsky (1998). In this article

I am grateful to Bob Ladd, Colin Wilson, and two anonymous reviewers for insightful comments and suggestions.

I defend an alternative conception, which has no serial derivations and is instead fully parallel. I have argued elsewhere (Burzio 2000b) that a single modification of Prince and Smolensky's (1993) original formulation provides a unitary solution to a number of outstanding problems, including the ones examined by Rubach. The modification consists of allowing the calculation of an output in OT not just to rely on a unique "input," but to make comparisons with other representations as well, as in the output-to-output faithfulness approach that Rubach argues against. Unlike reintroducing derivations, this modification is a natural one, since it merely exploits the inherent character of OT, namely, its parallelism.¹ To put it differently, it would actually be an accident for OT if outputs were always calculated from a single input, since—in that architecture—they can be calculated from multiple representations simultaneously just as well, by concurrent application of multiple faithfulness constraints. The view that there is a single input is arguably just a carryover from rule-based systems, in which one-to-one mappings were a contingency of the rewrite rules. If one makes a commitment to that element of rule systems, it is not surprising that other elements may be found necessary as well, just as Rubach claims. My aim in this article is to remove the arbitrary premise that mappings are only one-to-one and, with it, the conclusion it entails—that there are derivations.

I will examine the cases that Rubach adduces in support of derivations, dividing them into three categories: (a) cases that can alternatively be handled by means of output-to-output faithfulness (OO-FAITH), as Rubach acknowledges; (b) cases of opacity (counterbleeding/counterfeeding) that can alternatively be handled by some appropriate parallel theory of opacity—a possibility that Rubach also acknowledges by considering McCarthy's (1999) sympathy theory; (c) other cases. I will argue that the last category includes only spurious uses of derivations, that is, cases that are adequately handled even by standard OT means. If this classification is correct, Rubach's argument for derivations does not go beyond the claim that parallel solutions are heterogeneous (output-to-output faithfulness and some other theory (perhaps sympathy) to deal with opacity), while the serial solution is unitary (derivations). I will show that this claim is incorrect for two reasons: (a) there exists a well-motivated theory of opacity, developed by Wilson (2000, to appear), which can be naturally related to output-to-output faithfulness; (b) there is in fact no relationship between the derivational surrogates that Rubach proposes for the above two theories of output-to-output faithfulness and sympathy, other than the relatively trivial one that they are both derivational.

2 Output-to-Output Faithfulness Cases

Rubach (p. 289) proposes the following type of derivation for Bulgarian, to be utilized also for Czech and rural Polish:

(1)		<i>Gloss</i>	<i>Critical ranking</i>
a. Word level	/amerika/ → ?amerika	'America'	ONSET >> *?
b. Phrase level	s ?amerika → s ?amerika	'with America'	MAX >> *?

¹ Obviously, Rubach's proposed partial serialism is not *logically* impossible within OT. Rather, it only seems conceptually unnatural.

S \varnothing amerika in (1b) is a case of opacity: specifically, counterbleeding, since the presence of the preposition *s* ‘with’, which would provide the needed onset, fails to bleed \varnothing -epenthesis. Although not all cases of opacity can be accounted for by output-to-output faithfulness, some—like this one—can, as Rubach admits. His account is in fact virtually identical to an output-to-output faithfulness account, and no derivation is entailed by it. That is, the two calculations in (1) can be seen as proceeding in parallel without any difficulty. One need only postulate one principled difference between the grammars that apply in the two cases: calculation of phrases is constrained by faithfulness to the form that words have in isolation (the MAX of (1b) to be understood on this view as MAX_{OO}), while calculation of words in isolation is generally insensitive to the structure of the phrases that may contain them. This is Benua’s (1997) Base Priority Principle, which relates larger units to their parts, in a nonderivational way.² While Rubach’s derivational analysis and the nonderivational output-to-output faithfulness analysis of (1) may just seem inconsequential restatements of one another, other considerations distinguish the two approaches. Consider that phenomena like the one in (1b) are found as well with English level 2 affixes. For example, unlike in level 1 *metr-ic*, which is regular (Borowsky 1993), in level 2 *meteR-ing* the process that causes word-final *r* to be syllabic ‘overapplies’ just as \varnothing -epenthesis does in (1b). Here, Rubach’s derivational approach would postulate a level 1 stage, feeding level 2. At level 2—and only at level 2—faithfulness will dominate markedness as it does in (1b), ensuring that *meteR-ing* will mirror *meteR*, which is the output of level 1. This is insufficient, however. Level 1 and level 2 formations differ simultaneously along two independent dimensions: phonological (regular application vs. overapplication of phonological processes) and morphological. Unlike level 2 formations, level 1 words exhibit widespread morphological irregularity; this is illustrated in (2), where the capitalized portions are absent in the respective bases, (2e) and (2f) instantiating the phenomenon of bound stems.

(2) *Irregular level 1 morphology (augmented, distorted, and bound stems)*

- | | | |
|-----------------|---|--------------|
| a. crime | → | crimIN-al |
| b. problem + ic | → | problemAT-ic |
| c. horizon + al | → | horizonT-al |
| d. compel + ive | → | compUIS-ive |
| e. ?? | → | ARBORE-al |
| f. ?? | → | INEVIT-able |

Since such irregularities are systematically absent with level 2 formations (cf. *crime-less*, *compelled*, etc.), the generalization is thus that their stems are strongly required to equal words (Burzio 1994b:274). This requirement has the effect of suppressing both the kind of distortion illustrated for level 1 in (2) and certain phonologically driven alternations like *meteR/metr-ic* or the other

² Rubach (fn. 31) finds it ‘uncontroversial’ that words are derivationally prior to phrases. However, it is only the weaker conclusion that words are *epistemologically* prior to phrases that is—relatively—uncontroversial. The stronger conclusion that such priority is expressed by discrete derivational stages would need justification. Even the weaker conclusion can only be relatively uncontroversial, since it seems only partially true. For instance, the first person singular present of the verb *be* is expressed by different words depending on its phrasal context: *Am I not?* versus *Aren’t I?*

familiar ones in (3), all absent at level 2 (witness *cri:me-less* (no shortening), *frolick-ing* (no velar softening), *párent-hood* (no restressing)).

(3) *Regular level 1 phonology*

- a. *Vowel shortening*: *cri:me/crimin-al*, *na:ture/natur-al*, *admi:re/admir-able*, . . .
- b. *Velar softening*: *authentic/authenti[s]-ity*, *critic/criti[s]-ism*, . . .
- c. *Restressing*: *párent/parént-al*, *títan/titán-ic*, . . .

This inverse correlation of phonological and morphological regularities (*irregular* morphology/*regular* phonology for the level 1 lexicon vs. *regular* morphology/*irregular* phonology for level 2), is what the output-to-output faithfulness approach expresses: high-ranked OO-FAITH for the level 2 lexicon will both inhibit regular phonological alternations and suppress morphological idiosyncrasy, while low-ranked OO-FAITH for level 1 will permit both. As for the source of morphological idiosyncrasy, this is taken to be input-to-output faithfulness (IO-FAITH). In the approach developed in Burzio 1994b, et seq., morphologically complex forms are related to their bases only via their surface forms (OO-FAITH), rather than by a shared underlying representation (IO-FAITH). This allows IO-FAITH to be deployed with morphologically derived forms in the same capacity as with morphologically simple ones: to express what is item-specific and unpredictable (see Burzio 2000a, to appear a,b, for details, and Benua 1997 for a partially similar argument). In contrast, in the derivational approach the morphological subpart of the generalization requires the assumption that level 1 may only output actual words (no bound stems, no augmented or distorted words like *crimIN*, etc.). However, that is quite unrelated to the assumption about level 2 phonology needed to express the other half of the generalization. The two separate provisions thus form a conspiracy to the effect that stems of level 2 formations must equal words. Such conspiracies are a textbook argument against derivations and for surface constraints (Kager 1999: 56), in this case supporting output-to-output faithfulness.

In fact, derivations appear irrelevant even to the strictly phonological side of the generalization. What yields the difference between *metr-ic* and *meteR-ing*, or between shortening *crim-in-al* and nonshortening *cri:me-less*, is the assumption that FAITH (whether IO- or OO-) is higher-ranked with level 2 than with level 1 affixes. The sequential derivation itself contributes nothing, just as in the case of the Bulgarian example in (1b).

In sum, derivations are both unnecessary and insufficient: phonological generalizations hinge on high-ranked versus low-ranked FAITH (derivations are unnecessary); morphological generalizations hinge on the notion that faithfulness holds between words (outputs). This is true by definition under a theory that includes OO-FAITH, but needs stipulating under one based on derivations (derivations are insufficient).

Failure to distinguish IO- from OO-FAITH as in Rubach's system has further puzzling consequences, at least in English. The reason is that it is not a generalization about IO-FAITH that its rank increases at higher levels of composition like level 2 and phrases. While FAITH must be lower-ranked for level 1 than for level 2 so as to permit vowel shortening in (e.g.) *crim-in-al*, while blocking it in *cri:me-less*, it must nonetheless be relatively high-ranked for underived items in general so as to block the vowel shortening in those items (e.g., *cri:me*, *vi:tamin*, *di:nosaur*).

If all FAITH is IO, the question arises why its rank varies nonmonotonically over what would otherwise seem a descriptively well motivated progression of morphological domains—underived, level 1, level 2, phrases—that rank going from high, to low, to high again. This puzzle disappears if one recognizes *both* IO- and OO-FAITH. On this view OO-FAITH is simply irrelevant to underived items, which are thus calculated from IO-FAITH alone. In contrast, at least in the conception developed in Burzio 1996, 2000a, to appear a,b, in which there is no underlying representation, derived items are calculated solely from OO-FAITH (aside from the irregularities of (2), which require partial inputs as noted). On this general conception, OO-FAITH can then be seen as re-ranking upward monotonically from level 1 to level 2 to phrases. On this view the problem of nonhomogeneity of the grammar over different morphological domains, which is theory-independent (after all, phonology just works differently over level 1/level 2/phrases) will reduce to the question of why the rank of OO-FAITH rises (now monotonically) as the size of the units grows (phrases being larger than words, and level 2 formations apparently being quasi-phrasal from this perspective). As for the rank of IO-FAITH, it never varies at all. It is simply ranked high enough to block vowel shortening and other processes (Burzio 2000a). The next few remarks address the question of the apparent dependency between unit size and rank of OO-FAITH.

I have argued in Burzio, to appear a,b, that the effects encapsulated in the notion of OO-FAITH are the combination (by some appropriate operation—e.g., summation of their scalar values, or the ‘local conjunction’ of more familiar OT work) of two different effects. The first, which I refer to as *SELECT*, is simply a specific expression of the needed combinatorial principle. The second component, which I term *gradient attraction*, pertains to independent similarity of representations in multidimensional space. Beginning with the first component, consider for instance that *-less* attaches only to nouns. This is expressible as the conditional *LESS-SELECT*: $-less \rightarrow / N \text{ ____}$, requiring that, if *-less* is present, there must be a noun preceding it. If we simply take *LESS-SELECT*—which is effectively a subcategorization frame—to be a violable constraint in OT, it can then be assigned a high rank, given that *-less* is a level 2 affix (*crime-less*, etc.). In contrast, the similar conditional *AL-SELECT*: $-al \rightarrow / N \text{ ____}$ will be assigned relatively low rank, given that *-al* is level 1 (*crimin-al*, etc.). Since OO-FAITH equals *SELECT* + *gradient attraction*, this will result in OO-FAITH of two different ranks for level 1 and level 2 affixes (on *gradient attraction*, see below), consistent with Benua’s (1997) idea that affixes are subcategorized for OO-FAITH of different ranks. Similarly, the OO-FAITH imposed by the preposition *s* ‘with’ on its complement in Bulgarian can be directly related to the subcategorization frame for the preposition, *s-SELECT*: $s \rightarrow / \text{ ____} NP$, evidently high-ranked, making the resulting OO-FAITH dominant over *[?]. In the case of Russian [?] *Amerikals Amerikoj* ‘America/with America’, which, as Rubach (p. 289) notes, contrasts minimally with its Bulgarian counterpart in lacking the overapplication of ?-epenthesis, *s-SELECT* can be taken to be lower-ranked, resulting in OO-FAITH now being dominated by *[?] (hence reversing the rank in (1b), while *[?] remains dominated by *ONSET*, just as in (1a)).

As I argue in Burzio 1997, to appear a,b, larger units correspond to larger class sizes. For example, a preposition like *with*, which combines with NPs, produces a larger class of structures (PPs) than does an affix like *-al*, which combines with Ns (producing adjectives). The reason is

simply that there are more NPs than there are Ns, each N yielding multiple NPs when it itself combines with other constituents. Hence, the same combinatorial properties that yield larger units from smaller ones, also yield larger classes. Now the fact that there are more instances of *with*-NP than there are of *N-al* will—as desired—result in a higher-ranked conditional, *with* → / ____ NP, than the corresponding conditional *-al* → / N ____ on the assumption that the strength of such conditionals simply accrues over experience. Since there are more instances of *with*-NP than of *N-al* in the experienced data, the former conditional will acquire greater strength/higher rank than the latter. Since such SELECT conditionals are one of the bases of OO-FAITH, the latter constraints will then be correctly expected to be higher-ranked at phrasal levels than at the word level.³ In addition to their rank, the conditionals themselves can be directly linked with experience on the hypothesis I advance in Burzio, to appear a,b, that mental representations of linguistic forms are sets of entailments, each entailment expressing dependency of some aspect of the representation on some other, as if any cooccurrence were mentally registered as necessary. On this view a conditional such as *-al* → / N ____ is an “emergent” property of the lexicon, due to the collective effect of all representations of the type *N-al* (see Burzio, to appear a,b, for more exact motivation). In sum, on the proposed view output-to-output faithfulness is motivated independently (aside from gradient attraction, discussed below) as just a specific interpretation of the type of selectional dependencies that are generally referred to as subcategorization specifications, and which any theory needs to express in some form. The present innovation is in taking them to be violable constraints, and in taking their rank to be driven by the frequency with which they are instantiated.

Reliance on derivations in characterizing word-to-word relations is undercut by two further considerations, one of which will lead us to the other basis of OO-FAITH: gradient attraction. This consideration is the simultaneous existence of multiple word-to-word relations. This claim is advanced independently in Burzio 1998 and Steriade 1999, on the basis of examples like (4a–b), respectively.

(4) a. *Italian*

vínc-ere	vín-t-o	vinc-it-óre
‘win (INF)’	‘won (PART)’	‘winner’

b. *French (dialect)*

[grɔs]	[gro]	[groz] arbre
‘large (FEM)’	‘large (MASC)’	‘large (MASC) tree’

In (4a) the agentive noun in *-óre* reveals a relationship with the participle *vínto* by way of its participial affix *-it-*, related to the syncopated allomorph *-t-* of *vínto*. This relationship with the

³ The notion that OO-FAITH more strongly inhibits phonological alternations at phrasal levels than at word levels may seem challenged by languages that have extensive sandhi phenomena, such as Sanskrit (as one reviewer points out). I am not in a position to address such cases here. Note, however, that the existence of a generalization along the lines suggested in the text is confirmed by the independent existence of the Strong Domain Hypothesis once proposed in the context of Lexical Phonology (Kiparsky 1985). According to that hypothesis, no new phonological rules are permitted to be introduced at higher levels of compositionality. Rather, rules can only be discontinued at those levels, attesting to the increasing paucity of phonological alternations at higher levels.

participle is the normal one for *-óre* nouns in Italian (cf. *adatt-at-ol/adatt-at-ore* ‘adapted/adapter’, etc.). However, in (4a) the noun also reveals a relationship with the infinitive *vincere*, by sharing its consonant *c* [č], which the participle lacks. The case in (4b), from one dialect of French, is similar in that the liaison form [groz] shares the vowel of the masculine citation form, but the final consonant of the feminine form, aside from voicing (see Burzio 1999, Steriade 1999 for exact analyses). The existence of such multiple dependencies is expressible—in the form of multiple application of OO-FAITH—in the parallel system advocated here, but not by serial derivations. For instance, in (4a) the *-ore* noun could not be derived by calculating the structure of the inner participle first in cyclic fashion, since then the resurfacing of the underlying sequence /ci/ would be inexpressible ([vin-t]-o → [vin??-t]-ore). On the other hand, deriving the *-ore* noun from underlying material would miss important similarities between the participle and the noun and is therefore also excluded. As noted in Burzio 1998, syncope is variable in both the participle and the nouns, whence, alongside (4a), *tin-t-oltin-t-ore* ‘dyed/dyer’ (syncope in both), *aggred-it-ol/aggres-s-ore* ‘attacked/attacker’ (syncope in the noun only). In addition, syncope results in either an *-s-* or a *-t-* allomorph (from nonsyncopeated *-it-*), variably and unpredictably. Nonetheless, whenever both participle and noun syncope, they never fail to exhibit the same allomorph, as in *opprim-ere/oppres-s-oloppres-s-ore* ‘oppress/oppressed/oppressor’ (*-s-* allomorph) versus *redim-ere/reden-t-ol/reden-t-ore* ‘redeem/redeemed/redeemer’. Hence, at least for syncopeated nouns, the derivation would have to be cyclic. This is falsified, however, by the noted *aggred-it-ol/aggres-s-ore* case, with syncope in the noun only. Hence, there are no viable derivational routes for capturing the observed multiple dependencies (for more specific discussion, see Burzio 1998).

The surface-to-surface approach to (4) will of course raise the question of how an output-to-output faithfulness relation could exist between the *-ore* noun and the infinitive in (4a) if in fact *-ore* attaches to participles, and more generally how multiple correspondences arise on that approach. I argue in Burzio, to appear a,b, that such relations result from a general effect, *gradient attraction* (GA), by which similarity between representations generates pressure for further similarity. GA is the second component of OO-FAITH anticipated above, and, like SELECT, it results from the hypothesis that representations are sets of entailments. Hence, SELECT and GA are not independent provisions of the OO-FAITH machinery, but ultimately result from the same set of primitives (see Burzio, to appear a,b). In addition to these conceptual underpinnings, GA is independently motivated by the fact that, even for the same affixational class, faithfulness effects defy fixed ranking, instead being either stronger or weaker depending on overall similarity between trigger and target. For instance, in *remé:di-able* the constraint responsible for *CiV* lengthening (as in *CanadalCana:dian*, etc.) must outrank OO-FAITH relative to the base *remedy*, but the opposite must be true in *lévi-able* (short V) from *lévy*. The paradox is resolved by noting that *remedi-able* differs from its base verb in stress, a difference that weakens the ‘attraction’/OO-FAITH, while *leviable* has the same stress as *levy* and is thus subject to greater attraction/OO-FAITH (for further examples, and consequences, see Burzio 2000b, to appear a,b). Now in *vinc-it-óre* (4a), the stem *vinc-* will be subject to GA from both participle and infinitive, given its similarity to both in sound and meaning, but the participle will exert the greater attraction/higher-

ranked OO-FAITH because it is also “subcategorized for” by the affix *-óre* and hence also benefits from SELECT, the other component of OO-FAITH. As argued in Burzio 1998, one of the ways in which the participle exerts its influence/attraction is accentual: OO-FAITH_{stress}, demanding that the stem in *vinc-it-óre* carry the same stress as *vín-t-o*. Given the stress on *-óre*, this requires the presence of an intervening syllable (to avoid a stress clash), and that is where the infinitive—the weaker attractor—comes in, lending needed segmental material (again, see Burzio 1998 for full discussion). Somewhat similarly, in (4b) the liaison form is expected to utilize the masculine citation form as its primary base, though here this is not because of affixal effects, but simply because the liaison form is itself masculine (greater independent similarity entailing stronger attraction). At the same time high-ranked ONSET forces divergence from that base, giving the feminine form (a secondary base/attractor) the opportunity to exert its influence, again lending needed material (see Steriade 1999 for details).

The second argument against derivations, besides the one based on multiple correspondences, relates to the behavior of affixes, illustrated by examples such as these:

- | | | | |
|--------|---------------------|--------------------------|-------------------|
| (5) a. | propagánda/Américan | propa(gánd-is)tθ | A(mérica)(n-ístθ) |
| b. | títan | *(títan-i)cθ/ti(tán-icθ) | bar(bár-icθ) |

As argued in Burzio 1994a,b, the stress-neutrality of (level 1) affixes like *-ist* in (5a) is best accounted for by taking the “metrical consistency” of the stem (a specific form of OO-FAITH) to prevail over that of the suffix, whose right edge can thus be parsed in one of two possible ways as indicated (similarly to what happens, for example, with the two independent items (*hónes)tθ/ro(bústθ)*). In (5b), however, matters are reversed, as the metrical consistency (unique parse) of the suffix prevails over that of the stem, excluding stress-preserving **(títan-i)cθ*. In other words, the pressure for *títan-ic* to stress like *títan* is successfully countered here by the pressure for *-ic* to parse in the same way as it does elsewhere in the lexicon (e.g., as in *barbar-ic*). From the present perspective, this is just another GA/OO-FAITH effect: same segments, same meaning entail the same metrical parse. In a derivational system relying solely on IO-FAITH, the metrical invariance of *-ic* or any other stem-restressing suffix must be attributed to properties of the input. This solution is incorrect, however. As argued in Burzio 1994b:sec. 8.4, the divide between restressing and stress-preserving Latinize suffixes is predictable from output properties. The restressing suffixes like *-ic* and *-al* are those that do not provide enough material for a full metrical foot. Thus, unlike in *A(mérica)(n-ístθ)*, where both feet are well formed, thus allowing preservation of stem main stress, in (e.g.) **Na(póleo)(n-icθ)* the final foot would be ill formed. Rather, well-formedness requires one more syllable, as in *Na(pòle)(ón-icθ)*; but the resulting foot calls for main stress, with consequent stress shift from *Napóleon*. The generalization is that suffixes that are necessarily stem-restressing under some circumstances choose to avoid *their own* allomorphy by parsing in a fixed manner as in (5b). Hence, in the competition between stem and suffix OO-FAITH, the winner is whichever form of OO-FAITH is the more productive. While the problem of tying rank of OO-FAITH to productivity may be complex, requiring calculations over lexicons, the input approach is incorrect, because it predicts no generalization. Note also in this regard that the distinction between *-icθ* with fixed presuffixal stress and other suffixes like *-al* or *-ous* with penultimate/antepenultimate stress (e.g., *bar(bár-icθ)* versus *(bárbar-ou)sθ*) is also not a matter

of input specifications. Rather, as argued in Chomsky and Halle 1968, the special metrical properties of *-ic* are best attributed to its massive alternation with *-ical*, as in *aca(dém-ic∅)/aca(dém-ica)l*. As argued in Burzio 1994b, the special parse of *-ic∅* is the only option available to satisfy the expected stem OO-FAITH over such pairs. The fact that the same parse obtains in the absence of an *-ical* variant (*Napole(ón-ic∅)??Napole(ón-ica)l*) then requires OO-FAITH over different instances of *-ic*. In sum, as a matter of factual description, affixes tend to metrical invariance just as much as stems. Derivational approaches introduce an unwarranted asymmetry: the invariance of stems is the result of cyclic derivations, while the invariance of affixes is the result of input prespecification. This conclusion is incorrect, because the exact modalities by which affixal invariance obtains are predictable from comparing outputs.

Recapping this section: Rubach admits that the facts in (1) provide no argument for derivations over output-to-output faithfulness. Other facts, however, support output-to-output faithfulness over derivations.⁴ In particular:

- There is a morphophonological generalization about English level 2 affixation that has a unitary output-to-output faithfulness account but results in an unexplained conspiracy under a derivational account.
- Any theory of phonology has to postulate some grammar-internal modulation across different morphological domains such as underived; level 1; level 2; phrases. However, the failure to distinguish input-to-output faithfulness from output-to-output faithfulness inherent in Rubach's system yields a nonmonotonic fall-rise distribution in the ranking of IO-FAITH constraints distributed over what would otherwise seem a natural progression.
- Output-to-output faithfulness can be enforced simultaneously from multiple bases. It also applies across different instantiations of the same affixal material. In Burzio, to appear a,b, I lay out a framework in which the scope of output-to-output faithfulness can be defined in terms of GA and the selectional properties of affixes.⁵ In contrast, neither effect is expressible by derivations.

⁴ Rubach (p. 303) also advances the following argument against output-to-output faithfulness: Given an alternation like Bulgarian *obet/obed-i* 'dinner-SG/PL' due to final devoicing, output-to-output faithfulness predicts that a language may come to level all such alternations as in *obet*obet-i*, which 'is unheard of.' Such leveling would result from the ranking IDENT_{OO}([voice]), FINAL-DEVOICING >> IDENT_{IO}([voice]).

This argument is successfully challenged by the following considerations: (a) It is not clear why such a situation in which final devoicing overapplies would be 'unheard of' given its parallelism to the *meteR/meteR-ing* case reviewed in the text, which requires a ranking just like the one Rubach argues against. A similar case is that of younger-generation Korean discussed in Kenstowicz 1996, in which the alternation *kap/kaps-i* 'price (CITATION FORM/NOM)', due to underlying /kaps/ and final cluster simplification, is now being leveled to *kap/kap-i*. (b) More importantly, Rubach's own system allows this situation to arise under the following grammar: final devoicing at level 1; plural affixation with undominated IDENT([voice]) at level 2.

There seems no way to exclude *obet*obet-i* in the 'language change' scenario that Rubach considers, either. Whichever generation first drops final devoicing will have a choice of basing its underlying representation either on the old singular *obet*, producing exactly the unwanted *obet*obet-i*, or on the old plural *obed-i*, producing **obed/obed-i*. There is no reason to believe that learners privilege plurals in setting up underlying representations.

⁵ That framework also captures the noted 'base priority' effect; accounting for this effect is therefore not an advantage for derivations. More specifically, it captures the stronger tendency for morphologically simpler forms to influence their derivatives, without however categorically excluding 'back-copying' as in the previously noted *académicallacadémic* case (see Burzio, to appear a,b).

3 Opacity Cases

If conspiracies are a textbook argument against derivations, counterfeeding/counterbleeding effects have been an argument *for* them (Kenstowicz 1994:chaps. 2, 3). Rubach (pp. 280–283) uses derivations in the following two cases of phonological opacity in Slovak:

(6) <i>Opaque palatalization</i>		<i>Critical ranking</i>
a. Level 1	Palatalization: /pan-æ/ → pañ-æ	IDENT([-back]) ≫ æ-BACKING
b. Level 3	æ-backing: pañ-æ → pañ-a (*pan-a)	æ-BACKING ≫ IDENT([-back])
(7) <i>Opaque diphthongization</i>		<i>Critical ranking</i>
a. Level 2	Diphthongization: /ræ:sa/ → ryæsa	IDENT([-back]) ≫ æ-BACKING
b. Level 3	æ-backing: ryæsa → ryasa (*rwas)a	æ-BACKING ≫ IDENT([-back])

Rubach's use of level 1 for the case in (7) will be discussed below.⁶ In (6) æ-backing counterbleeds palatalization, which otherwise occurs only before front vowels. In (7) æ-backing counterfeeds diphthongization in the sense that the back vowel *a* would be expected to select a back rather than a front glide. It also counterbleeds diphthongization altogether in that long [a:] does not diphthongize in Slovak. In Rubach's analysis it is therefore necessary in both cases that an underlying /æ/ be changed to [a] only late in the derivation. This is done by the "last-minute" change in the critical ranking indicated in (6) and (7). The question is, how can a last-minute effect be captured nonderivationally? Here again Rubach acknowledges the existence of a nonderivational alternative, in the form of McCarthy's (1999) sympathy theory. As Rubach notes, this theory permits the nonderivational identification, among all the candidates, of a "sympathetic" candidate that effectively has next-to-the-last-minute properties. For the case in (6) this candidate is *pañ-æ* and for the case in (7) it is *ryæsa*, identical to the input of Rubach's level 3. In McCarthy's system the correct output forms *pañ-a* and *ryasa* then result by assuming that faithfulness to this candidate appropriately dominates faithfulness to the input, thus ensuring a palatalized *ñ* in (6) and a diphthong in (7). Assuming further that æ-backing is undominated will ensure the correct output [a] rather than [æ] in both cases.⁷

⁶ In (6) I have deliberately changed level 2 of Rubach's (20) to level 3 to amend what seems to be an inconsistency. If (6b) were level 2, its associated ranking would contradict that of (7a).

⁷ Rubach (sec. 9.2) also considers and rejects a MAX(FEATURE) analysis of (6) and (7). Such an analysis would view the palatalization in (6) and presumably the frontness of the glide in (7) as preservation of the input feature [-back] present in underlying [æ], mirroring a proposal in Lombardi 1998. While the value of such an analysis of (6)–(7) may be limited since it would not generalize to other cases of opacity, Rubach's argument against it is dubious. He suggests that, on this analysis, /tel-æ/ should give *[t'el-a] with nonpalatal *l* rather than actual [t'el'-a] 'calf' because the independent presence of [-back] in both [t'] and [e] should suffice to satisfy MAX([-back]). Surely this conclusion must be rejected, since features are not free to metathesize this way. Consider diphthong neutralization, as for instance in French diachrony: *fl[ew]r* → *fl[ø]r* 'flower'. Here the [+round] of the former glide is picked up by the vowel, thus motivating MAX([+round]). But no one doubts that this will happen regardless of the presence of some other [+round] in the string (which would already satisfy MAX([+round]) on Rubach's assumption): for example, *doul[ew]r* → *doul[ø]r* 'pain'. Hence, features, like segments, must be constrained by some form of contiguity. Rubach himself (fn. 39) notes the mirror image phenomenon in the Polish pronunciation of French [ü] as [yu], where the glide must be the effect of MAX([-back]) pertaining to the input vowel. Here again one can safely assume that the process will not fail in the presence of some other [-back] in the string.

Given Rubach's admission that there are nonderivational alternatives to both sets of cases discussed so far, at this point his argument for derivations can only rest on the disparate character of the output-to-output faithfulness and sympathy theories. However, such an argument is difficult to uphold, because Rubach's own analyses of the two sets of cases are rather unrelated to one another. In the case in (1) the two levels correspond to words and phrases as noted. But this is not true of the cases in (6) and (7). In particular, although applying "late," æ-backing appears to apply not at the phrase level, but "earlier." Since æ-backing in Slovak applies only after nonlabials, its application at the phrase level would predict word-initial [a]/[æ] alternations controlled by a preceding word-final segment. To the best of my knowledge, such alternations do not exist. Hence, the levels needed for (1) are unrelated to those needed for (6) and (7), leaving only the serial character of Rubach's solution to unify them. One could as easily claim that output-to-output faithfulness and sympathy are similarly unified by their common parallel character, leaving no argument in favor of Rubach's solutions.

More importantly, however, there exists a parallel account of opacity, which is well motivated, is superior to derivations, and can be naturally related to output-to-output faithfulness. This is the theory of targeted constraints developed by Wilson (2000, to appear). Wilson addresses the problem of opacity in OT by way of another problem, which he finds to be related. He notes that certain types of contextual markedness constraints of the form "Avoid X in context Y" are satisfied only by manipulating X, and never Y. Specifically, while one could say that it is marked for one obstruent to directly precede another, the crosslinguistically attested cluster simplification consists only of deleting the first, not the second, as in the Diola-Fogny example *let-ku-jaw* → *lekujaw*/**letujaw* 'they won't go'—this despite the fact that either outcome would equally satisfy the surface prohibition. Wilson accounts for this asymmetry by introducing a set of what he terms *targeted* constraints, which effectively condition the repair in the right way. Targeted constraints formalize the role of contrast in phonology, along the lines of the dispersion theory developed by Lindblom (1986) and Flemming (1995), and they build on Steriade's (1994, 1997) claim that the positions subject to neutralization of segmental contrasts are those that limit the expression of perceptual cues relevant to the contrast. Targeted constraints penalize weak contrasts, while expressing the fact that weakness of contrast favors neutralization, as exemplified in (8), which yields the correct cluster simplification in the Diola-Fogny case.⁸

(8) →**WEAK-C*

- a. Avoid an obstruent in preobstruent position.
- b. Candidate set: candidates that differ exactly by the weak element

This constraint expresses the fact that, owing to weakness of the relevant perceptual cues, an obstruent in preobstruent position effectively stands in a weak contrast with zero: (8a). By restricting the candidate set in the manner of (8b), it also expresses the fact that zero is a better alternative to the latter obstruent, and that nothing else is a better alternative for this particular constraint (other candidates, with more varied structural properties, are irrelevant). Diola-Fogny cluster

⁸ The specific notation and some of the examples used in this discussion differ slightly (and inconsequentially) from Wilson's (2000).

simplification can now proceed as in (9), where I ignore epenthetic candidates (violating DEP) to simplify discussion.

(9)

	let-ku-jaw	→*WEAK-C	DEP	MAX	*VELAR	*CORONAL
a.	letkujaw	*		a > b, c	*	*
b.	^{ERASE} lekujaw	b > a		*	*	b > a, c
c.	letujaw	—		*	c > b, a	*
	Harmonic ordering	b > a	same	b > a > c	same	same

In (9) the asterisks indicate constraint violation as in standard OT notation. The cells unoccupied by asterisks are used for the harmonic ordering of candidates given by each constraint. For constraints other than the targeted one, this information is redundant with the asterisks. For example, the notation ‘‘a > b, c’’ under MAX indicates that candidate (9a) is better than candidates (9b) and (9c) according to MAX. The reason for this is that (9a) does not violate MAX, unlike (9b) and (9c) (as shown by the asterisks). The targeted constraint →*WEAK-C, however, only compares candidate (9a), which violates it, with candidate (9b), in which the weak C has been removed, as required by (8b). Candidate (9c), in which something other than the weak C has been removed, is irrelevant on this approach. The bottom row in (9) then tallies the overall harmonic ordering of candidates, from left to right, prioritizing the higher-ranked constraints (i.e., conflicts are resolved by rank, as usual). On this analysis, candidate (9b) is correctly selected as the most harmonic and hence the winner (final harmonic ordering b > a > c). In contrast, as the distribution of asterisks clearly shows, standard OT would incorrectly pick (9c), so long as [coronal] is less marked than other places of articulation as expressed in the tableau and as generally assumed. The crucial factor in the correct selection is thus the exclusion of candidate (9c) by the targeted constraint.

Note that the cluster simplification problem, while unsolvable within standard OT, would appear to have a solution within a derivational theory of syllabification such as that of Itô (1986). In that framework one can postulate a syllabification stage banning medial codas, yielding *le<t>.ku.jaw*, followed by ‘‘stray erasure,’’ yielding the correct output. This analysis can presumably be reproduced in any framework that can duplicate Itô’s two stages, including DOT.⁹ The derivational route seems precluded, however, by Steriade’s (1994, 1997) arguments that the contexts of neutralization are not always correctly identified by notions of traditional syllable theory such as ‘‘coda.’’ Rather, she argues that proper perceptual cuing is the relevant factor, such that a coda would be properly cued—on a par with an onset—if followed by a sonorant. If such arguments are accepted, as they are here or in Wilson 2000, to appear, a derivational account is excluded by the fact that perceptual cues are plainly properties of outputs, not of intermediate

⁹ It can in fact be reproduced even in the original formulation of OT (Prince and Smolensky 1993), where faithfulness was defined in terms of PARSE and FILL, as shown in Zec 1995.

representations. Indeed, in Wilson's analysis the bound-for-neutralization candidate (9a) is a hypothetical output, not an intermediate representation in a derivation.¹⁰

Returning to opacity: Wilson shows further that targeted constraints can successfully deal with it as well. On that approach the [æ] → [a] neutralization of (6)–(7) would instantiate the work of a targeted constraint like (10), parallel to (8).

(10) →*WEAK-æ

- a. Avoid [–back] with [+syllabic, +low], except after labials.
- b. Candidate set: candidates that differ exactly by the weak element

This constraint is meant to express the fact that, except after labials, the [æ]/[a] contrast is a weak one in Slovak—a situation that is less harmonic than neutralization of the contrast. On this view—just as in Rubach's analysis—the reason why [æ] neutralizes to [a] rather than vice versa must be that [æ] is more marked than [a] in the traditional OT sense of markedness, which is still required in the present system along with the “weak contrast” notion of markedness. Similarly, the deleted C in (9) must be more marked than its zero alternative. One might ask for evidence that [æ]/[a] is indeed a weak contrast, but that evidence is less than necessary in the present context. The present approach is based on the proposition, advanced by Steriade and Wilson and motivated independently, that neutralization occurs under weakness of contrast. Invoking it in the present case is a legitimate use of independent assumptions. It is parallel to Rubach's use of derivations, which are a general resource of his framework. Evidence may be discovered that [æ]/[a] is in fact *not* a weak contrast, falsifying the present analysis. Similarly, evidence may be discovered that Rubach's specific derivation is untenable. Until that point the two approaches seem comparably cogent.

The constraint in (10) will now give rise to the opaque interaction of (6), as in (11).

(11)	pan-æ	→*WEAK-æ	IDENT([–back])	PALATN.	IDENT([–anterior])
a.	pan-æ	*	a>c, d	*	a>d
b.	pañ-æ	*	b>c, d	b>a	b>d
c.	pan-a	c>a	*	c>a	c>d
d. [ɛ̃]	pañ-a	d>b	*	d>a	*
	Harmonic ordering	c>a, d>b	d>b>c>a ¹¹		same

¹⁰ It is sometimes suggested that the type of phonetic grounding expressed in Wilson's analysis obtains only diachronically. The loss of the *t* in Diola-Fogny's *le(t)kujaw*, while indeed due to perceptual factors, would be the result of diachronic reanalysis. This view constitutes a weaker theoretical position, since it treats the relevant allomorphic variation as synchronically arbitrary. I see no need to abandon the stronger position, at the moment. One could equally well suggest that the stages of derivations like Itô's are to be interpreted only diachronically. The same objection would apply.

¹¹ IDENT([–back]) gives ambiguity between c>a>d>b and d>b>c>a, but PALATALIZATION selects the latter ordering.

Given (10b), the only harmonic orderings established by the targeted constraint in (11) will be $c > a$ and $d > b$ (i.e., here there are two candidate sets: $\{c, a\}$, $\{d, b\}$). While the overall harmonic ordering of candidates can be tallied in the bottom row step-by-step as before, we can more simply note that the last three constraints together favor the transparent, palatalized candidate (11b). The undominated targeted constraint can then be viewed as simply reasserting (11d) over (11b), causing it to win instead. As before, the success of the analysis hinges on the targeted constraint failing to make certain comparisons, specifically here between candidates (11b) and (11c). In standard OT, which makes all the comparisons, the transparent, nonpalatalized candidate (11c) would win, incorrectly, as is easy to see.

Calculation of the other opaque case, $/r\text{æ}:sa/ \rightarrow [ryasa]$ in (7), would be similar and is left as an exercise for the reader.¹²

While Wilson's theory shares formal similarities with McCarthy's sympathy theory, it has certain empirical advantages over it. One is that it is independently motivated by the cluster simplification problem discussed above and by the need to formalize the critical notion of weak contrast. Another is that it is relatable to the output-to-output faithfulness theory, which is also independently motivated. I argue in Burzio 2000b (though in slightly different terms) that the critical characteristic of targeted constraints—the restriction on the candidate set, as in (8b) and (10b)—is essentially a restatement of the gradient attraction (GA) effect discussed above. The reason is that saying that a representation A attracts a similar representation A' (GA) is indeed equivalent to saying that harmony would improve if A' were to be replaced by A as in the pairwise comparisons made by targeted constraints. Unlike candidates (11b,d), $[pañ\text{-}\text{æ}, pañ\text{-}a]$, candidates (11b,c), $[pañ\text{-}\text{æ}, pañ\text{-}a]$, are not relatable by GA, since attraction of $[\text{æ}]$ by $[a]$ should have no effect on the rest of the string.

Attraction under a high degree of similarity, which thus seems to capture the nature of segmental neutralizations, is also detectable in patterns of allomorphy, as shown by the pair in (12) mentioned earlier.

- (12) a. $lévy \rightarrow lévi\text{-}able$ (OO-FAITH \gg *CiV*-LENGTHENING)
 b. $rémedy \rightarrow remé:di\text{-}able$ (*CiV*-LENGTHENING \gg OO-FAITH)

These cases give rise to an apparent ranking paradox, with OO-FAITH taking on two contradictory ranks. The generalization underlying this and several other similar paradoxes examined in Burzio, to appear b, is that the dominated version of OO-FAITH is always found when base and derivative already differ along some other independent dimension. In (12b) the relevant difference, not shared by (12a), is stress. It thus appears that, just as with segmental contrast, with allomorphy proximity promotes neutralization. The apparent reranking of OO-FAITH in (12) follows from the fact that one of its main components, GA, varies in strength, while the other component, SELECT, enforced by the suffix, remains constant. One could see the high-ranked OO-FAITH/GA of (12a)

¹² Note, however, that *WEAK- æ will have to refer to short $[\text{æ}]$ only, as reference to $[\text{æ}:]$ would allow $*ra:sa$ to win. This is consistent with the reasoning in the text: vowel length enhances contrast; hence, only short $[\text{æ}]$ and $[a]$ contrast weakly.

as in fact equivalent to a targeted constraint *WEAK-LENGTH that bans the weak contrast *lévy/lé:vi-*, two structures differing only by the length of one vowel, while the corresponding contrast *remedy/remé:di-* is tolerated, because the two structures differ in stress as well and thus fail to occur in the same candidate set: (8b), (10b). Conversely, one could as well interpret the pressure to neutralize weak segmental contrasts as a form of OO-FAITH holding between close—mentally instantiated—segmental alternatives: [t] and zero in (9), and [æ] and [a] in (11). As noted above, the GA effect (which describes pressure for identity as being inversely related to the multidimensional distance between two representations) is derivable from a specific assumption on the nature of mental representations: that they constitute sets of entailments, a point discussed in Burzio, to appear a,b.

Once this relationship is established between output-to-output faithfulness and the targeted constraints, it is no longer surprising that both can account for cases of opacity, like (1) and (11), respectively. That account is essentially unitary from the present point of view.

In sum, the facts discussed so far would provide no real argument for derivations even if McCarthy's (1999) sympathy theory was adopted alongside output-to-output faithfulness. These two extensions of OT may seem unrelated, but so do the two derivational extensions proposed by Rubach. On the other hand, as also argued in Burzio 2000b, Wilson's (2000, to appear) approach to opacity finds a direct conceptual link with output-to-output faithfulness within the perspective developed in Burzio, to appear a,b. It also benefits from the independent motivation provided by the cluster simplification asymmetry of (9), which resists derivational treatments.

4 Other Cases

The above discussion has defused all of the potential arguments for derivations in Rubach's analyses—essentially the ones that rely on opacity effects. In this section I review two further cases that Rubach invokes as arguments for derivations, but where the use of derivations seems spurious.

One such case concerns the contrast in standard Slovak between *ryasa* (← /ræ:sa/) 'cassock', discussed in the previous section, and *dialekt* 'dialect'—specifically, the contrast between the diphthong [ya] and the bisyllabic sequence [ia]. Rubach (pp. 280–282) proposes the derivational account in (13).

(13)				<i>Critical ranking</i>
	UR	/ræ:sa/	/dialekt/	
	a. Level 1 Syllabification:	ræ:.sa	di.a.lekt	NO-DIPH >> ONSET
	b. Level 2 Diphthongization:	ryæ.sa	di.a.lekt	IDENT(Nuc) >> ONSET >> NO-DIPH *[æ:] >> NO-DIPH
	c. Level 3 æ-backing:	rya.sa	di.a.lekt	See (7b)

It is clear that no derivation would be needed to account just for *di.a.lekt*. So long as the ranking NO-DIPH >> ONSET holds just as in Rubach's level 1 grammar, the input /dialekt/ will give the correct output. The derivation is only called for by the following incorrect assumption: If the

ranking NO-DIPH \gg ONSET held throughout, then level 2 *ryæ.sa* (with a diphthong, satisfying ONSET) would become *ri.æ.sa* (with a hiatus, violating ONSET), yielding surface **ri.a.sa*. On this assumption, that ranking has to be reversed at level 2 to maintain *ryæ.sa*, while *di.a.lekt* is also kept syllabically frozen by IDENT(Nuc). The assumption is incorrect because long vowels do not choose to turn into diphthongs rather than bivocalic sequences, because compelled to do so by ONSET. If they did, then, for example, Italian [fwɔ.ko] ‘fire’, diphthongized from Latin [fɔ.ku], might just as well have become *[fu.tɔ.ko] instead—an alternative satisfaction of ONSET, but an unimaginable outcome. Rather, long vowels diphthongize because of incompatibility between their prosodic length and some specific feature (often [–ATR]). Hence, the vowel with that specific feature stays short, while prosodic length is achieved by adding the glide, as shown in (14).

(14)

	fɔko	STRESS-TO-WEIGHT	*[ɔ:]	DEP	NO-DIPH
a.	fɔ.ko	*			
b.	fu.(t)ɔ.ko	*		*	
c.	fɔ:.ko		*		
d.	 fwɔ.ko			*	*

In (14) prosodic length is imposed by a STRESS-TO-WEIGHT constraint (Burzio 1994b:sec. 2.4). Diphthongization occurs under prosodic equivalence, both [ɔ:] and [wɔ] satisfying STRESS-TO-WEIGHT. The breakup of a long vowel into a bivocalic sequence would violate that equivalence—a crucial point that Rubach himself notes in his footnote 15, yet does not take into account. Thus, while Proto-Romance lengthening without diphthongization would give *[fɔ:.ko] (14c), there would be no reason to choose either variant of (14b) (*[fu.ɔ.ko] or *[fu.tɔ.ko]) instead, since both violate the lengthening imposed by STRESS-TO-WEIGHT in addition to violating some form of DEP. Either structure would automatically lose to simpler [fɔ.ko] (14a), which only violates the lengthening. Hence, the short vowel option (14a) is guaranteed to ‘harmonically bound’ the bivocalic option (14b), ONSET playing no role whatever. This is true in the Slovak case as well, as shown in (15), where length is now imposed by a faithfulness constraint, since length is underlying in Slovak. (I ignore æ-backing here.)

(15)

	ræ:sa	IDENT(Length)	*[æ:]	DEP	NO-DIPH
a.	ræ.sa	*			
b.	ri.æ.sa	*		*	
c.	ræ:sa		*		
d.	 ryæ.sa			*	*

As in (14), candidate (15b) *ri.æ.sa* from input *ræ:sa* directly loses to (15a) *ræ.sa* with a short vowel, since *ri.æ.sa* has a short vowel anyway (harmonic bounding). Hence, in addition to including a constraint banning length with certain features, the grammar of diphthongization must include a prosodic constraint requiring a long vocalic nucleus, dominating NO-DIPH. This is STRESS-TO-WEIGHT in (14) and IDENT(Length) in (15). The conclusion is that ONSET plays no role in diphthongization. Hence, there is no ranking paradox like the one in (13a–b). Hence, there is no need to rerank the constraints at different levels. Hence, there is no need for the levels.

The second case concerns the asymmetrical outcomes from the symmetrical inputs *Vu* and *uV* in standard Polish.

- (16) a. $Vu \rightarrow Vw$ (p[aw]sa ‘pause’)
 b. $uV \rightarrow uwV$ (pap[uwa]s ‘Papuan’)

Rubach proposes the derivations in (17).

- (17) *Critical ranking*
 UR /pauas/ /papas/
 a. Level 1 Syllabification: paw.sa pa.pu.as *ONSET([u]) >> ONSET
 b. Level 2 Resyllabification: — pa.pu.was IDENT(Nuc), ONSET >> *ONSET([u])

However, these derivations have no other function than stating an exception to the generalization that [w] in Polish cannot be an onset or part of an onset: namely, when [w] is adjacent to [u], as in [pa.pu.was]. Crucially, unlike in the opacity cases, this generalization is surface true. It can be expressed by means of the constraint in (18).¹³

- (18) *INDEP([w])
 The features of [w] must be dependent on an adjacent segment.

The critical cases are now simply calculated as in (19) and (20) under the given ranking, with no need for derivations.

(19)

	/papas/	ONSET	*ONSET([u])	*INDEP([w])
a.	pa.pu.as	*		
b.	pa.pwas		*	*
c.	pa.pwu.as	*	*	
d.	^u pa.pu.was		*	

¹³ The constraint in (18) needs to be dominated by others, including the one(s) responsible for vocalizing “dark” /s/, as in /skoła/ → [skowa] (Rubach 2000:292–295). No difficulty arises in this connection. Constraint (18) is also violated by the winner in (20).

(20)	/pausa/	ONSET	*ONSET([u])	*INDEP([w])
a.	pa.u.sa	*		
b.	pa.w.sa			*
c.	pa.wu.sa		*	
d.	pa.uw.sa	*		

In (19) violation of *ONSET([u]) is forced by ONSET (epenthesis must be excluded by high-ranked ONSET, just as in Rubach's analysis). *INDEP([w]) then chooses between the candidates satisfying ONSET, (19b) and (19d). In (20) realization of /u/ as [w] is optimal because it only violates bottom-ranked *INDEP([w]), *ONSET([u]) being satisfied by [w]'s parse as a coda—an option not available in (19). As Rubach notes, structures like (20c) (e.g., [xì.drà.wú.lik] 'plumber') are also allowed, however, in addition to (20b) (or [xì.dráw.lik]). Rubach's analysis of such cases in terms of input specification carries over to the present discussion straightforwardly: in [xì.drà.wú.lik] the [u] is syllabic (nuclear) in the input. The correct output is then guaranteed by the ranking IDENT(Nuc), ONSET >> *ONSET([u]).

Note that the constraint in (18) also seems to have some independent justification, as it is reminiscent of the one in (21) needed for Italian.

(21) *INDEP([−sonor])

The features of an obstruent *in coda position* must be dependent on an adjacent segment. (Latin *nocte* 'night', *optimu* 'optimal' → Italian *notte*, *ottimo*, etc.)

Within the licensing-by-cue approach proposed by Steriade (1997) and Wilson (2000), preobstruent codas are weak positions for other obstruents because they obscure perceptual cues to place and other features. This impasse is resolved by sharing features with a following onset, which is itself well cued, as reflected in (21). Similarly, one could take onsets to be weak positions for glides ([w] in particular) because of their weak closure, and it seems plausible that sharing features with an adjacent vowel would help perceptual identification. While (18) does not specifically refer to onset position, it could do so without affecting the discussion (the violation in (20b) would just be removed). Whatever the strength of this independent justification, it compares favorably with Rubach's derivational account, which will not carry over to (21).

5 Conclusion

Rubach (2000) proposes derivational analyses for a number of phonological alternations in Slavic. His main argument against nonderivational alternatives is that none of them can cover the full range of cases. That argument is undercut by the fact that his own derivational analyses do not constitute a unitary solution. It is further undercut by the fact that unity can be established within a nonderivational approach. I have divided the cases in point into three sets: those amenable to

an analysis in terms of output-to-output faithfulness; those analyzable in terms of Wilson's (2000, to appear) targeted constraints; and those that can be analyzed directly within simple OT. I have argued for a link between output-to-output faithfulness theory and targeted constraints. The link is provided by the fact that one basis for output-to-output faithfulness is similarity of representations—precisely the basis for targeted constraints (Burzio 2000b). This suggests that ‘‘attraction’’ under proximity in multidimensional space is a general property of mental representations (Burzio, to appear a,b). In addition to their unitary basis, output-to-output faithfulness and targeted constraints each enjoy independent motivation that the derivational solutions lack. Unlike derivations, output-to-output faithfulness captures the generalization that those domains that exhibit limited phonological alternations are also morphologically more regular (e.g., English *crime-less*, lacking both vowel shortening and morphological idiosyncrasy, compared with *crimIN-al*). Unlike derivations, targeted constraints relate neutralization processes to the same considerations of perceptual distance that are needed to successfully characterize the structure of segmental inventories. Unlike reintroducing derivations, comparing outputs with ‘‘other’’ representations, which both output-to-output faithfulness and targeted constraints perform, is a conceptually natural deployment of OT's parallel resources. The above discussion leads to the welcome conclusion that what seems conceptually natural is also empirically supported.

References

- Benua, Laura. 1997. Transderivational identity: Phonological relations between words. Doctoral dissertation, University of Massachusetts, Amherst.
- Borowsky, Toni. 1993. On the word level. In *Phonetics and phonology 4: Studies in Lexical Phonology*, ed. Sharon Hargus and Ellen Kaisse, 199–234. San Diego, Calif.: Academic Press.
- Burzio, Luigi. 1994a. Metrical consistency. In *Language computations*, ed. Eric S. Ristad, 93–125. Providence, R.I.: American Mathematical Society.
- Burzio, Luigi. 1994b. *Principles of English stress*. Cambridge: Cambridge University Press.
- Burzio, Luigi. 1996. Surface constraints versus underlying representation. In *Current trends in phonology: Models and methods*, ed. Jacques Durand and Bernard Laks, 123–141. Salford, England: University of Salford Publications.
- Burzio, Luigi. 1997. Strength in numbers. In *University of Maryland working papers in linguistics 5: Selected phonology papers from H-OT 97*, 27–52. Department of Linguistics, University of Maryland, College Park.
- Burzio, Luigi. 1998. Multiple correspondence. *Lingua* 103:79–109.
- Burzio, Luigi. 2000a. Cycles, non-derived-environment blocking, and correspondence. In *Optimality Theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 47–87. Oxford: Oxford University Press.
- Burzio, Luigi. 2000b. Segmental contrast meets output-to-output faithfulness. *The Linguistic Review* 17: 368–384.
- Burzio, Luigi. To appear a. Missing players: Phonology and the past-tense debate. *Lingua*. [Preliminary version in *University of Maryland working papers in linguistics 10: Proceedings of the Maryland Mayfest on Morphology 1999*, 73–112. Department of Linguistics, University of Maryland, College Park.]
- Burzio, Luigi. To appear b. Surface-to-surface morphology: When your representations turn into constraints.

- In *Many morphologies*, ed. Paul Boucher. Somerville, Mass.: Cascadilla Press. [Preliminary version: Rutgers Optimality Archive ROA-341-0999; <http://roa.rutgers.edu>.]
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Flemming, Edward. 1995. Auditory representations in phonology. Doctoral dissertation, UCLA, Los Angeles, Calif.
- Itô, Junko. 1986. Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, Amherst. [Published New York: Garland Press (1988).]
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, Mass.: Blackwell.
- Kenstowicz, Michael. 1996. Base identity and uniform exponence: Alternatives to cyclicity. In *Current trends in phonology: Models and methods*, ed. Jacques Durand and Bernard Laks, 363–393. Salford, England: University of Salford Publications.
- Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2:82–138.
- Kiparsky, Paul. 1998. Paradigm effects and opacity. Ms., Stanford University, Stanford, Calif.
- Lindblom, Björn. 1986. Phonetic universals in vowel systems. In *Experimental phonology*, ed. John J. Ohala and Jeri J. Jaeger, 13–44. New York: Academic Press.
- Lombardi, Linda. 1998. Evidence from Max feature constraints from Japanese. In *University of Maryland working papers in linguistics 7: Papers in phonology*, 41–62. Department of Linguistics, University of Maryland, College Park.
- McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16:331–399.
- Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, N.J., and University of Colorado, Boulder.
- Rubach, Jerzy. 2000. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31:271–317.
- Steriade, Donca. 1994. Positional neutralization and the expression of contrast. Ms., UCLA, Los Angeles, Calif.
- Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., UCLA, Los Angeles, Calif.
- Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison. In *Formal perspectives on Romance linguistics*, ed. Barbara Bullock, Mark Authier, and Lisa Reed, 243–270. Amsterdam: John Benjamins.
- Wilson, Colin. 2000. Targeted constraints: An approach to contextual neutralization in Optimality Theory. Doctoral dissertation, Johns Hopkins University, Baltimore, Md.
- Wilson, Colin. To appear. Positional faithfulness, licenser avoidance, and harmonic ordering. *Phonology*.
- Zec, Draga. 1995. The role of moraic structure in the distribution of segments within syllables. In *Frontiers of phonology: Atoms, structures, derivations*, ed. Jacques Durand and Francis Katamba, 149–179. New York: Longman.

Department of Cognitive Science
Johns Hopkins University
Baltimore, Maryland 21218
burzio@jhu.edu