Surface-to-Surface Morphology: when your Representations turn into Constraints[*]

Luigi Burzio
Department of Cognitive Science
Johns Hopkins University
Baltimore, MD 21218
burzio@jhu.edu

(Paper presented at the '1999 Maryland Mayfest', University of Maryland, College Park.)

Abstract
Traditional morphology, that cranks out 'underlying representations' destined for phonology, has a serial character that makes it an odd bedfellow to parallel OT. The problem becomes all the more acute once one recognizes the need for 'surface-to-surface' faithfulness constraints (Benua 1997, Burzio 1994a, Kenstowicz 1996, Steriade 1996), given the ability of the latter to do -in parallel- some of the work that the old morphology did in series. This paper proposes a theory of surface-to-surface relations that has full morphological capabilities. It does so by introducing the seemingly radical but actually natural assumption that representations are clusters of entailments that directly condition other representations. That assumption is shown to yield a faithfulness function that automatically manages distance among representations, by pressuring close neighbors to 'neutralize'. The latter function is then shown to reduce otherwise paradoxical patterns of allomorphy to an effect that is independently attested at the segmental level, where weak contrasts are eliminated. The proposal thus relates well to the 'Dispersion Theory' of Flemming (1995) and the theory of neutralization of Steriade (1994, 1997b), which account for those segmental effects. It also connects with the theory of Wilson (1999, in preparation) which in turn establishes an important link between phonological neutralization and phonological opacity (counterfeeding/ counterbleeding effects).

August 15, 1999

---

[*] Acknowledgments.

1. Introduction

In work going back almost ten years (see Burzio 1991, et. seq.) I have argued that surface forms of words, such as those in (1b), can be correctly calculated only by making reference to other surface forms, such as their respective counterparts in (1a), within a system of violable constraints that evaluates surface forms.

(1) a. phenómenon b. phenòmenólogy
américan américan-ist

While the words in (1a) are perfectly regular for stress, the ones in (1b) should rather be **phènomenólogy* (cf. *àbracadábra*) and **amerícan-ist* (cf. *antágonist*), respectively, leading to the conclusion that 'consistency' with the words in (1a) is disturbing the normal effects of the phonology. The popularity of the notion of such 'surface-to-surface' consistency or 'Output-to-Output' faithfulness (henceforth 'OO-F') has received a considerable boost by the work of Benua (1997), who developed it independently within Optimality Theory ('OT', Prince and Smolensky, 1993), as an extension of important work by McCarthy and Prince (1993, 1995) on 'reduplication', where surface-to-surface identity is also involved (see also independent work by Steriade 1996, Kenstowicz 1996). The introduction of OO-F into the theory has major reverberations for the general conception of morphology not pursued by Benua, however. One reason is the redundancy between ℜ-1 and ℜ-2 in (2).

(2)

Underlying Representation		Surface Representation
/**americ-an**/		[am**é**rican]
⇓ ℜ-1 (WFR)	⧫⧫⧫⧫⧫⧫	⇓ ℜ-2 (OO-F)
/**americ-an**-ist/		[am**é**ricanist]

In a system that combines OO-F with traditional morphology, the pattern of similarity across pairs like *américa*n/ *américa*n-*ist* has two sources. One is the word-formation rule (WFR) that yields partially overlapping underlying representations (URs); the other is the OO-F system, that enforces surface-to-surface similarity. Given the redundancy, conceptual parsimony suggests that the relation that has proven necessary --ℜ-2, should also be sufficient. What this means is that 'morphology' should be reworked as a set of surface-to-surface relations --a proposition long advocated in important work by J.Bybee (1985, 1988, 1995). Another way to set the compass in the same general direction is to consider how allomorphic variation and allophonic variation constitute similar problems, witness the similar solutions within the serial theory (UR feeding into context-sensitive rules in both cases). It would seem incongruous then, for OT to supplant the serial solution to allophonic variation, reducing it to constraint ranking with no reference to UR (Smolensky 1993, Kirchner 1997, Hayes 1999), while letting the serial solution to allomorphic variation stand, with WFRs still feeding the phonology via UR. That too should rather reduce to constraint ranking, with no reference to UR.

There are additional reasons, beside the redundancy in (2), and the allomorphy/ allophony parallelism just noted, to replace WFRs with constraints. WFRs are 'rigid'/ inviolable devices, a property shown to be incorrect by the phenomenon of morphological 'irregularity', as in *compel/ compUlS-ive, problem/ problemAT-ic,* etc. Traditional approaches require 'readjustment' devices

to fix the output of the WFR's, but those are not an independent part of the system, and thus only 'state' the problem, rather than solve it. On the other hand, in the system to be developed here, morphological irregularity comes from the fact that OO-F (now taking over the job of the old WFRs) is dominated by Input-Output faithfulness (IO-F), as in (3).

(3)

Input: /...**U**..**S**-ive/; *Base*: /compel/	IO-F	OO-F
a. compell-ive	*	
b. ☞compuls-ive		*

That is, in the proposed system, derived forms are no longer related to their bases by a common UR (ℜ-1 of (2)), but rather only by OO-F. This makes it possible to assign to derived forms their own independent input, which will have an effect only under the ranking in (3). Unlike readjustments rules, IO-F *is* an independent part of the system, and the problem of morphological irregularity is thus properly resolved, reducing to constraint ranking (see Burzio, in press, for further discussion). This account of morphological irregularity is parallel to that of phonological markedness (also a form of 'irregularity') in standard OT, as both will result from IO-F dominating the relevant source of regularity: OO-F, and Markedness constraints, respectively.

The idea to be pursued, then, is that there is, in the grammar, a general 'associative' component, which subsumes both the OO-F constraints that have been proposed in the literature, and the former word-formation machinery, in an overall architecture of the lexicon which is fully parallel, with no UR, and exhausted by the three sets of constraints in (4), the first two being, of course, just those of standard OT.

(4) a. IO-Faithfulness b. Markedness c. Associativity/ OO-Faithfulness

The paper will proceed as follows. In the next section I make two general observations on OO-F effects: (i) their strength/ rank is modulated by the independent degree of similarity between trigger and target; (ii) they come from multiple triggers, simultaneously --a fact only consistent with a 'parallel' approach. In section 3, I formulate the principle needed to account for (i), arguing that the source of OO-F is in the fact that each representation 'attracts' its neighbors, by generating entailments that they be identical to it. In section 4, I relate the proposal that morphological representations thus interact on the basis of 'distance' to recent theories of segmental neutralization. In section 5, I consider how traditional notions of morphology, such as morphemes and of word-formation-rules are rendered in the proposed system. Section 6 reviews the gains thus achieved over less structured theories of OO-F. Section 7 returns to the variable strength of OO-F and the role of independent similarity, offering more examples. Section 8 concludes.

2. Lexical Attraction and Lexical Conservatism

The task, then will be to reinterpret morphology as a set of constraints. Before turning to it, I will consider the effect described in (5), which will help us define the nature of word-to-word relations.

(5) **Gradient Attraction**

a. The overall structure of a word *w* (in both its phonological and semantic components) is influenced by that of other words in the lexicon to which *w* is independently similar, and which can be thought of as 'attractors' of *w*.

b. Attraction is stronger where independent similarity is greater.

There is a significant collection of cases that illustrate this effect. Here I will consider only two, reserving others for section 7 below. As Anderson (1992, 193) notes, there are two variants of the *-able* adjective based on the verb *compare*. One is *compárable*, which is faithful to the verb both semantically, meaning 'able to be compared', and for stress. The other is *cómparable*, unfaithful both ways, meaning approximately 'roughly equal'. One way to look at this is to consider that there is pressure from the phonology to stress the first syllable, as also manifested in *admíre/ ádmirable*. The exact nature of this pressure combines several factors. One is the long status of the last vowel of *compáre* (like that of *admíre*) and the pervasive tendency of vowels to shorten in level 1 formations. Another is the special (extrametrical) behavior of the final syllable *ble*, which makes the stresses of both *compára<ble>* (long V, penultimate stress), and *cómpara<ble>* (short V, antepenultimate stress) just regular (see Burzio 1993; 1994a for details). For present purposes, I will conflate these factors into a single constraint STRESS *cóm*. The variation is then analyzable as in (6a, b).

(6) a.

'compare-able'	OO-F good-fit	STRESS *cóm*
☞compárable	✔	*

b.

'roughly equal'	STRESS *cóm*	OO-F poor-fit
☞cómparable	✔	*

Taking the semantic difference between *compárable* and *cómparable* as given, the variation in (6) then reduces to Gradient Attraction (5). The good fit on the semantic side makes the verb *compáre* a relatively strong attractor in (6a), enabling OO-F to overcome the pressure from the phonology. In (6b), however, the attraction is weaker, allowing the phonology to prevail. The same reasoning holds if we proceed in the other direction. The semantic deviancy of *cómpar-able* is on the present theory similar to the formal deviancy of *compuls-ive* in (3) and due to some 'input', here semantic, being administered to the derived form independent of its base. Faithfulness to that input (IO-F) prevails over OO-F only if the phonology has also prevailed, revealing again that OO-F does not have a fixed rank, but one that is relativized to independent similarity --once violated, it becomes easier to violate it further. Direction of inference is in fact irrelevant here: both IO-F and stress can find a low-ranked OO-F simultaneously in (6b), while they both find a high-ranked OO-F in (6a).

The second case instantiating (5) is an instance of 'non-derived environment blocking' discussed in Łubowicz (1998).[1] She notes that, in Polish, structures that have been differentiated

[1] These are environments which are phonologically 'derived', by independent application of some phonological process, and are thus different from the purely morphologically derived ones considered in Burzio (in press), which also deals with 'non-derived environment blocking'.

from their bases by palatalization (k ⇒č; g ⇒ǰ) will additionally undergo spirantization when the affricate is voiced (ǰ ⇒ž), e.g. *dron[g]* 'pole'/ *dron[ž]-ek* 'little pole'. In contrast, items in which the voiced affricate is already present in the base, do not undergo spirantization (ǰ ⇒*ž), e.g. *bry[ǰ]* 'bridge'/ *bry[ǰ] -ek* 'little bridge'. This situation is analyzable as in (7), parallel to (6).

(7) a.

bry[ǰ] -ek	OO-F good-fit	SPIRAN (ǰ ⇒ž)
☞bry[ǰ] -ek	✔	*

b.

dron[ǰ]-ek	SPIRAN (ǰ ⇒ž)	OO-F poor-fit
☞ dron[ž]-ek	✔	*

In (7b), the upper left-hand corner gives the form resulting from palatalization, which applies in any event, as shown by the unvoiced velar: *kro[k]*/ *kro[č]-ek* 'step/ little step'. Because of this, OO-F applies here under 'poor fit' conditions, since the base is *dron[g]*, not *dron[ǰ]*, and here spirantization prevails. In (7a), however, OO-F finds a good fit, since the palatal is already present in the base *bry[ǰ]*, and thus blocks spirantization.

The 'stress preservation' effects in (1) are of course consistent with (5). The derived forms are similar to their bases segmentally and semantically, and this predicts some influence on the stress pattern. However, because (5) only refers to independent similarity, it does not limit such influences to what in traditional morphology would be the 'base'. Rather, it predicts that OO-F effects should be more general. While we will see below that the traditional 'base' does indeed have a special status, the greater generality of (5) is supported by the 'Lexical Conservatism' effects of Steriade (1997b, 1998) and the 'Multiple Correspondence' effects of Burzio (1998a), which in fact describe the same phenomenon. Steriade notes that the adjective *remédi-able* is uncharacteristically unfaithful to its verbal base *rémedy*, a fact that seems attributable in part to the phonology's dislike for pre-antepenultimate stress (**ré.me.di.a.<ble>*, in the analysis of Burzio, 1994a). However, as she argues, the presence of the adjective *remédi-al* seems also to play a role, given that *párody* does not equally give ??*paródia-ble*, in the absence of **paródial*. The tableau in (8) illustrates this situation.

(8)

remedy-able	OO-F1 *rémedy*	*preAP stress	OO-F2 *remédial*
☞ remédiable	*	✔	✔

What we are seeing is that what we may call a 'secondary' base or attractor, imposing its own OO-F2 in (8), is able to join forces with the phonology to overcome the influence of the primary base that imposes OO-F1. In Steriade's terms, unfaithfulness to a primary base is 'lexically conservative', in the sense that it complies with some independently existing form: what I have called the secondary base. For the actual form to be optimal in (8), the two individually lower-ranked constraints must therefore be top-ranked as a conjunction. I will return to such 'conjunctive' effects shortly.

Steriade gives other examples, including one characterizable by the tableau in (9) parallel to (8).

(9)

nouveau an	OO-F1 *nouveau*	ONSET	OO-F2 *nouvelle*
☞nouvel an	*	✔	✔

French masculine adjectives like *nouveau* 'new' borrow heavily from their feminine counterparts when occurring in certain prevocalic positions (the 'liaison' contexts), here as in *nouvel an* 'new year (MASC.)', where *nouvel* is homophonous with *nouvelle* 'new (FEM.)'. This follows from combining the phonology, that requires an onset, with 'Lexical Conservatism', which privileges independently existing forms. The feminine is the independent form that has the needed consonant. In (9), we can thus see the citation masculine form *nouveau* as the primary base, and the citation feminine form *nouvelle* as the secondary base, matters then working just as in (8).

In Burzio (1996, 1998a) I similarly note the Italian cases in (10), (11), respectively.

(10)

comi[k]-itá	OO-F1 *comi[k]o*	**[ki]* (palatn.)	OO-F2 *comi[č]i*
☞ comi[č]itá	*	✔	✔

(11)

vinto-óre	OO-F1 *vínto*	STRESS STEM	OO-F2 *víncere*
☞vìncitóre	*	✔	✔

The case in (10) relates to the fact that the pattern of palatalization in cases like *comi[k]o/ comi[č]-ità* 'comical/ comicalness' is generally unpredictable, compare *anti[k]o/ anti[k]-ità* 'ancient/ antiquity', and yet is quite consistent with the outcome in the plural form of the adjective: *comi[č]i*, versus *anti[k]i*. Assuming that the singular form of the adjective is the primary base for the abstract noun, the plural form of the adjective must nonetheless be able to contribute its crucial influence, similarly to the other 'secondary' bases above. In the case in (11), we see that the agentive noun *vincit-ore* 'winner' exhibits both a *t* from the participle *vinto* --the 'primary' base for such formations, and a *c* from the infinitive *vincere*, evidently acting as a secondary base, thus revealing the 'Multiple Correspondence' (=Lexical Conservatism) at work in the system. The independent incentive for straying from the segmentism of the primary base is encapsulated in (11) as STRESS STEM, in fact a form of metrical faithfulness (OO-F), to both the participle *vín-to* and the infinitive *vínc-ere*, both stem-stressed. Given the inherent stress of the affix *-óre* and the general prohibition on consecutive stresses, only insertion of a pre-affixal syllable will permit satisfaction of STRESS STEM. It is then the nature of the inserted syllable *ci* that reveals the influence of the infinitive. The latter lends the consonant, while the *i* is from participial allomorph *-it-* (another 'Multiple Correspondence' effect. See Burzio 1998a).

The characterization in (5) thus recognizes both an important modulation effect on the rank of OO-F constraints, and the multiplicity of relations that seem relevant to their application. Given that multiplicity, WFRs are thus not only redundant with OO-F constraints in the general sense of (2), but would also be insufficient do define the domain of application of OO-F constraints, since there is no WFR relating *remedial* to *remediable*, etc. This multiplicity of relations also reveals the inadequacy of the 'cycle' as an account of OO-F effects (like those in

(1) above), since obviously no cyclic derivation could involve multiple bases (see also Burzio 1994a, 1998a, in press). Even further argument against WFRs as the basis for of OO-F constraints, as well as against cyclicity, is given by the fact that attraction/ OO-F effects hold not only over stems but also over affixal material (consistently with (5), that draws no distinction). For instance, as discussed in Burzio (1994a, b), *titán-ic* is metrically inconsistent with *títan*, and this is due to the metrical consistency of the affix *-ic*, shown for example by the stress-wise identity of *titánic*/ *barbáric*, despite the lack of any WFR or cycle relating such pairs.

There are still some respects in which taking the 'base' of traditional WFRs as the trigger of OO-F, as in Benua (1997), may seem desirable, however. One is that OO-F effects from that 'base' are undoubtedly more prominent and and systematic than from any 'secondary' base. Another is that 'backward' transfer of properties from derived forms to bases is rather rare, and could be argued not to exist. I will address these issues below.

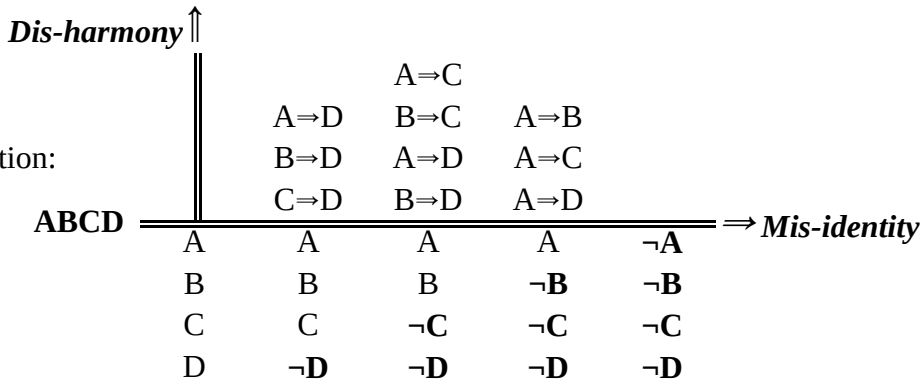
3. Hebbian Learning?

An account of 'Gradient Attraction' (5) can be attained by means of the simple assumption in (12), which also has a wide range of further theoretical consequences (Burzio 1998b).

(12) **Representational Entailments** (RE)
Mental representations of linguistic expressions constitute sets of entailments --a representation with the structure AB generating the entailments A ⇒B, B ⇒A.

On this assumption, representations which are similar but not identical will violate some of each other's entailments, giving precisely an invariance or 'attraction' effect. The illustration in (13) reveals the general pattern of disharmony thus produced by mis-identity of two representations.

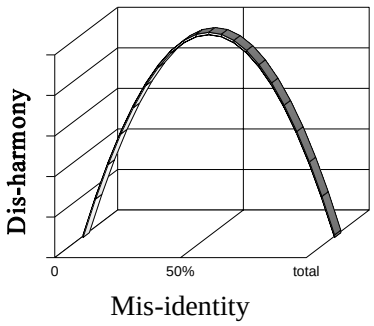
(13)



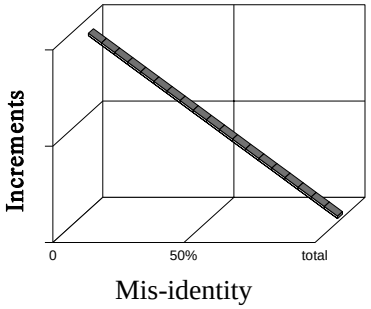
The diagram in (13) considers a four-component representation ABCD and tabulates entailment violations on the part of another four-component representation as the similarity between the two changes from total identity to total misidentity (left-to-right). Disharmony is measured by the total number of entailments violated --effectively, the height of the column above the horizontal axis. At the leftmost point, the representation is ABCD, identical to the target, so all entailments generated by the target are satisfied. As D is changed to ¬D at the second leftmost point, however, three entailments are violated, as indicated. In general, violated entailments are tallied

by taking each of the still identical components as an antecedent and combining those with each of the changed components as a consequent. The other points in the tabulation will be straightforward. Hence, as misidentity increases linearly, disharmony varies non-linearly as in (13), and as more clearly visible in (14a), which simply works out the pattern over a larger number of points, namely over representations that have more components in them. The curve in (13)-(14a) is one of linearly decreasing increments, as shown by (14b), which plots the first derivative of (14a) (pattern of increments).

(14) a.



b.



As shown in (14a), disharmony peaks at 50% misidentity, dropping back to zero at total misidentity. As shown by both (14a) and (14b), increments in disharmony are initially high, dropping to zero at 50% misidentity, and going negative beyond that point.

This accounts for Gradient Attraction (5) in the following way. Suppose the characterization in (5) is in fact based on cases of 'allomorphy', namely representations which are relatively similar in both sound and meaning. In terms of (14), such representations would presumably be more than 50% similar, thus implicating only the left-half of (14a, b). Over that half, a one-component mis-identity will be penalized maximally (maximal increments in disharmony) when independent identity is high, near the left edge, but minimally when independent identity is low, near the peak of (14a). Put differently, the 'attraction' (pressure for identity on a specific dimension) is greatest when the two representations are closest (already identical on other dimensions). The *compárable/ cómparable* contrast of (6) can now receive the analysis in (15), extendable to the other instantiation of Gradient Attraction in (7), and those of section 7 below.

(15)

	compáre	a. **compár**-able	b. **cómpar**-able
segmental structure:	A	A	A
	B	B	B
	C	C	C
meaning: 'compare'	D	D	¬D
stress: *pá*	E	☜¬E****	☞¬E**

For the sake of illustration, (15) takes the structure of the verb *compare* to be a five-component

vector, where the first three components, A, B, C are stand-ins for the segmental structure; the fourth, D, for its semantics, and the fifth, E, for its stress pattern. In considering the adjective, I put aside the contribution of the affix *-able*, for the moment, assuming it simply 'adds' further components (both semantic and segmental) to the vector, and limit discussion to the stem, which is the structure being calculated. Then, an adjective *comparable* will share the same segments of the verb: components A, B, C, and, if it means 'able to be *compared*', it will also share the semantics of the verb, namely D, as in (15a). The question is whether it will also share the stress of the verb. Remember that the phonology provides pressure to stress the first syllable, as in *cómparable* (= ¬E). That pressure is evidently stemmed here by the potential four-entailment violation, represented by the four stars in (15a). Rather than incurring that violation, stress is here faithful to that of the verb (= E). Matters are different, however, if the adjective has a different semantics (= ¬D), as in (15b). Now only a two-entailment violation stands in the way of re-stressing, thus accounting for why here it occurs in compliance with phonology's demands. The count of two comes from the fact that, in (15b), E is entailed only by A, B, C (3 entailments), while at the same time a choice of 'E' over '¬E' would itself give rise to a one-entailment violation, of 'E ⇒D' (because of the presence of ¬D). Hence the total disharmony difference (between 'E' and '¬E') in (15b) stands at only two entailments. This account will be partially revised below to take account of the affix, but the general reasoning given for (15) will stand.

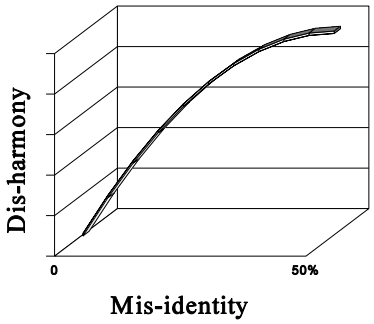
The 'Representational Entailments' assumption (12) thus provides a general framework for understanding word-to-word relations, by postulating a conversion (automatic, inherent in the character of mental representations) between representations and constraints (the entailments being just a type of constraint operative in the system). Words influence each other because they generate constraints which other words are then subject to. On the present conception, calculation of an individual word's structure presupposes taking the rest of the lexicon as a given. OO-F constraints do therefore not emanate from a unique 'base', but are rather provided jointly by the lexicon at large. The largest contributions will be provided by words that are independently most similar to the one being calculated but also, as we see below, by the one that would be the 'base' in traditional terms.

The above appeal to 'harmony', measurable inversely from the number of entailments violated, brings up of course a certain theoretical point. Most work in OT follows Prince and Smolensky (1993) in taking constraints to be 'strictly-ranked', in the sense that lower-ranked constraints cannot -by joining forces - prevail over higher-ranked ones. We have seen, however, that in the cases (8)-(11) above this is not true, as some phonological constraint did join forces with OO-F2 to prevail over OO-F1. A good number of others such conjunctive effects are known. Smolensky (1997) proposes in this connection that constraints can, under special circumstances, indeed form 'conjunctions' of higher rank than each of the conjuncts. In the present work, I will proceed from the opposite direction, however, and assume that constraints are inherently summatory, i.e. that they have numerical weights, as in Burzio (1994a), and in fact in Smolensky's (1986) 'Harmony Theory', leaving 'strict ranking' effects, rather than their absence, to be accounted for. The Representational-Entailments assumption in (12) has in fact the potential for deriving strict ranking effects. Since, under (12), constraints are non-distinct from the representations that satisfy them, a constraint that, for any reason, tends to prevail, will be relatively more successful in recruiting representations through the lexicon, which will in turn

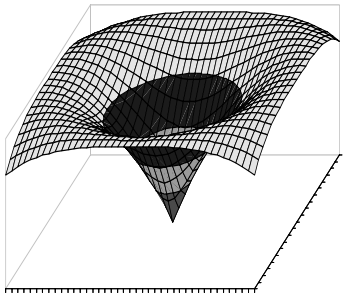
reinforce that constraint by means of their representational entailments. This 'strengthening-of-the-strong' effect will work to create substantial differences in weight among conflicting constraints (in fact proportional to the domain of conflict), thus approximating the effect of strict ranking.[2]

The effects of RE (12) on representations that are independently similar can thus be compared with those of a slope shaped like the left half of (14a), given in (16a), or -to see it three dimensionally- those of a hole, as in (16b).

(16) a.



b.



A representation being calculated, like *compar* in *compar-able* can be thought of as an object being pushed by gravity to the bottom of the slope of (16a)/ center of the hole of (16b), a point which is in fact another representation, in this case, *compáre* (the 'attractor'). The closer the candidate representation gets to it, the steeper the slope, and thus the stronger the push. Two other forces can deflect this attraction, and thus give rise to 'allomorphy'. These come from the other two components of the system in (4): IO-F and Markedness. The opportunity for them to have an effect will be greater on the flatter than on the steeper part of the slope, however (greater distance from the attractor), as we saw was the case in (6)-(7). We can think of IO-F as a kind of friction, that works to hold the object in its initial position (the 'input'), thus resisting the downward push, as in *compuls-ive* of (3) above, which successfully resists the pull from *compel*. As for markedness constraints, we can think of them as a set of springs pulling the representation, on each relevant dimension, towards a favorite point, as in *divin-ity*, which resists the pull from *divi:ne* due to a markedness constraint that favors short vowels (Burzio, in press).

In sum, the assumption in (12), radical from the point of view of generative work, but otherwise rather simple and natural, provides the basis for word-to-word interaction which can characterize OO-F effects including Gradient Attraction (5). It can also reconstruct basic concepts of traditional morphological theory to the extent that they are useful, as we see below. It is worth considering as well that RE (12) is a potential link between generative work, traditionally focused on turning empirical generalizations into elements of 'macro'-architecture,

---

[2] Another respect in which the present view may appear unorthodox is that it does not maintain that all constraints are universal. However, the class of constraints that can be thought to be universal remains in fact unaffected by the present approach. It is only the new class of constraints, that come from representations, and that supplant other resources like word-formation rules, that cannot be universal since representations are not.

and connectionism -in turn linked to neuroscience- with its alternative goals of mapping elements of 'micro'-architecture (weights and connections) into behavior. The reason is that RE (12) is a virtual re-statement of 'Hebb's rule of mental learning' in (17).

(17) When an axon of a cell A is near enough to excite cell B or repeatedly or persistently takes part in firing it, some growth or metabolic change takes place in both cells such that A's efficiency, as one of the cells firing B, is increased (Hebb 1949, 62)

The co-firing of cells A and B is taken in (17) to trigger a change that makes that co-firing necessary, just as co-occurrence of A and B in (12) is interpreted as being necessary.

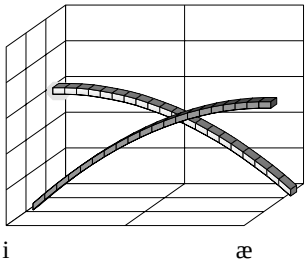
4. The Theoretical Scope of Neutralization

Before turning to implementing (12) in a system of combinatorial morphology, I will briefly consider certain relations of (12) to other contemporary work.

The harmony function produced by RE (12) and depicted in (16a, b) can be thought of as a general 'neutralization' function, characterizing the tendency of weak contrasts to be eliminated. At one particular level, the function in (16) can characterize the notion of 'morpheme', namely the fact that expressions that do not contrast in meaning, also tend not to contrast in their sound structure. Since that correlation is, in the present system, both violable and imposed on outputs, at the same time as it characterizes the notion of morpheme, the function in (16) also characterizes the notion of 'anti-allomorphy'/ paradigm uniformity --the OO-F of most recent work. At a finer-grained level of representation, that of segments, the function in (16) can capture the notion of 'minimal distance' among members of segmental inventories postulated in the 'Dispersion Theory' (DT) of Flemming (1995), and will relate in turn to the theory of 'positional' neutralization of Steriade (1994, 1997b), based on the distribution of perceptual cues. As Steriade argues, positions affected by neutralization, e.g. of laryngeal or place contrasts for consonants, are those in which perceptual cues for that contrast are weak, as in pre-obstruent position. What that means in terms of (16) is that, for perception-based representations, each member of the weakly contrasting pair will be strongly attracted to the other.

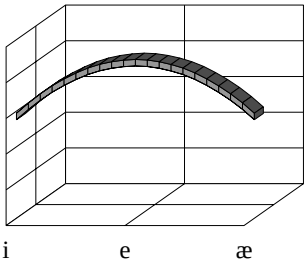
So, for representations which are intended not to contrast, like allomorphs of the same morpheme, the curve in (16a) is one of *dis*-harmony: difference is bad. But for representations which are intended to contrast, like different phonemes, the same curve is one of harmony: difference is good, because it lessens attraction, i.e. the possibility of neutralization. The even-spacing of phonemes in perceptual space will in fact follow from (16a) as a harmony function. Consider two points in perceptual space corresponding to front vowels *[i]* and *[æ]*. What we want ifs for a third front vowel, *[e]* to turn up exactly centered between those two points, rather than bunched up with either *[i]* or *[æ]*. Now, the two harmony functions generated by *[i]* and *[æ]* over the same front axis will be as in (18a). Then, the best third vowel on that axis will simply be where total harmony (summation of the two curves), given in (18b), is maximal -- indeed in the middle.

(18) a.



Harmony of contrast with each of *[i]* and *[æ]*.

b.



Overall harmony for points between *[i]* and *[æ]*.

If the two functions in (18a) were linear, summation would give a horizontal line in (18b), with no way to identify the middle. Hence, indeed, Gradient Attrtaction (5) and 'even spacing' within segmental inventories appear to be parallel phenomena.

The Flemming/ Steriade theory of contrast that the present work ties into, thus introduces a notion of 'markedness' which is relative to the existence of other neighboring representations, alongside of the standard OT markedness, which is absolute. Both types of markedness are needed: when contrasts are neutralized, absolute markedness determines how. So, English unstressed vowels neutralize to schwa: the least marked vowel, involving minimal articulatory activity. Roughly speaking, the absolute markedness is grounded in production/ expression, while the relative one, imposing contrast, is grounded in perception/ interpretation. The above notion of OO-F as gradient attraction then converges with the relative markedness of Dispersion Theory: unfaithfulness under conditions of close similarity is highly unstable, and hence 'marked'. The present work extends the DT conception from segmental inventories to lexicons, via the RE assumption in (12), applicable to representations of all sizes.

C. Wilson (1999, in preparation) has discovered further that introducing such relative markedness into OT (which he argues for in both phonology and syntax) yields 'opacity' or 'derivational' effects, like counterbleeding/ conterfeeding ordering of rules. Intuitively, the reason is that, once relative markedness is introduced, the input-output mapping is no longer a simple interpolation between the input and a least marked structure. Rather, it is a more complex calculation in which that interpolation is combined with the effects of the nearest attractor, which will deflect the output, either towards itself: neutralization; or away from itself: enhancement of contrast, depending on specific constraint ranking. Although Wilson provides (and defends) a fully parallel characterization of this effect, the pattern of input-output dependency is 'multi-stage', whence the illusion of derivation. In general, derivational/ opacity effects would thus be produced by the presence in space of a 'deflector', so that , given the input and general markedness, one first needs to know whether one is within the range of a deflector (relative markedness) before knowing where to end up. For the class of 'deflectors' that are words, their ability to produce opacity effects in a parallel system was recognized as early as Burzio (1994a, 190f, fn. 16; 1992 manuscript). There, I suggested -terminology aside- that calculation of *wr[ʌɪD]ing*, with 'raised' *[ʌɪ]*, in the relevant Canadian dialects, is affected by the presence of *wr[ʌɪ]te* (OO-F), where raising is normal, induced by voiceless *[t]*. McCarthy (1998) has shown, however, that not all cases of phonological opacity are amenable to such a (word-based) OO-F account. His 'Sympathy Theory' provides a way to characterize a larger class of attractors (the

class of ‘flower’ candidates) and thus a larger class of opacity effects. While Sympathy Theory lends a sharp insight into the dynamics of optimization, the gain in expressive power it brings comes withouth benefis in explanation, in that ‘deflectors’ (flower candidates) can be constructed ad-hoc. The significance of Wilson’s finding lies in the potential for an independent characterization of the class of deflectors. This would come from a general theory of neutralization based on Steriade’s (1995, 1997b) --a result that would raise the explanatory level of McCarthy’s general conception considerably and tie together two important lines of inquiry that might have seemed completely unrelated.

In hindsight, then, opacity could be said to be indeed due to OO-F as suggested in Burzio (1994a) after all, except that the class of ‘deflectors’ is not limited to words, but is rather the class of neighboring representations more generally --morphemes, segments, or other, as now expected under RE (12). For related -and more extensive- discussion of these issues, see Wilson (1999, in preparation). For a different approach to opacity in OT, see Goldrick and Smolensky (1999).

5. Morphology and Entailment Structures

The first step in utilizing the RE assumption in (12) and its effects diagramed in (14) in a general theory of morphology is to recognize the need for language to be combinatorial. From the present perspective, combinations of morphemes are inherently disharmonic. The reason is that RE (12) penalizes variation in general, thus also when, for instance, an affix combines with multiple stems. Hence, it has to be the case that failure to combine, as in a lexicon that only had monomorphemic words, would be even more disharmonic. Indeed, this is due to the fact that the latter situation would produce a worse overpopulation of the relevant space, with higher degrees of attraction between neighbors. The same is true of phonemes for that matter, and of the fact that words are not in general ‘monophonemic’. Phonemic inventories are limited by the need to maintain sufficient contrast as noted above. A lexicon of monophonemic words would thus be correspondingly limited, disharmony growing very rapidly with its size. Combining phonemes into sequences increases the size of the space by adding dimensions to it, and thus relieves overpopulation/ disharmony. Disharmony is only reduced this way, though, never eliminated, since combinations are themselves disharmonic as noted, so that segmental sequences like *ta, ba* will violate some of each other’s entailments. Transposing everything to morphemes leads to the conclusion that their combinations are also disharmonic, but less so than the alternative, which is a monomorphemic lexicon.

Consider then a typical combination, as in (19), where suffix *-al* combines with different stems.

(19)

	Lexicon:	Entailments: I.	Entailments: II.
a.	parental	*al* ⇒/*parent*___	*al* ⇒/N ___
b.	natural	*al* ⇒/*nature*___	*al* ⇒/N ___
c.	...	...	...
d.	Entailments summation over the lexicon: $\sum_{i=1}^{n}$	??	>2
e.	Macro entailment (constraint):	___	*al* ⇒/N ___

Calculation of a word featuring affix *-al* (like *nation-al*), will face two types of entailments from the rest of the lexicon, given in columns I. and II. in (19). The entailments in I. refer to the specific structure of individual words. So, the word *parent-al* will generate, among others, a set of entailments that can be abbreviated as 'if there is a structure *-al*, there must be a structure *parent* preceding it', as in (Ia). Similarly for *natur-al* in (Ib), etc. At the same time, both *parent-al* and *natur-al* will also generate a set of entailments to the effect: 'if there is a structure *-al*, there must be a noun preceding it', as in column II. This second set of entailments is a subset of the first, in that the semantic property 'noun' is part of the representation of both *parent-* and *natur-*. More accurately, a stem like *parént-* of *parént-al* entails the property 'noun', by virtue of its close association (shared entailments) with the word *párent*, whose semantic property it is to be a noun. With large enough classes of stems, the entailments from column I. will reduce to noise, as they largely contradict each other --the disharmony alluded to earlier. As such, they will have little or no effect on the calculation of any individual word. The entailments in column II., however, are identical, yielding a significant summation effect over the lexicon as indicated in (d). The result of the summation is referred to in (e) as a 'Macro entailment' and is simply a constraint like more familiar ones, but expressing a morphological, rather than phonological, generalization. The latter is in fact perfectly similar to a subcategorization frame for the affix or to a word-formation rule (WFR), except that: (i) it is violable; and (ii) it can apply in parallel with the phonology, both desirable qualities as it turns out.

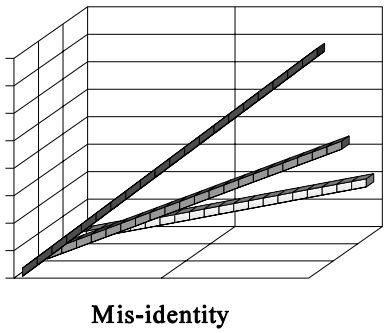
Morphemes are thus structures that tend to be internally invariant, associating the same sound with the same meaning --a foreseeable consequence of RE (12). Some of them combine with others with some productivity. Such combinations are disharmonic under RE because they produce partially overlapping representations: e.g., same affix/ different stems. Weaker forms of stem invariance are characteristically still achieved however, as in (19), with *-al* combining only/ predominantly with nominal stems. Such weaker form of invariance is still attributable to RE. The latter, however, makes no actual prediction as to which features or components of a representation will be selected to remain invariant in a stem. But that is true of WFRs as well, which can in principle be formulated with reference to any aspect of the representation. The fact that lexical category is characteristically selected for invariance simply suggests that the latter plays a major role in the semantic representation.

Given macro-entailments like the one in (19e), presence of an affix will then be a factor in determining the harmony of its stem. This is in addition to the harmony due to similarity between that stem and some other representation ((14), (16) above). So, continuing to abstract

away from possible scalar differences among the dimensions, we can take each increment in misidentity between the stem and a member of the targeted category (e.g. nouns) to give a corresponding increment in disharmony-as-violation of (19e), yielding an affix-based harmony function which is essentially linear. This characterization is an approximation, since in fact each violation of stem-base identity would simultaneously also reduce the force of the macro entailment in (19e), which only reflects the overall truth of that stem-base identity over the lexicon. However, for a sufficiently large class of affixed stems, we can take the contribution of any individual word to be insignificant, and the misidentity/ disharmony function generated by the affix to be indeed linear, as in (20), where the different lines represent effects of different strengths, as generated by different affixes.

(20)

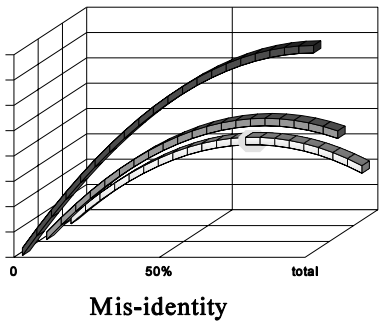
Affix-induced disharmony on unfaithful stems



The effect of stem-base misidentity in (20) must now be compounded with the one in (14), (16), namely the attraction (positive above 50% identity, negative below 50% identity) that any representation has on any other, given by the upside-down-*U* curve in (14a). Such compounding, i.e. summation, gives, for the three lines in (20), the three curves in (21).

(21)

Overall disharmony for unfaithful affixed stems



This means that, in effect, affixes are ‘attraction enhancers’. By imposing that their stem be a member of a certain class, they ‘direct’ their stems towards that class, thus enhancing the attraction effect that any member of that class exerts on the stem under calculation, by virtue of their shared entailments. Recall here that the account of Gradient Attraction for (15) above had abstracted away from the effects of the affix, which now turn out to be significant. However, what was crucial for that discussion was only that the disharmony curve (of (14), (16)) be one of decreasing increments, which is still true of (21). Hence the earlier conclusions will stand.

We may again visualize the effects in (21) three-dimensionally by turning the curves into 'holes' as in (14) above, but in fact the hole is here iterated through the whole relevant class, e.g. nouns. This iteration will produce something like the 'valleys' in (22).

(22)

a. Attraction for an affixed stem: strong b. Attraction for an affixed stem: weak

In these illustrations, the centerlines stand for the class of items of the category selected by the affix, e.g. nouns, which exert their attraction on stem representations. The latter are in this analogy objects on the sloping surfaces. The attractors form a line in the sense that they share the same value: noun, on a certain semantic coordinate. They are taken to be contiguous as in (22) only for the sake of graphic ease, however, the more appropriate figure being presumably a set of intersecting holes, lined up but at random distances from one-another. This approximation will do for our purposes. The two figures in (22), then, like the three curves in (21), represent affixes with stem-selecting properties of different strengths/ ranks. Note that the upside-down-*U* curve in (14a), and hence also the diagrams in (21)-(22) which are based on it, exhaust the representational space, since they cover all possibilities, from total identity, to total misidentity to a target representation (a point on the centerline in (22)). Now one can imagine that, for an affix whose effects give sufficiently steep surfaces like those in (22a), there may be no representation, i.e. point on those surfaces, where an object could fail to slide onto the centerline, thus becoming identical with one of the items on that centerline (e.g. an existing noun in the lexicon). As I argue in Burzio (1999c), that situation is the one instantiated by 'regular' morphological systems, like that of English 'level 2' affixes. For example, in contrast to level 1 'irregulars' like *arbore-al* (bound stem) or *symptomAT-ic*, level 2 affixes can be taken to 'neutralize' such potential stems, that would be somewhere on the slopes in (22a) to their nearest centerline neighbors (i.e. existing words), thus yielding, correspondingly, *tree-less*, not **arbore-less*, *symptom-less*, not **symptomAT-less*. Irregular morphology, like that of English level 1 affixes, on the other hand, can be associated with a weaker attraction system like that of (22b), where representations can more easily/ frequently 'hang on' to the slopes, being kept away from the centerline by IO-F (discussion of (3) above), interpretable in this model as friction. The other force that can oppose the attraction/ downward push in (22) is of course phonological markedness, which will also be expected to succeed more in (b) than in (a), as in level 1 shortening *crimin-al*, or re-stressing *parént-al*, compared with level 2 *cri:me-less*, *párent-less*, both faithful to their stems. As argued in Burzio (1999c), this thus accounts for the general fact that morphological regularity correlates inversely with phonological regularity, sectors of the derived lexicon that are morphologically

regular containing phonologically irregular/ marked structures (like *párentless*), while those that are morphologically irregular (*arboreal*, etc.) are rich in regular phonological alternations (shortening, re-stressing, etc.).

The RE assumption in (12) is thus the key to understanding OO-F in a way that can fully take over morphological relations, thus eliminating the redundancy of (2) above. On the latter assumption, morphological relations are RE-induced forms of invariance that obtain across representations. For a large enough lexicon, words will end up overlapping in their representations, resulting in disharmony under RE. Morphological organization can be seen as an effective way to manage representational overlaps, substantially limiting them to combinations of units --the morphemes, thus satisfying two forms of invariance: (i) internal invariance of the morphemes; (ii) invariance of their combinations by some major categorial feature. Both forms of invariance are violable in the present system. General allomorphy violates (i) by the upside-down-*U* function in (14a), while allomorphy specific to a selected stem violates both (i) and (ii), in a pattern of disharmony given by the lopsided curves in (21), whose steepness depends on the 'strength' of the selector, i.e. the affix.

6. Counting Assets

The present approach is thus consistent with that of Benua (1997), Burzio (1994a) in postulating high-ranked OO-F for English level 2 affixes and low-ranked OO-F for level 1 affixes, but is more principled on a number of fronts, yielding the advantages in (23).

(23)
- I. It provides an account of 'base-identity', without limiting OO-F effects to a unique base.
- II. It accounts for the fact that OO-F constraints pertaining to different dimensions cluster in rank (all low-ranked for level 1; all high-ranked for level 2 affixes).
- III. It accounts for Gradient Attraction (5).
- IV. It attributes the 'rank' of OO-F constraints to entailment summation.
- V. It does away with an independent morphological module.

Beginning with I., the present system identifies the 'primary' base, just like a system based on traditional word-formation rules, since it has the constraint analog of WFRs, in the form of macro-entailments like the one in (19e) above. Given a word that includes an affix, like *parent-al*, the primary base is the attractor for which the macro-entailment *al* $\Rightarrow /$ *N*__ holds, here *parent*. In the present system, the primary base in this sense is the source of the highest-ranked OO-F constraints, precisely because the affix thus identifies it and enhances its attraction effect as argued. Other, 'secondary' bases, like *remedi-al* for *remedi-able*, will still act as attractors, but not under comparable enhancement. The reason for the difference is essentially that, while pairs like *remedy/ remedi-able* (primary base) reflect a systematic distributional fact, pairs like *remedi-al/ remedi-able* do not, but are rather rare and accidental. Cross-lexical summation will thus yield a strong macro-entailment imposing a verb before *-able*, but virtually no macro-entailment imposing *X* before *-able*, where *X* also appears before *-al*. Hence the effect of secondary base *remedi-al* on *remedi-able* is only the 'regular', unenhanced attraction of (14a) rather than the one

in (21)-(22). The observed asymmetry between primary and secondary bases thus follows this way. The important asymmetry between base and derivatives will also follow, as we see next.

Given the fact that affixes enhance attraction by 'directing' their stems towards a certain lexical class, one may wonder whether the opposite could also be true. That is, given that certain nouns have derivatives in *-al*, could the lexical class of *parent* correspondingly 'direct' this item to *parent-al*? It seems reasonable to suppose that, if this effect exists at all, it will, in general, be weaker. The reason is that the existence of a noun for an adjective in *-al* is a relatively strong generalization (despite the class of *ic-al* items discussed below), while the existence of a corresponding adjective in *-al* given any noun is a relatively weak one, since many nouns lack such adjectives: **student-al*. This will then constitute an asymmetry in attraction enhancement. More importantly and clearly, however, there is an independent asymmetry in the attraction. Consider that a complex form like *parent-al* can potentially satisfy all of the entailments of the simple form *parent,* essentially because it contains it. In general, a structure ABC will satisfy all the entailments of a substructure like AB (A ⇒B; B ⇒A). In contrast, a simple form like *parent* cannot satisfy all of the entailments generated by *parent-al*. In particular, it will violate the cluster of entailments that abbreviate as *'parent ⇒/ __ -al'* (if there is a structure *'parent'*, there must be a structure *'-al'* following it). Similarly, AB will violate ABC's entailments A ⇒C; B ⇒C. This means that the word *parent* is a strong attractor for the stem in *parent-al*, while the latter is only a weak attractor for the former. In other words, the attraction 'hole' generated by *parent* has the stem *parent-* inherently near its center, while the one generated by *parent-al* has the word *parent* away from the center. This then accounts for the fact that 'backward' transfer of phonological properties is rarely attested. For example, transfer of stress from *napóleon* gives only mildly deviant *napòleón-ic*, with secondary stress on the second syllable similarly to *phenómenon/ phenòmenólogy* of (1) above (regular stress: *àbracadábra, wìnnepesságukee,* etc.). In contrast, there is no transfer from *parént-al* to **parént*, which would be no more than mildly deviant, given attested *cemént*. However, the above considerations also make sense of the two apparent cases of back-copying in English stress identified in Burzio (1994a, b). One is the case of adjectives in *-ic*, e.g. *académic* whose stress is irregular, but in synch with the regular one of their *ic-al* counterparts, like *académic-al*. The other is that of uninflected verbs, e.g. *pervért*, whose stress is again irregular, compare the noun *pérvert*, but in synch with that of the gerund, e.g. *pervért-ing*, which is regular. First, in both cases, the distributional asymmetry that yields asymmetrical enhancement is lacking. So, the majority of adjectives in *-ic* do have variants in *-ic-al*, and all the verbs have inflected *-ing* forms. Second, in both cases, the addition of the affix does not contribute a change in lexical category. This places the base form relatively closer to the derived one when the latter is the attractor. Concretely, a word like *read-able* will generate the entailment: 'if there is a structure *'read-'*, the word is an adjective' (part of the entailment: *read ⇒/ __ -able*), which is then violated by the base *read*, which is a verb. In contrast, the corresponding entailment from *read-ing*: 'if there is a structure *'read-'* the word is a verb', is satisfied by the base form *read*. Hence *read* will be closer to *read-ing* than it will be to *read-able* when each of the latter is an attractor. Similarly for *academic/ academic-al*, both adjectives. If lexical category plays a major role in the representation as already suggested above, then it should make a significant difference in controlling 'backward' transfer possibilities. The two cases in question seem to confirm that.

In sum, in general, bases influence derivatives rather than the other way round because derivatives inherently involve entailments that a base will violate, while being inherently capable of satisfying all of the base's entailments. In addition, the attraction enhancements are asymmetrical, in that affixal material entails the category of its stem, in a way in which a stem does not in general entail attachability by an affix. Both cases of backward transfer in English word stress of Burzio (1994a, b) involve special circumstances that temper both asymmetries.

The second property in (23) that the present approach sheds light on is the fact that OO-F constraints pertaining to different dimensions cluster in rank. For example, level 2 affixes impose uniformly high-ranked OO-F, blocking re-stressing: *éffort/ éffort-less*; vowel shortening: *cri:me*/ *cri:me-less*; and other alternations at the same time. In contrast, their level 1 counterparts feature uniformly low-ranked OO-F, whence re-stressing as in *párent/ parént-al*; shortening as in *cri:me*/ *crimin-al*; etc. When the different dimensions are implicated by the same lexical item, the clustering is of course just Gradient Attraction (5): if OO-F for vowel length is high-ranked, *cri:me* will yield *cri:me-less-ness* with a long vowel, and identity for vowel length will create greater pressure for identity in stress, whence the clustering effect. When different items are involved, however, as in *cri:me-less*, faithful for vowel length, versus *éffort-less*, faithful for stress, clustering in rank no longer follows in this simple way, but is rather effected indirectly, via the affix. In essence, it is due to the fact that items like *crí:me-less*, faithful for both stress and vowel length, will generate entailments like: *-less* ⇒/ N ___ in both dimensions simultaneously, i.e. *-less* entails a stem which equals a noun for vowel length, and one that equals a noun for stress. The affix will thus come to 'encode' the clustering of OO-F effects produced by this class of items under Gradient Attraction, and then essentially 'transfer' it to other items, like *éffort-less*, faithful for stress.

The third point in (23) is that only the present approach derives the Gradient Attraction effect (5), namely the fact that independent similarity affects the rank of OO-F constraints --a fundamental generalization, it seems.

The fourth point in (23) is that the rank of OO-F constraints, aside from the effects of similarity just discussed, is attributable to entailment summation. This means that that rank ultimately comes from the size of the class over which summation occurs. Consider in this connection that phonological alternations found with phrases are fewer than those found with words, so the flapping of *a[ɾ]om*, *permi[ɾ]ing* is also found with *he sho[ɾ] a man*, but the 'de-syllabification' of *cyc[l̩]*/ *cyc[l]-ing* is not found in *cyc[l̩] incessantly*. The 'Lexical Phonology' of Kiparsky (1982a, b) and later work captured this fact, and more generally the fact that phonological alternations are fewer and fewer with larger and larger units (words, compounds, phrases), by assuming, as seemed natural, that smaller units are formed first, and that phonology of different sorts applies at each stage of composition. Smaller units go through more stages of composition and thus more points at which phonology can apply. So the flapping of both *a[ɾ]om* and *he sho[ɾ] a man* can be attributed to phrasal level phonology (if only by overlooking the fact that *a[ɾ]om* is actually not a phrase), while the different syllabifications of *cyc[l̩]*/ *cyc[l]-ing* are attributable to word-level phonology, long past when formation of *cycle incessantly* occurs. But there is of course no need to assume that larger units are formed 'later' than smaller ones. We can perfectly well assume that all units --segments, syllables, words and phrases, are calculated in parallel. It is in fact not order of derivation, but rather another property, that necessarily correlates

with the size of the units. It is the size of their classes. Affixal configurations *X-A*, where A is a specific affix, have a potential class size *n*, where *n* is the number of items of class *X*. On the other hand, phrasal configurations *XY*, where both *X* and *Y* are lexical items have a potential class size *n*m*, where *n, m* are the number of items of categories *n, m*, respectively. Assuming that actual class sizes (reflecting restrictions of various sorts) are commensurate with potential size, the number of entailments $X \Rightarrow / \_\_ Y$; $Y \Rightarrow / X \_\_$ generated by phrasal structures will be one order of magnitude greater than the number of entailments $A \Rightarrow / X \_\_$ generated by affixed structures. This will make words in phrases inherently more stable than affixed stems, which seems correct. Note that summation of entailments generated by phrases must be assumed to occur 'in use', rather than 'over the lexicon' as was assumed for *-al* structures in (19) above, since presumably phrases are not lexically stored. The same assumption seems necessary for productive affixational classes, like level 2 affixation. This will not affect the essence of the proposal. Note as well that since phrasal dependencies can be long-distance, as in *I heard the professor's lecture*, where a mutual dependency holds between *hear* and *lecture*, stability can have a long-distance source. That is, here, the phonological integrity of *lecture* is sponsored by *read*, that selects (entails the presence of) a noun as the head of its complement. This will also not affect the above conclusion that words will be relatively more stable in phrasal structures.

This approach makes sense of the fact that alternations within phrases often affect closed-class items (like determiners, pronouns, auxiliaries), e.g. *a pear/ an apple*; *[ðə] pear/ [ðiy] apple*. The reason is that phrasal combinations *XY*, where *X* is a member of a small class, will themselves constitute a small class, tending to the size of the *Y* class, as with affixes, hence their relative vulnerability to variation. The allomorphy of both the definite and the indefinite determiner is in fact not too different from the *t/ d/ ɪd* allomorphy of the regular past-tense, or the similar one of the regular plural. All cases combine a closed-class item with an open-class one. The remaining question is then why, given relative vulnerability of such combinations, it is the closed-class item (past-tense/ plural affix, or determiner) that alternates. The answer is again 'numbers'. Alternations in the open class items would produce a large number of allomorphs, while alternations in the closed-class items produce only few, in proportion to the size of each class. A different way of putting it is that the *an* of *an apple* can serve *orange* as well, as in *an orange*, that is it will act as a general alternative attractor, beside *a*, for the expression of the indefinite determiner. In contrast, hypothetical *napple* alternating with *apple*, will not correspondingly help *norange* by attracting it away from *orange*, since *napple* and *norange* are different points in space (see also Burzio 1999c). From a derivational point of view, there is no reason for either of these two facts. The phrase *[ðiy] apples* is presumably formed at the same level as *I loath apples*. If phonological activity is controlled by level of derivation, one should find comparable **I loa[ðiy] apples* ('I loath apples'). There is no reason as well why the closed-class item should be targeted, thus making **lo[ft]* a fine alternative to *lo[vd]* (loved), etc. Note that *le[ft], bere[ft]* etc., with stem alternations, do exist, but these indeed form a small class of verbs, as the above considerations predict. As noted in Burzio (1997), the numerical basis of morpho-phonological stability is also supported by Romance verb conjugations, with the overwhelmingly more numerous first conjugation exhibiting virtually no alternations, in contrast to the other, less numerous conjugations, that exhibit many alternations (for further evidence from English word stress see also Burzio 1994a, index item: 'Constraints, reflecting numerical or

statistical factors).

Thus, in general, larger units form larger classes, and as such they generate larger classes of entailments that enforce stability on their components. Alternations with larger units such as phrases typically crop up under two circumstances: when one of the components is a closed class item as just noted, and when the relevant phonological constraint is particularly high-ranked, as with the 'flapping' of American English, as in *sho[ɾ] a man*. Other alternations will be blocked. Hence the different subsystems of the 'Lexical Phonology', associated with different levels, indeed reduce in a parallel system to different rankings of OO-F constraints (Benua 1997, Burzio 1994a). In turn, this difference seems to reduce to class size, which clearly distinguishes affixal from phrasal formations. The difference between level 1 and level 2 affixes can then also be naturally attributed to a difference in the class of stems, smaller for level 1, larger for level 2 affixes, yielding a difference in the 'enhancement' effects in (20), resulting in the weaker/ stronger attraction effects of (21)-(22), which in turn account for the multiple differences in behavior between the two groups of affixes (Burzio in press, 1999c).

The final point in (23), that this approach does away with an independent morphological module was the stated goal of the article. The present perspective is in agreement with Benua's notion that the rank of OO-F constraints imposed by affixes on their stems is part of the 'subcategorization frame' of the affix, but in fact it ties the two together more intimately. OO-F constraints and subcategorization frames for affixes are non distinct notions in the present system, in the form of macro-entailments like: *al* ⇒/ *N* __, which are violable. Hence, here it is the subcategorization frame itself that is ranked. The paradox that the present work aims to dispel is in the idea that there is a subcategorization frame, or equivalently a WFR, that is enforced inviolably upstream of the phonology, in an independent box called 'morphology' and that the product is then shipped to the phonology with some license agreement that specifies the degree of phonological modification accepted (rank of OO-F). It seems obvious that there is no point in having an inviolable subcategorization/ WFR under these circumstances. Further proof of this is in cases like *arbore-al, symptomAT-ic, spiritU-al* that violate that subcategorization for reasons other than phonological. Again, violable subcategorization frames arise in the present system as a reflex of RE (12). When an affix A occurs in an environment E it generates an entailment that its presence be confined to E, or A ⇒/ E. Entailments of this sort compound over the lexicon into a macro-entailment/ constraint, whose rank/ weight reflects their number. When numbers are small, macro-entailments are weak/ low-ranked, resulting in violability by the other two components of the system: phonological markedness, as in *na:ture/ natur-al*, *cri:me/ crimin-al*, and IO-F, as in **arbore*/ *arbore-al*, *spirit/ spiritU-al*, etc. (Burzio, in press, 1999c). When the numbers are large, the macro-entailments are strong/ high-ranked, resulting in relative inviolability, whence *cri:me-less*, **arbore-less*, etc.

7. Gradient Attraction Effects

In this section I consider further Gradient Attraction effects, namely cases where identity over one dimension appears to promote identity over others, making it impossible to assign a fixed rank to OO-F constraints. All cases are amenable to an account along the lines of (6), (7) above, except for the last one, which will require additional discussion. Space prevents me from

discussing the analyses offered in each of the references. Suffice it to say that the present solution aims to be more principled in each case, as well as more general.

The first set of cases includes the already discussed *comparable* of (6) above.

(24)	**Base**	**More similar**	**Less similar**	**Interacting dimensions**
a.	compáre	compárable	cómparable	stress; semantics
b.	divíde	divídable	divísible	vowel length; segmentism
c.	applý dený	deníable	ápplicable	vowel length; segmentism
d.	rémedy lévy	léviable	remé:diable	stress; vowel length
e.	lárynx	lárynxes	larýnges	stress; segmentism

Here a derivative which is 'more similar' to a base, differs from another 'less similar' derivative across at least two dimensions simultaneously, as we saw for (a) (=(6) above). In (b), *divis-ible* has idiosyncratic segmentism (*d* ⇒*s*) and is affected by vowel shortening (a phonological regularity), while *divid-able* is instead faithful both ways. Similarly, in (c) *applic-able* is affected by shortening, and also features an inserted *c* (for partly phonological, partly idiosyncratic reasons), while *deni-able* is faithful both ways. The difference in vowel length is then responsible for the different stresses. In (d), *remedi-able* is re-stressed, for the reasons discussed for (8) above, and is also affected by regular *'CiV*'-lengthening (like *cana:dian*), while *levi-able*, that does not re-stress, also fails to undergo *'CiV*'-lengthening. In (e), *larynx-es* has an irregular stress pattern, forced by OO-F. When segmental identity is lost, however, as in (segmentally idiosyncratic) *laryng-es*, the stress is regularized to the heavy penultimate.

The next set is from Steriade's (1998) discussion of French liaison:

(25)			**Interacting dimensions**
a.	nouv[ɛ]l ami (*nouv[o]l)	'new friend'	Final C, preceding V
b.	proch[ɛ̃]n-arrêt (dialectal)	'next stop'	Gender, V-quality (here nasalization)
c.	proch[e]n-arrêt (dialectal)	,,	Final C, preceding Vquality
d.	gro[z] arbre	'big tree'	Gender, Vquality (cf. *grɔs*)
e.	*for[t] accent	'strong accent'	Final C, gender

Liaison contexts are prevocalic and require a final C to serve as a syllable onset. Such C is generally supplied by the feminine allomorph. The case in (a) shows that the final C does not come by itself, but is rather accompanied by the feminine allomorph's adjacent vowel *[ɛ]* that supplants the *[o]* of the masculine (spelled '*nouveau*'). Hence faithfulness for final C entails faithfulness for a preceding V as well. However, since the liaison form is masculine, hence faithful for gender to the masculine citation form, it will be 'attracted' to that form as well, in competition with the feminine allomorph. That competition shows up when the vowel of the feminine is closely similar to that of the masculine, as in the dialectal variation of (b): vowel of

masculine *proch[ɛ̃]*, and clustering of gender and V quality, versus (c): vowel of feminine *proch[e]n*, and clustering of final C and V quality. That competition is absent in (d) despite the closeness of the vowels of Masc. *gr[o]*/ Fem. *gr[ɔ]s*, because of independent factors that voice the final C (see Steriade's discussion), thus distancing the feminine allomorph and weakening its attraction. The masculine vowel, associated with the right gender, thus wins. In (e), the [r] of the masculine is sufficient to provide an onset, and the C of the feminine is thus no longer compelled. It is then excluded by clustering of gender and final C.

The next set of cases, given only schematically in (26), is from Łubowicz (1998), and includes the already discussed (26a).

(26)			**Interacting dimensions**
a. Polish	i. k ⇒č		coronality, continuancy
	ii. g ⇒ǰ ⇒ž		
	iii. ǰ ⇒*ž	(underived)	
b. Slovak	i. i ⇒i:		length, diphthongization
	ii. e ⇒e: ⇒ie		
	iii. e: ⇒*ie	(underived)	
c. Campidanian Sardinian	i. s ⇒z		voicing, continuancy
	ii. p ⇒b ⇒β		
	iii. b ⇒*β	(underived)	
d. Tiberian Hebrew	i. i ⇒i:		length, height
	ii. i ⇒i:⇒e:		
	iii. i: ⇒*e:	(underived)	

In each case, there is a phonological change that occurs as in the second step of (ii), only if some other phonological change, given in the first step of (ii) and in (i), has already 'distanced' the derivative from its base. Otherwise there is no change, as in (iii), which thus constitute *phonologically* 'underived' environments (they are all *morphologically* derived).

A case involving purely morphological properties is provided by Icelandic unaccusative verbs like those in (27).

(27) a. *Vegurinn breikkaði*
the road-N widened

b. *Bátinn fyllti*
the boat-A filled

The two verbs in (27) are representative of the more general fact that unaccusatives that have transitive counterparts split into two subclasses in Icelandic: those that assign Nominative to their subject like their English counterparts: (a); and those that assign Accusative to their subjects --the same Case that that argument would receive in the transitive structure, where it is an object. Zaenen and Maling (1990) have found that the difference in (27) correlates with another difference: verbs of group (b) inflect identically to their transitive counterparts, while those of group (a) exhibit independent inflectional idiosyncrasies. This the familiar clustering pattern, as indicated in (28), where 'P' is some inflectional property, and 'A' is ability to assign Accusative to a patient. For further discussion, see Burzio (to appear), where I propose an OO-F approach to

Case-marking patterns.

(28)	**Base**	**More similar**	**Less similar**	**Interacting dimensions**
	TransitiveV with P A	Unacc. V with P A	Unacc. V with ¬ P ¬ A	Inflectional properties Case assigning properties

We may pause here to note the heterogeneity of the interacting dimensions in (24)-(26), (28), consistent with RE (12), under which each dimension in a complex representation generates entailments for all others.

The last case I consider is a more complex and particularly interesting one. According to Fukazawa (1998, ch. 4), Yucatec Maya exhibits patterns of alternation that, idealizing somewhat, can be characterized as in (29).

(29)	Input	Output	Repair	**Interacting dimensions**
a.	...t-t...	...h-t...	Total dissimilation	place, manner
b.	...ʔ-t... ...k-t...	✔ ✔	No repair	
c.	...s-t...	✔	No repair	

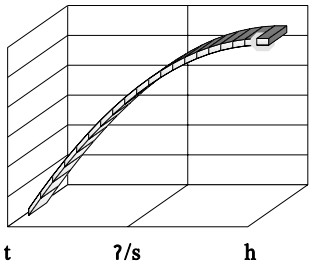
The reported generalization is that combinations of homorganic stops dissimilate the first member in both place and manner (oral stop ⇒glottal fricative), as in (a), while combinations that already differ either in place (b), or manner (c), undergo no repair. Assuming dissimilation is due to the 'OCP', it is easy to see that absolute ranking of constraints yields straight paradoxes: the OCP for each of place and manner must dominate the corresponding IO-F constraint for (a), but must instead be dominated by it for each of (b, c). Combining the two OCP constraints into a conjunction of higher rank for (a) offers little help in itself, because: first, there is no reason why the two IO-F constraints should not also combine and then block the OCP jointly just as they do individually in (b, c); and, secondly, because the repair in (a) should in any event stop at partial dissimilation, enough satisfy the OCP conjunction (*SAME MANNER& SAME PLACE).

Turning then to the present perspective, we note first that the structures in question are subject to OO-F, since they are phrasal, combining words (see references). Assuming then that the rank of OO-F constraints is modulated by independent similarity as before, we make some sense of the fact that a change in place occurs only if a change in manner also occurs (and vice-versa). This will not be sufficient, however. If the OCP had a fixed rank, it could not prevail in (a) without also prevailing in either (b) or (c). It will thus be necessary to assume that the rank of the OCP is also modulated, as if it were itself part of the OO-F family. As I note in Burzio (1998b), the OCP can indeed be naturally related to OO-F. 'OCP' effects simply reveal that the notion of 'contrast', which is a property of inventories as discussed above, is also a property of sequences. We can relate this to the fact that sequential contiguity as a form of similarity -- similarity in time. Elements that combine this similarity with structural similarity will then be under special pressure to neutralize, i.e. become identical, including in time, effectively

eliminating one of the elements. Similarity under sequential contiguity is now ‘marked’ for the same reasons that similarity of contrasting elements is marked --because under threat of neutralization. This makes the harmony function in (18a) above applicable here as well, now under the label ‘OCP harmony’, as in (30).

(30)

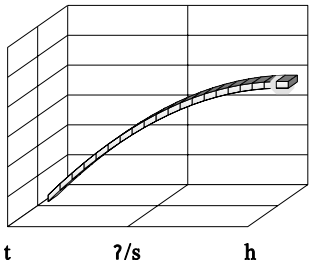
OCP harmony /_t



In the context: ___*t*, it is worst (minimal harmony) to have a *t*, better to have a *ʔ* or *s* (difference in place or manner), and better still to have an *h* (difference in both). As always, the optimal output is given by maximal harmony, and that will depend on both (30) and the OO-F harmony, given in (31).

(31)

OO-F *dis*-harmony
for a base ...*t*...

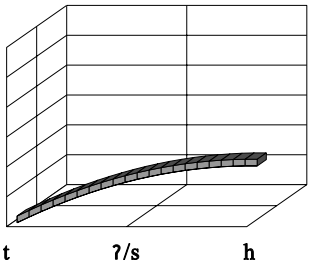


The effect in (31) is the usual OO-F curve of (16a) above, disharmony growing with distance at slower and slower rates. Hence (30) and (31) are essentially similar curves, just in opposite directions (positive vs. negative harmony). I take this to hold despite the fact that important details remain sketchy. So, there is no obvious reason why the curve in (30) should top at *[h]*, and not at *[a]*, for example --a vocalic segment even more different from *[t]*. There is also no obvious reason (and it is probably not the case) that (31) should top at *[h]*, given that what is being compared are word-sized representations which remain invariant except for the segment in question. Still, the two curves in (30) and (31) are expected to have somewhat similar patterns over a certain range. Overall harmony, then, should result from the algebraic sum of the two curves, and that result will be a curve of similar shape, either positive or negative, depending on the relative magnitude of (30) and (31). Now, the generalization in (29a) (*t-t* ⇒*h-t*) would simply reveal that the OCP curve has the greater magnitude, the result being positive, as in (32), with maximal harmony indeed at *[h]* (total dissimilation).

(32)

Overall Harmony /_t
for a base ...*t*...

[equals: (30) minus (31)]

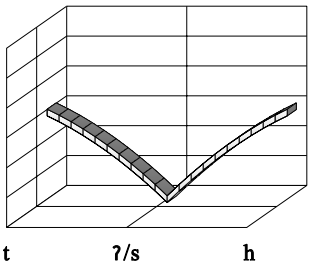


That is, the two curves being similar, if the OCP prevails over OO-F, it will do so all the way through, yielding total dissimilation as in (29a).

Matters will be different if the base has either *ʔ* or *s* in the same context __*t*, however. In that case, the relevant OO-F curve will not be (31), but (33).

(33)

OO-F *dis*-harmony
for a base ...*ʔ*... or ...*s*..

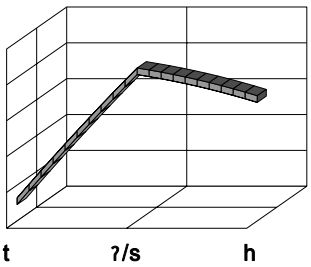


That is, the OO-F curve itself has the same shape as before, but its origin (zero disharmony) is here shifted to *ʔ* or *s*, the present form of the base, with disharmony then growing with distance from that point, here in either direction. But now summation of the OCP curve (30) with (33) (which is negative) yields (34), with maximal harmony indeed at *ʔ* or *s* as in (29b, c).

(34)

Overall Harmony /_t
for a base ...*ʔ*... or ...*s*...

[equals: (30) minus (33)]



That is to say, in both (32) and (34) the OCP and OO-F are in competition. However, for (32) the two curves (30) and (31) are essentially 'in synch', so that if the OCP prevails at one point, it will prevail at all others, yielding total dissimilation. In contrast, for (34), the two curves (30) and (33) are 'out of synch', with the relatively flat portion of the OCP curve facing the steep portion of the OO-F curve. This is why here it is the latter that wins --the solution to the ranking paradox. No principled solution appears to exist in which either the OCP or the OO-F fails to be non-linear. (Burzio 1999a).

8. Conclusion

Phonological output forms cannot be correctly calculated without reference to other output

forms. The introduction of direct relations among surface forms, however, tends to undercut the role of traditional morphology whose function was to express word-to-word relations indirectly, via underlying representations. I have proposed to expand the role of output to output relations in a way that can take over the function of morphology completely, reworking it into a system of violable constraints, and yielding an overall theory of the lexicon that is more fully consistent with the parallel approach advanced by OT. What makes the reconceptualization possible is the 'Representational Entailments' assumption, under which representational notions like that of 'morpheme' as well as distributional properties once captured by word-formation rules, can be converted into violable constraints, which are then better able to interact with the rest of the system. The latter assumption yields a faithfulness function which is both necessary to account for the patterns of allomorphy that I referred to as 'Gradient Attraction', and independently attested by the structure of phonemic inventories, being essentially implicit in Flemming's (1995) 'Dispersion Theory' of such inventories. Such parallelism of effects provides support for the present attempt to unify morphology and phonology under a single system of violable constraints. The two domains of morphology and phonology are already defined by the different 'granularity' of the phenomena they deal with. It would be an accident if they were also defined, independently but concurrently, by different internal architectures.

References

Anderson, S.R. (1992) *Amorphous Morphology*, Cambridge University Press.

Benua, L. (1997) *Transderivational Identity: Phonological Relations between Words*, Ph.D. Dissertation, U. Mass. Amherst.

Burzio, L. (1991) 'On The Metrical Unity of Latinate Affixes,' in G. Westphal, B. Ao, and H.R. Chae (eds.) *Proceeding of the Eighth Eastern States Conference on Linguistics*, Department of Linguistics, Ohio State University. Reprinted in *Rivista di Grammatica Generativa* 16, 1-27. Revised version in H. Campos and P.M. Kempchinsky (1995), (eds.) *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, Georgetown University Press.

Burzio, L. (1992) 'Principles in Phonology,' in E. Fava (ed.) Proceedings of the XVII Meeting on Generative Grammar (Trieste, February 22-24, 1991), 97-119. Rosenberg and Sellier, Torino.

Burzio, L. (1993) 'English Stress, Vowel Length and Modularity,' *Journal of Linguistics*, 29.2, 359-418.

Burzio, L. (1994a) *Principles of English Stress*, Cambridge University Press.

Burzio, L. (1994b) 'Metrical Consistency,' in E.S. Ristad (ed.) *Language Computations*, American Mathematical Society, Providence RI.

Burzio, L. (1996) 'Surface Constraints versus Underlying Representation,' in: J. Durand and B. Laks (eds.) *Current Trends in Phonology: Models and Methods*. European Studies Research Institute, University of Salford Publications. (123-141).

Burzio, L. (1997) 'Strength in Numbers', in V.Miglio and B. Morén (eds.) *University of Maryland Working Papers in Linguistics 5 (1997): Selected phonology papers from H-O-*

T 97, 27-52.
Burzio, L. (1998a) 'Multiple Correspondence', *Lingua 103*, 79-109.
Burzio, L. (1998b) 'What's Hebb's Learning Rule doing in Phonology?, ms. Johns Hopkins University.
Burzio, L. (1999a) 'The Incremental but Non-linear Nature of Faithfulness', ms. Johns Hopkins University.
Burzio, L. (1999b): 'Italian Participial Morphology and Correspondence Theory,' in G. Booij, A. Ralli and S. Scalise (eds.) Proceedings of the First Mediterranean Conference of Morphology. University of Patras.
Burzio, L. (1999c) 'Missing Players: Phonology and the Past-tense Debate,' ms. Johns Hopkins University.
Burzio, L. (to appear): 'Anatomy of a Generalization,' in Eric Reuland (ed.) *Arguments and Case: Explaining Burzio's Generalization*, John Benjamins, Amsterdam.
Burzio, L. (in press) 'Cycles, Non-Derived-Environment Blocking, and Correspondence,'in J. Dekkers, F. van der Leeuw and J. van de Weijer (eds.) *Optimality Theory: Syntax, Phonology, and Acquisition.*
Bybee, J.L. (1985) *Morphology: A Study of the Relation between Meaning and Form*, J. Benjamins, Amsterdam.
Bybee, J.L. (1988) 'Morphology as lexical organization,' in M.Hammond and M. Noonan (eds.) *Theoretical Morphology*, San Diego, CA: Academic Press.
Bybee, J.L. (1995) 'Regular Morphology and the Lexicon', *Language and Cognitive Processes*, 10 (5), 425-455.
Flemming, E. (1995) *Auditory Representations in Phonology* Ph.D. Dissertations, UCLA.
Fukazawa, H. (1998) *Theoretical Implications of OCP Effects on Features in Optimality Theory*, Ph.D. Dissertation, University of Maryland, College Park.
Goldrick, M. and P. Smolensky (1999) 'Opacity, Turbid Representations, and Output-based Explanation,' paper presented at the University of Alberta, June 11, 1999.
Hayes, B. (1999) 'Phonological Acquisition in the Early Stages', ms. UCLA. Rutgers Optimality Archive 327.
Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory*, John Wiley and Sons.
Kenstowicz, M. (1996) 'Base-identity and Uniform Exponence: Alternatives to Cyclicity,' in: J. Durand and B. Laks (eds.) *Current Trends in Phonology: Models and Methods*. European Studies Research Institute, University of Salford Publications. 363-393.
Kiparsky, P. (1982a) 'Lexical Phonology and Morphology,' in I.S. Yange ed. *Linguistics in the Morning Calm*, Seoul.
Kiparsky, P. (1982b) 'From cyclic phonology to lexical phonology,' in H. van der Hulst and N. Smith (eds.) *The Structure of Phonological Representations (Part I)*, 131-175. Dordrecht: Foris.
Kirchner, R. (1997) 'Contrastiveness and Faithfulness,' *Phonology* 14, 83-113..
Łubowicz, A. (1998) 'Derived Environment Effects in OT', Rutgers Optimality Archive 239.
McCarthy, J. (1998) 'Sympathy and Phonological Opacity'. To appear in *Phonology*. Rutgers Optimality Archive 252.

McCarthy, J., and A. Prince (1993) 'Prosodic Morphology I: Constraint Interaction and Satisfaction,' ms., University of Massachusetts, Amherst and Rutgers University, New Brunswick.

McCarthy, J. and A. Prince (1995) 'Faithfulness and Reduplicative Identity,' in J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, Amherst: GLSA.

Prince, A. and P. Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, ms., Rutgers University, New Brunswick, and University of Colorado, Boulder. To appear, MIT Press.

Smolensky, P. (1986) 'Information Processing in Dynamical Systems: Foundations of Harmony Theory. In D.E. Rumelhart and J.L. McClelland (eds.) *Parallel Distributed Processing: Exploration in the Microstructure of Cognition. Volume 1. Psychological and Biological Models*:.194-281. MIT Press.

Smolensky, P. (1993) 'Harmony, Markedness and Phonological Activity,' Rutgers Optimality Archive 87.

Smolensky, P. (1997) 'Constraint Interaction in Generative Grammar II: Local Conjunction', paper presented at the Hopkins Optimality Theory Workshop/ University of Maryland Mayfest 1997, May 8-12, 1997.

Steriade, D. (1994) 'Positional Neutralization and the Expression of Contrast,' ms. UCLA.

Steriade, D. (1996) 'Paradigm Uniformity and the Phonetics-Phonology Boundary,' to appear in M. Broe and J. Pierrehumbert (eds.) *Papers in Laboratory Phonology 5*, Cambridge University Press.

Steriade, D. (1997a) 'Lexical conservatism,' in SICOL 1997: *Linguistics in the Morning Calm*, Hanshin, Seoul, Korea.

Steriade, D. (1997b) 'Phonetics in Phonology: The Case of Laryngeal Neutralization,' ms. UCLA.

Steriade, D. (1998) 'Lexical Conservatism in French Adjectival Liaison,' ms. UCLA.

Wilson, C. (1999) 'Consonant Clustser Neutralization and Targeted Constraints,' ms. Johns Hopkins University.

Wilson, C. (in preparation) *Targeted Constraints: An Approach to Neutralization in Optimality Theory*, Ph.D. Dissertation, Johns Hopkins University.

Zaenen, A. and J. Maling (1990) 'Unaccusative, Passive, and Quirky Case,' in J. Maling and A. Zaenen (eds.) *Modern Icelandic Syntax. Syntax and Semantics, Vol. 24*, Academic Press, San Diego.