4

Sources of Paradigm Uniformity

Luigi Burzio

4.1 INTRODUCTION

Generative Phonology was only in its teens, when J. W. Harris (1973) argued on the basis of facts like those in (1) that Paradigm Uniformity (henceforth 'PU') must be recognized as a principle of grammar.

(1)

love-Imperf. Indic.	Latin	Spanish
1SG	amá:bo	amába
2SG	amá:bas	amábas
3SG	amá:bat	amába
1PL	ama:bá:mus	amábamos
2PL	ama:bá:tis	amábais
3PL	amá:bant	amában

Since stress in the Romance languages is quite generally faithful to the original stress of Latin, there seems no reason for the shifts in the first and second plural in (1) unless 'paradigms' are under pressure to be uniform. The theoretical machinery of the day had no way of incorporating PU, however, and Harris's proposal was quickly forgotten. Years later, Roca (1992) will argue that Spanish simply has two different stress

I am grateful to the editors for insightful comments that led to significant improvements of both the form and content of this chapter.

systems, one for verbs and one for nouns. The former is essentially 'lexical', associating stress with a specific morpheme, in (1) the thematic vowel *-a-*, while the latter calculates stress from the right edge of individual words, though not always in the same way: *señór* 'gentleman', *pistóla* 'pistol', *fábrica* 'factory', etc. While theoretically troublesome, Harris's solution would have been superior to Roca's, since it required no stipulation: PU obtains only with verbs because in Spanish only verbs form this kind of inflectional paradigm.[1]

In Burzio (1994*a*) (henceforth *PES*) and earlier work leading up to it (1991 *et seq.*) I proposed an account of the English stress-transfer effects in (2*a'*, *b'*) based on the three blocks of violable constraints of (2.I, II, III).

(2)

	I. MWF (Metrical Well-Formedness)	II. MC (Metrical Consistency)	III. MA (Metrical Alignment)
a. an(tágonis)tø	√	n/a	√
a'. a(mérica)(nìstø)	√	√	*
b. (ònoma)(tólogy)	√	n/a	√
b'. phe(nòme)(nólogy)	√	√	*

Example (*a*) shows the preferred right-edge Metrical Alignment (III) for nouns: a foot boundary coincides with the end of the last syllable with an overt nucleus, the final consonant *t* forming a degenerate syllable with a null nucleus. Example (*a'*) shows that this requirement can be violated when so compelled by 'Metrical Consistency' – the competing requirement that morphologically complex forms stress consistently with their bases, the base here being *américan*. Example (*b*) shows the preferred Metrical Alignment at the left edge of a word, with a foot boundary coinciding with the left edge of the first syllable. This requirement is violated in (*b'*) again as compelled by Metrical Consistency, this time to maintain the stress of *phenómenon*. In (2), top-ranked MWF defines the range of possible feet, thus accounting for the absence of Metrical Consistency effects under foot degeneracy, as in **ca(tàs)tróphic* (cf. *catástrophe*); **(cómpensa)tòry* (cf. *cómpensàte*). In the *PES* analysis these cases are excluded by the ill-formedness of unary feet and of ternary feet with a heavy median syllable, respectively.

The *PES* approach, which independently introduced basic elements of Prince and Smolensky's (1993) 'OT', directly accounts for the facts in (1). Latin primary stress reflected a strictly deterministic system: a penultimate syllable was always stressed if heavy, and never if light (except in disyllables). In terms of (2), MWF left no window of opportunity for MC, whence the lack of PU in Latin. In contrast, Spanish MWF does have certain degrees of freedom: a light penultimate can either be stressed—*pistóla*—or not—*fábrica*. With top-ranked MWF equally satisfied, MC will

[1] To the limited extent that nouns also form inflectional paradigms, these are also uniform, as Harris had noted: *señór/señóres* 'gentleman/gentlemen', etc.

be correctly expected to make the choice, hence accounting for the uniform Spanish stress in (1)—the illusion of 'lexical' stress. In the Optimality Theoretic framework as further developed by McCarthy and Prince (1995, 1999), the 'Metrical Consistency' of *PES* and PU more generally can be formally treated as a specific instance of 'faithfulness'—the same type of relation that exists between an output and its input, or between a reduplicant and its base, as argued in Benua (1997*a*). The introduction of such Output–Output Faithfulness/Consistency into OT raises important questions, however, notably the ones in (3).

(3) *a.* Exactly which other output is the source of Output–Output Faithfulness for any form under calculation?
b. How may the existence of Output–Output Faithfulness (OO-F) constraints affect our view of morphology? Specifically, what must we make of the fact that the type of surface-to-surface resemblances established by traditional morphological relations are insufficient, requiring the addition of OO-F?
c. May 'cyclic' derivations, once employed to capture some transfer of phonological properties (such as the one in (2*b*′)) still be viable within OT after all as an alternative to OO-F, should questions (3*a*, *b*) prove troublesome?

In this chapter I address the critical issues in (3). I will argue, as I also do in Burzio (2002*b*), that PU effects ultimately show that, while the move towards parallelism initiated by OT was much on the right track, the degree of parallelism in the system is in fact more extensive, concerning not only the internal structure of phonology, but also the relationship between phonology and morphology. I will claim further that the mental calculations or 'processing' required by phonology and morphology are not only parallel, but also 'distributed', indeed as in 'Parallel Distributed Processing'—the formal term for connectionism. The specific organization of the chapter is that in ss. 4.2 and 4.3 I argue that representations influence or 'attract' each other to a degree that is inversely related to their geometrical or global distance. In s. 4.4, I argue that the role of global distance reveals that mental representations have fundamental properties of neural nets, and also that a major factor behind the OT notion of OO-F is precisely the noted attraction effect. Section 4.5 elaborates further on the proposal of s. 4.4, pointing to additional consequences of significance. In s. 4.6 I suggest that a coherent and improved approach to morphology can also be developed from this perspective, one that further defines the OO-F relations at work within the phonology in terms of morphological relations. In s. 4.7 I review the typology of PU effects, showing that it is captured by the present approach though not by traditional means. In s. 4.8 I conclude.

4.2 THE ROLE OF DISTANCE

PES (276 and *passim*) noted certain clustering effects. For instance, items like *larynx*, *pharynx*, *syrinx* have two attested plurals: a regular one like *lárynxes*, and an irregular

one like *larýnges*. The irregular plural has regular stress, on a heavy penultimate, while the 'regular' one has the stress of the singular. What one observes here is thus co-variation of two types of identity: segmental identity and identity in stress. The regular plural is identical to the singular both ways, while the irregular one is different both ways. Anderson (1992: 193f.) had noted similar clustering effects with *-able* adjectives. For instance, *cómparable*, which is different in stress from its base *compáre*, is also divergent from it semantically, while *compárable*, which retains the stress of the verb, is also strictly faithful to the verb semantically. Anderson suggested that *-able* may be ambiguous between being either a Level 1 or a Level 2 affix within Lexical Phonology (Kiparsky 1982*a*, *b*), given that both re-stressing and semantic divergence are known properties of Level 1 formations while being both absent at Level 2. While this view might be extended to the previous case, (Level 1 v. Level 2 plural), it does not really solve the clustering problem, but merely subsumes it under a fairly general stipulation of Lexical Phonology, since—as I note below—the latter framework is not particularly successful in accounting for this clustering of properties across the postulated levels. In addition, this account would not extend to other clustering cases such as the one in (4).

(4)

Base	Derivative	Faith (stress)	Faith (V-length)
a. rémedy	remé:di-able	*	*
b. lévy	lévi-able	√	√

In (4*a*) the derivative is unfaithful to its base with respect to both stress and vowel length, while in (4*b*) faithfulness obtains both ways. Here, it would not do to appeal to a Level 1/Level 2 distinction, since the difference between re-stressed *remédiable* and non re-stressed *léviable* is fully predictable from Level 1 properties alone. In the *PES* analysis, the metrical structure *(ré.me.di.a)<ble>, with extrametricality of the final 'weak' syllable and pre-antepenultimate stress is not well formed, barred by the MWF of (2) which excludes quadrisyllabic feet, while that of (lé.vi.a)<ble> is well formed. Sameness of stress is then evidently able to induce sameness of V-length in (4*b*), blocking the otherwise regular phenomenon of *CiV*-lengthening of (4*a*) and many other cases, like *Caná:dian*, *Mongó:lia*, etc. The Lexical Phonology machine provides no clue for this clustering. On that approach, *CiV*-lengthening would be expected to apply to both (4*a*, *b*) equally, whether or not it is a cyclic process. Hence, while inheritance of properties can sometimes be handled by 'cyclicity', the inheritance of the short vowel in *leviable* cannot be.

The cases illustrated in (5), from Steriade (1997*a*, 1999*a*), are also beyond the reach of any traditional means.

In French, when occurring in certain prevocalic contexts, masculine adjectives ending in a vowel resolve the potential vowel hiatus by taking on a form similar to that of the feminine allomorph. In the forms in (5), dialect (*a*) borrows just the final consonant [n] from the feminine form, while dialect (*b*) borrows both the final consonant

(5) French Liaison (Steriade 1997*a*, 1999): *prochain arrêt* 'next-M stop'
Dialect (*a*): [pʁɔʃɛ̃n aʁe]
Dialect (*b*): [pʁɔʃɛn aʁe]

Base 1 (Masc.)	Base 2 (Fem.)	Liaison form/ dialect	Faith-1(M)		Faith-2(F)	
			gender	V-quality	C#	V-quality
[pʁɔʃɛ̃] *prochain* 'next-M'	[pʁɔʃɛn] *prochaine* 'next-F'	*a.* pʁɔʃɛ̃n	√	√	√	*
		b. pʁɔʃɛn	√	*	√	√

and the non-nasalized preceding vowel of the feminine: [ɛn]. We can see this variation as the result of two competing clustering effects, expressed by the enclosed pairs of cells in (5). On the one hand, the liaison form is masculine, in agreement with the head noun, and is thus faithful to the masculine citation form (Base 1) in this respect. On the other hand, by borrowing the consonant from the feminine allomorph (Base 2), the phonology places the liaison form in a partial faithfulness relation with the latter as well. The dialectal variation in (5) would show that V-quality may cluster either with the final C, yielding [ɛ] in conformance with the feminine allomorph (case (*b*)), or with gender, yielding [ɛ̃] in conformance with the masculine one (case (*a*)). The 'clustering' interpretation is confirmed by the fact that, where Faith-2 (C#) is breached, as in [gʁoz aʁbʁ] *gros arbre* 'big-M tree' compared with feminine [gʁɔs] *grosse* 'big-F' ([z] v. [s]), V-quality is no longer influenced by the feminine allomorph, *[gʁɔz aʁbʁ], but only by the masculine [gʁo] *gros* 'big-M'. A special restriction imposes [z] rather than [s] in such liaison contexts, as discussed by Steriade.

Clustering of faithfulness effects characterizes as well one of the syndromes referred to as 'Non Derived Environment Blocking' (NDEB), exemplified in (6).

(6) Sanskrit ruki rule (Kiparsky 1982*a*)
a. No retroflection: sisa 'lotus' (non-derived)
b. Retroflection: siṣ-ṭa 'taught' (base: sās 'instruct')

The assimilatory process that turns [s] to retroflex [ṣ] after *r*, velars or high vowels (r, u, k, i) applies morpheme internally only if some other change has applied, as in (6*b*), where the [ā] of the base has ablauted to [i], but not as in (6*a*), where no other change has occurred. Hence faithfulness for retroflection of the [s] ([±high] according to Kiparsky) appears to cluster with faithfulness to other aspects of the representation of the morpheme.

What emerges from these cases and others (reviewed in Burzio 2002*b*) is the descriptive generalization in (7).

(7) Allomorphs attract each other in a way that is inversely related to their global distance.

This generalization describes the noted clustering effects: when faithfulness is satisfied on some dimension (shorter global distance), it tends to be satisfied on other dimensions as well (stronger attraction). The notion of 'allomorph' in (7) seems in fact superfluous. Since allomorphs are by definition structures that are relatively similar in sound and meaning, the definitional similarity of allomorphs may simply be subsumed under the notion of global distance of (7), and the specific term 'allomorphs' may be replaced by the more general 'representations'.

Summing up, we find that the faithfulness constraints that are at work in allomorphy defy unique ranking. Their rank appears to be rather modulated by distance, such that the closer the target representation is to the trigger of the faithfulness, the higher is the rank. The above discussion has also shown indirectly that the faithfulness that controls allomorphy cannot in general be faithfulness to some 'Underlying Representation' (UR). The reason is that such faithfulness is sensitive to surface properties that would not be present in UR. For instance, we have seen that lengthening in *remé:diable* is crucially aided by the fact that its stress is different than that of *rémedy*, yet the latter stress cannot be in UR since it is perfectly regular and predictable. Similarly, in the French case in (5), the liaison allomorph is not in a faithfulness relation to some UR. For one thing, a unique UR for both masculine and feminine allomorphs seems unlikely, the [ɛ̃] (MASC.)/[ɛn] (FEM.) alternation being more plausibly analyzed as just part of the morphology (suppletion) rather than phonologically induced. Even if there were such a common UR, say /pʁɔʃɛn/, the dialect (*a*) liaison form would not be derivable from it, since vowel nasalization would have to be contingent on deletion of the nasal to correctly derive masculine and feminine allomorphs, but this relation would automatically exclude [pʁɔʃɛ̃n], where both the nasal and nasalization coexist. Rather, the faithfulness relation in question appears to hold on related outputs, implicating no UR. Finally, we find that, while 'cyclicity' may capture some surface-to-surface faithfulness effects (since each cycle produces a surface form), it is challenged by the cases at hand. No form of cyclicity can relate liaison forms to masculine and feminine allomorphs simultaneously as just noted, and none can distinguish *remé:diable* from *léviable* as also noted. The conclusion is thus that it is the extent to which surface representations themselves intersect (global distance) that is relevant to allomorphy, not the extent to which their derivational histories do.

4.3 DISTANCE IN PHONOLOGY AND PHONETICS

While the previous section has illustrated the role of distance in allomorphy, that is distance between morpheme-sized units, it is well known that distance plays a central role with segment-sized units as well. This is the essence of the 'Dispersion Theory' of Lindblom (1986) and Flemming (1995) (see also Padgett 1997), embodying the claim that segmental inventories are shaped by maximal perceptual distance between their members. That claim is expressible as the same notion of attraction over distance

of (7) that was applicable to allomorphs. Reduced distance between segments results in greater attraction, forcing weakly contrasting segments to neutralize, thus bringing about greater distance between the remaining segments. Steriade (1994, 1997*b*) has argued that not only segmental inventories, but segmental neutralizations also attest to the role of perceptual distance. For instance, coda devoicing in Dutch ([bɛt] 'bed', versus [bɛdən] 'beds') results in her view from the fact that the perceptual cues for voicing in the positions that we generally refer to as 'codas' are weaker, thus heightening the similarity/reducing the distance between members of the voiced/unvoiced contrasts. In this conception, segmental neutralizations are the expected inventory reductions in contexts that effectively shrink the perceptual space.

Wilson (2000) has shown further that a proper OT characterization of neutralization effects resolves formal problems that had long remained recalcitrant within OT, including the 'opacity' (counterfeeding/counterbleeding) effects. While the reader is referred to Wilson (2000, 2001) for detailed discussion, and Burzio (2000*b*, 2001) for a synopsis, a further succinct characterization can be provided for present purposes by comparison with McCarthy's relatively well-known 'Sympathy' Theory (McCarthy 1999; Kager 1999*b*: s. 9.2). McCarthy shows that opacity effects can be produced within OT by requiring an output to be faithful not only to its input, but simultaneously also to another specially designated output candidate, referred to as the 'sympathetic' or 'flower' candidate. The essence of Wilson's contribution is in showing that the needed further source of faithfulness or 'attractor' is independently provided by the theory of segmental neutralizations. On Wilson's solution, opacity effects result from the fact that, while a word or morpheme-sized representation is being faithful to its own input (which, on the present approach, is just another output), some segment-sized substructure within it is simultaneously also in an attraction relation with a competing representation of that substructure. In the Dutch example, the degenerate [d] of output candidate [bɛd] is being attracted to the independent well-formed representation [t], whence [bɛt].

Assimilatory phenomena also seem to be driven by distance. Thus, it seems rather generally the case that segments that assimilate are independently similar. For example, in Spanish, voiced stops spirantize post-vocalically, but unvoiced ones do not: *Cu*[ß]*a* 'Cuba'/*co*[*p*]*a* 'wine glass'. One can take this to indicate that continuancy is shared with a preceding vowel only if voicing is already shared. Similarly, the lack of nasal place assimilation in standard English *i*[*n*]*famous* can be attributed to clustering of place and continuancy. If this clustering obtains, then assimilation will be blocked by the markedness of the resulting nasal (labiodental) fricative *[ɱ] (although fast speech does tolerate the marked pronunciation, [iɱfəməs]). Similarly, in Sudanese Arabic a stop assimilates in continuancy to a following obstruent: *kitaa*[*b*] 'book' /*kitaa*[*f*]*Fáṭḥi* 'Fáṭḥi's book', but only if trigger and target already share place: *kitaa*[*p*] *Samíir* (not **kitaa*[*f*])'Samíir's book' (Kenstowicz 1994: 54). Vowel harmony phenomena are also often 'parasitic' on some degree of independent identity. For instance rounding harmony in Yawelmani depends on identity in [±high] (Kenstowicz 1994: 108).

It is well known that alongside of assimilatory effects, dissimilatory ('OCP') effects

also exist. They are also modulated by distance. For instance in Yucatec Maya (Fukazawa 1998), only sequences of obstruents that are homorganic in both place and manner (e.g. *tt*) undergo dissimilatory repairs. Sequences sharing only one of those features (*ṭt*, *kt*, *st*) undergo no repair.

The existence of both assimilation and dissimilation seems rather mysterious on a standard approach to markedness: if it is 'marked' to agree in features (OCP), how could it also be marked to *dis*agree (assimilations)? The notion of attraction sheds light on this mystery in so far as it entails two different regions of relative stability: identity, where attraction is satisfied, and greater distance, where attraction is reduced (see also discussion below). The locality or proximity conditions on both assimilatory and dissimilatory processes can themselves be understood in terms of attraction-inducing similarity by regarding sequential proximity itself as a form of similarity, as I discuss further below.

Hence attraction between segments plays a central role in several areas of phonology as currently understood and its links to perception. It seems also central to the notion of perception more generally. Our ability to map degenerate auditory inputs into a set of mental canons can be viewed in terms of attraction. If attraction over distance is a property of mental representations in general, then, in order to be interpreted, a representation formed on a perceptual input need not match a mental category exactly, but need be merely within its range of attraction. Mental categories can thus be viewed as 'perceptual magnets', indeed as independently proposed by Kuhl and Iverson (1995).

Repp (1982) describes as 'phonetic trading relations' the fact that the weakening of one perceptual cue can be offset by the strengthening of another. If global distance is what matters, then indeed distances over individual dimensions will be in trade-off relations with one another. Note that trading relations are implicit in the clustering effects observed above in so far as we found no absolute restriction on what dimensions could cluster together (segmental structure, prosodic structure, semantics). The ability of some identity effect I_A to cluster together equivalently either with some other identity I_B or yet some other identity I_C: I_A & $I_B \approx I_A$ & I_C constitutes a type of trading relation in the domain of allomorphy comparable to that described by Repp for perception. The effect identified in Ganong (1980) reveals further that the same perceptual mechanisms span over phoneme and word-size units. The latter effect consists of

(8)

Base 1 (Masc.)	Base 2 (Fem.)	Liaison form/ dialect	Faith-1(M)		Faith-2(F)	
			gender	V-quality	C#	V-quality
[nuvo] *nouveau* 'new-M'	[nuvɛl] *nouvelle* 'new-F'	*a.* *nuvol	√	?	√	*
		b. nuvɛl	√	*	√	√

the fact that a phoneme identification task is facilitated if the stimulus that it is part of is a lexical item. This is consistent with the present view that attraction over distance is a property of mental representations of any size. The Ganong effect can be characterized by saying that a phoneme-size perceptual unit is a candidate for attraction by a phoneme-sized mental category while at the same time the word-sized perceptual unit it is part of may also be attracted by a member of the word lexicon, and that the two attractions are cumulative. In fact, a similar conclusion would seem warranted by the French liaison case in (8) to be compared with that of (5).

The data in (8) are laid out identically to those in (5), but here there is no dialectal variation, as the form that takes the final C from the feminine and the preceding vowel from the masculine: **nuvol*, is not found. To make sense of this divergence, we must revise the above discussion slightly. We must assume that there are two simultaneous attraction effects from the masculine allomorph (Base 1): one at the level of the whole word, and one at the level of individual segments, exactly along the lines of the Ganong effect. The word-level effect is evidently by itself insufficient to impose the vowel of the masculine, whence (8*a*) **nuvol*, while the segment-level attraction is absent in (8) because [o] is not a perceptual neighbor of [ɛ]. FAITH-1(V-quality) is thus voided, as indicated in (8) by the question mark, and consequently so is the clustering of (8*a*). In contrast, in (5*a*) above the vowel [ɛ] of feminine [pʁɔʃɛn] *is* a perceptual neighbor of the [ɛ̃] of the masculine [pʁɔʃɛ̃n]. The resulting attraction by the masculine form thus results in satisfaction of FAITH-1(V-quality) and the clustering of (5*a*) above. Segment-level attraction cannot be sufficient to turn [ɛ] to [ɛ̃], however, since it is not a general property of French that [ɛ]s merge with their nasalized counterparts. Hence the word-level attraction must therefore also be at work, drawing the liaison form to the masculine allomorph under gender identity. Thus it appears that the purely distributional data of the French liaison effectively reproduce the properties of the Ganong effect: a segment-sized representation is influenced simultaneously by two types of neighbors, segment-level and word-level.

In conclusion, the notion of attraction over distance (7) finds a central role to play in phonology/phonetics as well as perception. In phonology/phonetics it is key to the characterization of segmental inventories, segmental neutralizations, opacity effects, assimilations, and dissimilations. In perception, the 'perceptual magnet' effect, the 'phonetic trading relations' and the Ganong effect, all seem to embody such a notion as well.

4.4 HEBB'S RULE

The attraction over distance of (7) is beyond the expressive power of the symbolic computation of traditional generative linguistics. The reason is that symbols, such as distinctive features in phonology, are by definition means to isolate individual dimensions of the representational space while abstracting away from others. As such, they

are insufficient to capture a notion of distance that is defined over multiple dimensions simultaneously. This is in contrast to neural computation, in which interactivity among dimensions is the natural by-product of the pattern of connectivity among units. In Burzio (2002*b*) I propose to capture the role of distance by means of the hypothesis in (9), with transparent links to principles of neural computation.

(9) Representational Entailments Hypothesis (REH): Mental representations of linguistic expressions are sets of entailments; e.g. a representation consisting of A and B corresponds to the entailments: $A \Rightarrow B$, $B \Rightarrow A$.

The REH (9) is a virtual restatement of Hebb's (1949) rule of mental learning according to which neurons that are active simultaneously develop synaptic connections that render their future co-activation necessary. The latter is parallel to the claim in (9) that simultaneous instantiation of A and B in some representation would make their co-instantiation necessary in other representations. Since the entailments of (9) are elementary constraints, the REH blurs traditional distinctions such as the one between representations and grammar and the one between grammar and lexicon. Under the REH all representations are effectively part of the grammar. Conversely, we will see below that formerly grammatical notions, such as that of the word-formation rule, and OT notions of both Output–Output faithfulness and Input–Output faithfulness can be reduced to representational entailments. The REH yields the attraction over distance effect of (7) in the way illustrated in (10).

(10) Effect of Entailments over Neighboring Representations

Entailments violated by ¬D in R2, for R1= A, B, C, D:	$A \Rightarrow D$ $B \Rightarrow D$ $A \Rightarrow D$ $C \Rightarrow D$ $B \Rightarrow D$ $A \Rightarrow D$			
	1	2	3	4
Varying R2:	A	A	A	¬A
	B	B	¬B	¬B
	C	¬C	¬C	¬C
	¬D	¬D	¬D	¬D

The diagram in (10) reports the effects of a representation R1 on a second representation R2, as the distance between them increases. R1 consists of the four components A, B, C, D. At point 1 in the diagram, R2 differs from R1 by a single component: ¬D. In this situation, the presence of ¬D in R2 gives rise to violations of three of the entailments generated by R1, as indicated. The reason for those violations is that R2 has each of A, B, C in common with R1, but not D. The diagram tracks the entailments violated by the presence of ¬D at each further step, as distance from R1 increases. At point 2, the number of entailments violated by the presence of ¬D decreases to two, because now R2 no longer shares C with R1, which was the source of one of the violations at point 1. Comparable decreases obtain at points 3 and 4. The violated entailments in (10) constitute the pressure on ¬D of R2 to become identical to its counterpart D of

R1—what we referred to above as 'attraction'. Such attraction is directly proportional to the number of components that R1 and R2 already share, namely their overall proximity, and hence inversely proportional to their overall distance. The attraction over distance effect, which I will also refer to as 'Gradient Attraction' (GA) can be further visualized as in (11).

(11) Gradient Attraction effect from entailments

R2

R1

In (11), the R1 of (10) is a point in multidimensional space which can be thought of as the center of a hole with variably sloping walls. The R2 of (10) can be thought of as an object on the slope, affected by gravity. The closer it moves to R1, the steeper the slope, and hence the stronger the force pushing it to the center. The linear function of (10) (linear decrease in the number of violated entailments) is the first derivative of the sloping function in (11) (steepness decreases linearly with distance from R1).

Representational entailments are understood to be elementary constraints that are violable like standard OT constraints, and will in fact be taken as the basis for various classes of OT constraints as noted. At the same time, entailment summation will be taken as the basis for the ranking of those constraints. Specifically, the class of Input–Output Faithfulness constraints (henceforth 'IO-F') will be characterized as a form of attraction in which the attractor, R1 of (11) is the input representation, while for the Output–Output Faithfulness constraints (henceforth 'OO-F') such an attractor will be an independently existing output form. At the same time, the attraction resulting from the entailments can also subsume various forms of markedness in standard OT, since it is held responsible for triggering various repairs, as in the assimilation/dissimilation and neutralization cases.

The clustering effects reviewed in s. 4.2 will now follow in the manner illustrated for the case of (4) above in (12)–(14) below.

(12)

R1: Attractor verb *rémedy/lévy*		R2: Forms of the stem	
A	Basic meaning	A	A
B	and segmentism	B	B
C	Stress	C	¬C
D	Length of stressed V	¬D	¬D
		OO-F (V-length)$_1$	≫ OO-F (V-length)$_2$
		(A, B, C ⇒ D)	(A, B ⇒ D)

In (12), it is assumed that the base verb, either *remedy* or *levy* of (4), plays the role of the attractor R1 in (10) and (11). For present purposes it will be sufficient to assign to

such representations the rough internal composition shown, using the four components A, B, C, D. We take A, B as stand-ins for the basic meaning and segmental composition of the verb, while C and D represent the exact position of the stress and the length of the stressed vowel, respectively. The stem of the respective *-able* adjectives is now R2 in (10), (11), subject to attraction by R1. Such attraction constitutes a form of OO-F in OT, except that, unlike standard OT constraints, its rank is modulated by distance. In (12), I consider OO-F for vowel length, and the modulation shown is due to the stress. When R2 has the same stress as R1 (C), a difference in vowel length (¬D) causes a three-entailment violation, reported in (12) as a higher ranked OO-F. However, when R2 has a different stress than R1 (¬C), then a difference in vowel length (¬D) only causes a two-entailment violation, reported in (12) as a lower ranked OO-F. The two variants of OO-F (V-LENGTH) can now be deployed in standard OT fashion to correctly obtain the optimal candidates, as in (13) and (14) below. Shaded cells express irrelevance of either OO-F variant.

(13)

rémedy	MWF	MC	OO-F $(\text{V-LENGTH})_1$	*CiV*-Length	OO-F $(\text{V-LENGTH})_2$
a. rémediable	*			*	
b. remédiable		*		*	
c. ☞ remé:diable		*			*

In (13), candidate (*a*) satisfies the Metrical Consistency of (2) above, also a form of OO-F, but only by violating top-ranked Metrical Well-Formedness for the reasons discussed in connection with (4) above, and is thus not optimal. Candidate (*b*) is also not optimal because, once stress is shifted, only OO-F2 is in force, and that is lower-ranked than the constraint imposing lengthening in the context '__CiV'. Candidate (*c*) is thus optimal. Note that, just like OO-F (V-LENGTH), so MC = OO-F (STRESS) should be given in different variants, because equally modulated by distance. In particular, V-lengthened candidate (13*c*) should be dealt with by a lower-ranked MC. This would only improve its status, however, and hence not change the conclusions. Consider now (14).

(14)

lévy	MWF	MC	OO-F $(\text{V-LENGTH})_1$	CiV-Length	OO-F $(\text{V-LENGTH})_2$
a. ☞ léviable				*	
b. lé:viable			*		

Here, the winning candidate satisfies both MWF and MC, under the parse (lé.vi.a)<ble> of the *PES* analysis. OO-F1 is thus in force, and candidate (*b*) is correctly excluded under the given ranking.

In sum, the REH (9) directly yields the attraction over distance effect that we earlier

found to be pervasively present. The latter is directly relatable to principles of neural computation and in fact coincides with the claim that representations are 'distributed'. The reason is that taking an entailment A ⇒ B due to a representation R1 to be relevant also to some other representation R2 presupposes that the A and B of R1 and those of R2 are the same entities, and therefore that representations such as R1 and R2 are 'distributed' over a general pool of units that include A and B, rather than being autonomous or local. The foregoing discussion has provided a partial answer to question (3*a*) above on the sources of OO-F effects: the latter effects obtain between representations that are independently similar. However, in the case of allomorphy, the independent similarity has a specific source that we still need to deal with. For example, in the case of *remediable*, the independent similarity with *remedy* is partly a semantic fact: the intended meaning is 'able to be remedied', not 'able to be read', or other. But it is also a morphological fact. It is a general property of the affix *-able* that it attaches to verbs; hence the presence of *-able* enforces some identity relation between its stem and some verb in the lexicon. Our residual question is then how could this be expressed within the framework of the REH (9), or to put it differently, what is the role of morphology in a system that sees words as related to one another 'surface to surface' in the ways prescribed by the REH. This is in fact question (3*b*) above, and the topic of s. 4.6. Section 4.5 next takes us on a brief detour to address some of the more pressing questions for the REH.

4.5 WHY SOME ENTAILMENTS MAY BE MORE EQUAL THAN OTHERS

One of the questions that arises naturally for the above discussion is what the dimensions of the relevant space are, namely what A, B, C, D of (10)–(12) may be allowed to refer to. The answer is that, at least as a first approximation, the nature of the dimensions involved is orthogonal to the present claim, which merely addresses issues of the general architecture of the system. For present purposes one may well assume decomposition of sound structure in terms of the usual distinctive features or other traditional notions, although there are significant inadequacies to the traditional distinctive features, stemming from the now established role of perceptual distance. For instance, the common rounding of back vowels cannot be adequately accounted for by simply allowing distance between front and back vowels to compound over the two standard features of backness and rounding, since this would predict the same distance if front rather than back vowels were rounded. The inadequacy is in the almost purely articulatory basis of the standard distinctive features (as noted by Boersma 1998). What predicts that rounding should go with back vowels is their acoustic and hence perceptual properties, related to length of the front cavity, already greater with back vowels and hence suitably enhanced by rounding (Lindblom 1986). Similarly, the same feature

of voicing measures different perceptual distances depending on whether it pertains to an onset or a coda and is hence in itself inadequate to capture the role of perception. I must put these issues concerning the proper decomposition of the perceptual space aside, along with the corresponding issues concerning the decomposition of the semantic space. The present claim is merely that there is a notion of global distance to which all dimensions can in principle contribute.

A second question is the apparent commitment to uniform interaction among all components of a representation. For example, given a word of a certain length, the above discussion may suggest that the first segment in the word could equally entail the presence of the segment that immediately follows it and that of the very last segment in the word. However, it is not part of the present claim that all components of a complex representation must necessarily entail each other equally. For one thing, I will argue below that the strength of individual entailments can compound over different representations, so that ultimately relative strength of entailments will come to reflect the overall structure of one's mental lexicon or experience. At the same time, there is also no doubt that the system of mental representation has inherent biases, as reflected by language universals, and this will translate in certain entailment structures being privileged over others. The present claim is—again—only that general associative mechanisms must be available, not that they are sufficient. Mapping out the finer structure of the representational space will be an additional task for future work, perhaps different than it might have been previously, but not completely. There is, however, one general prediction that the present approach makes, which provides an important second approximation refinement of the REH (9). This is the prediction that in general the strength of an entailment relation between two co-occurring components A and B will depend on their own degree of similarity, as I discuss next.

Consider first that given any two components of the same representation A, B, they as well as a resulting entailment may be analyzed into sub-parts, as in (15).

(15) From entailments to sub-entailments for a representation A, B

I. Components	II. Sub-components	III. Entailment	IV. Sub-entailments
A	w, x	$A \Rightarrow B$	$w \Rightarrow y$; $w \Rightarrow z$
B	y, z		$x \Rightarrow y$; $x \Rightarrow z$

In (15), supposing that components A, B of (I) are analyzable into their sub-components *w*, *x* and *y*, *z* respectively, as in (II), then the entailment $A \Rightarrow B$ of (III) will correspondingly be analyzable into the four sub-entailments shown in (IV). That is to say, if A entails B, then presumably each sub-component of A: *w*, *x* will entail each sub-component of B: *y*, *z*, as in (IV). Now suppose further that, beside being part of the same representation, A and B are themselves not unrelated, but are rather partially similar by virtue of $x = y$. Then (16) will hold.

The internal structures of A, B assumed in (15) will give rise to the respective internal entailments in (16.I) and their mirror images, which I put aside. However, given the identity of *x* and *y*, this will also result in the entailments labeled 'transitivity effect' in (16.II), namely if *w* entails *x*, and if $x = y$, then *w* entails *y*, etc. Such transitivity effect

(16) Sub-entailments for a representation A, B when A, B are neighbors

I. Internal entailments		II. Transitivity effect (x = y)	III. Sub-entailments, overall
A (w, x)	B (y, z)	$w \Rightarrow y; x \Rightarrow z$	$(w \Rightarrow y)^2; w \Rightarrow z$
$w \Rightarrow x$	$y \Rightarrow z$		$x \Rightarrow y; (x \Rightarrow z)^2$

results because A and B are neighbors, and is in fact a formal analog of what is referred to as a 'neighborhood effect' in psycholinguistics, describing the fact that representations which are active cause their neighbors also to be partially active. Within the present discussion, this follows from their entailment structure. Neighboring representations share some components by definition, and hence the entailments produced by one representation will partially instantiate the other representation as well. We have noted that, for this reason, neighboring representations are in an attraction relation, which may sometimes result in their neutralization, but we are abstracting away from that here. What we are focusing on is rather that the sub-entailments from the transitivity/neighboring effect in (16.II) will now compound with those of (15.IV) that analyze the more general case, giving rise to the overall structure of sub-entailments in (16), where the exponents refer to multiple instantiations of the same entailment. In other words, while—under the REH—two components A and B will entail one another by being part of the same representation, they will also entail one another to the extent that they are neighbors. The conclusion is then that not all components of the same representation will entail each other equally. Rather, they will do so in a way that is further enhanced by proximity. One can see a number of more specific consequences in this connection, some of which I develop next.

Consider a representation like (17) where the sequential or temporal scale is characterized schematically by a system of *n* binary features, where *n* is the number of phoneme-sized time units minus 1.

(17) Entailments under sequential proximity:

Phonemes	i	n	e	p	t
Time elapsed	0	1	2	3	4
Time Vector	¬1	1	1	1	1
	¬2	¬2	2	2	2
	¬3	¬3	¬3	3	3
	¬4	¬4	¬4	¬4	4

What the above discussion implies for a representation like (17) is that, while *i* co-occurs with both *n* and *t*, the entailment $i \Rightarrow n$ will be stronger than the entailment $i \Rightarrow t$. The reason is that unlike *i* and *t*, *i* and *n* have maximally similar time representations or vectors by virtue of being sequential neighbors, respectively the vectors:

(¬1, ¬2, ¬3, ¬4) and (1, ¬2, ¬3, ¬4). These three phonemes *i*, *n*, *t* are compared in (18), where 'Time'' refers to the portion of the time vector over which *i* and *n* are identical, namely the last three components, and 'Other' refers to the rest of each representation: the first component of the time vector and the structure of each phoneme in space (let us say in terms of distinctive features).

(18) Spatio-temporal analysis:

Phonemes	i	n	t
Other	O-i	O-n	O-t
Time'	T'-i: ¬2, ¬3, ¬4	T'-n: ¬2, ¬3, ¬4	T'-t: 2, 3, 4

The two entailments $i \Rightarrow n$ and $i \Rightarrow t$ due to co-occurrence of these phonemes in the structure [inept] will now map into sub-entailments in the manner of (19) under the criteria of (15) above.

(19) From entailments to sub-entailments

$i \Rightarrow n$ O-i $\Rightarrow$ O-n; (O-i $\Rightarrow$ T'-n)
(T'-i $\Rightarrow$ O-n); T'-i $\Rightarrow$ T'-n

$i \Rightarrow t$ O-i $\Rightarrow$ O-t; O-i $\Rightarrow$ T'-t
T'-i $\Rightarrow$ O-t; T'-i $\Rightarrow$ T'-t

However, the two entailments parenthesized in (19) will be effectively boosted by the spatio-temporal proximity of *i* and *n*, not paralleled by the proximity of *i* and *t*. The reason is the transitivity effect in (20), an instance of the effect in (16) above.

(20) Sub-entailments added by representational overlap

Internal entailments		Transitivity Effect
i: O-i, T'-i	n: T'n, O-n	due to: T'-i = T'-n
O-i $\Rightarrow$ T'-i	T'-n $\Rightarrow$ O-n	O-i $\Rightarrow$ T'-n; T'-i $\Rightarrow$ O-n

That is, given their internal structures, *i* and *n* will generate the internal entailments given, respectively. However, given that *T'-i* is identical to *T'-n* as shown in (18), those entailments will translate into the ones in the rightmost column in (20), respectively, exactly identical to the parenthesized ones in (19), which are thereby boosted. No such boosting effect will occur for *t* since *i* and *t* do not share any temporal structure within the time frame considered. If the frame were extended, e.g. to instants 5, 6, etc. some sharing would occur, but always less than occurs between *i* and *n*. Hence both components of *n*, *O-n* and *T'n*, are entailed by *i* in virtue of the simple co-occurrence of *i* and *n* in the same structure [inept] (the sub-entailments in (19)), but they are both additionally entailed in virtue of the sequential proximity of *i* and *n*.

What the foregoing means effectively is that segments should be better able to 'bind' together in recurrent patterns when they are contiguous—exactly the generalization that describes 'concatenative' morphology. We are thus shedding light on the non-existence in languages such as English of discontinuous morphemes like a hypothetical *in-. . .-y*, such that for example, given *accurate*, *in-accurac-y* would just mean

'inaccurate'. A well-known alternative to concatenative morphology is of course the Semitic-type 'root and pattern' morphology (McCarthy 1981). From the present point of view this would in fact involve an alternative type of binding also expected, in the way illustrated in (21).

(21) Semitic morphology (Arabic)
katab-a *'he wrote'*
kutib-a *'it was written'*
aktub-a *'he was writing'*
. . .

Phonemes	k	a	t	a	b	a
Other	O-k	O-a	O-t	O-a	O-b	O-a
Major Class [+ conson.]	+C	-C	+C	-C	+C	-C

The case in (21) is simply the spatial counterpart to the temporally based (18), as the segments that bind together here are spatial rather than sequential neighbors. In particular, the member of the root segmentism *ktb* share the major class feature [+consonantal] and are thus expected in (21) to tie to each other more strongly than to the vowels by parity of reasoning with (18)–(20).[2] The conditions under which elements can bind together in morphological systems—spatio-temporal proximity—are thus predicted to be very much analogous to those that underlie the phonological attraction relations manifested in assimilatory and dissimilatory phenomena from this point of view.[3]

To recap, while in general clustering effects seem able to implicate unrelated dimensions, there is also reason to believe that co-occurring elements that are sequential or spatial neighbors cluster more strongly. The required refinement of the REH appears to involve simply expanding components into sub-components, and correspondingly entailments into sub-entailments. The sub-entailments yield neighborhood effects that overlay the basic co-occurrence effect under the same basic principle of the REH, thus accounting for the stronger clustering or binding of neighbors.

[2] It also seems less than an accident from this point of view that the only sequentially discontinuous morpheme in English is constituted of segmentally identical parts as in *en-light-en*, etc.

[3] The binding together of elements under spatio-temporal proximity thus predicted by their entailment structure is also likely to cover effects formerly addressed by the 'Feature Geometry' program (Clements 1985; Sagey 1986). While features tend to cluster in assimilatory processes for the reasons discussed earlier, features sharing certain subspaces, like the features [±voice], [±spread glottis], both pertaining to the same articulator: the glottis, appear to cluster more systematically. This would follow from a representation along the lines of (i) for an aspirated [t].

(i) Featural Analysis, [t^{h}]

Features	[-voice]	[+spread gl.]	[-anterior]	[-contin.]
Feature Class	Glottal	Glottal	Coronal	

The shared feature class [Glottal] predicts stronger binding of the elements involved just as in the previous cases of shared elements, thus accounting for such complex assimilations as /b t^{h}/ → [p^{h}t^{h}] (Ancient Greek, Kenstowicz 1994: 156 f. and refs.). Note as well that, because entailments are violable, we predict that clustering by feature class should be violable, just as independently argued in Padgett (2002), and contrary to the rigid clustering predicted by feature trees.

4.6 WHAT TO DO WITH MORPHOLOGY

4.6.1 How affixes encode OO-Faith relations

We return here to question (3*b*) above concerning the relationship between OO-F constraints and morphology. A fundamental characteristic of the way in which affixes combine with stems is that they do so violably, as shown in (22).

(22)

		Phonological alternations	
	I. Morphological and semantic irregularity	II. Contrastive variation	III. Allophonic variation
a.	compUlS-ive, derivAT-ive	der[i]vAT-ive	
b.	regrett-able, INEVIT-able	adm[i]r-able	
c.	symptomAT-ic	paras[i]t-ic	parasi[ɾ]-ic
d.			permi[ɾ]-ing

In (22), the suffixes *-ive*, *-able*, *-ing* all select verbs as their stems, but rather than the forms in (22), the actual verbs are: *compel* (*a*), *der*[*ay*]*ve* (*a*), *adm*[*ay*]*re* (*b*), and *permi*[*t*] (*d*). In addition, there is no verb (*in*)*evit* (*b*), and *regrettable* (*b*) does not have the expected meaning 'able to be regretted', but rather refers to things that *must* be regretted instead. Similarly, *-ic* selects nouns, but the actual nouns are rather *symptom* and *paras*[*ay*]*te* instead of the forms in (22*c*). Such distortions can be divided into the three categories of (22). There are simply irregularities, both of form and meaning, as in column (I), although irregularities may in turn correspond to various degrees of subregularity, e.g.: *symptomAt-ic*, *problemAT-ic*, *axiomAT-ic*, etc. Then there are regular phonological effects, further subdivided into contrastive (column (II)), and allophonic (column (III)). English vowel shortening is a contrastive type of variation because it neutralizes a contrast, that between short vowels and their long or diphthongal counterparts. These contrast because they have non-complementary distributions: [bayt]/[bɪt], etc. Instead, American English flapping is allophonic as the alveolar flap [ɾ] is in complementary distribution with regular alveolar stops [t/d]. Each individual suffix appears to control the general degree to which its stem can fail to match an independent word in the lexicon. The affixes that were once classified as 'Level 1' in 'Lexical Phonology' (Kiparsky 1982*a*, *b*) and exemplified in (22*a*, *b*, *c*) tolerate mismatches rather well, hence featuring both sheer irregularities and regular phonological alternations. In contrast, those that were classified as Level 2, such as *-less*, *-ness*, *-ful*, and *-ing* of (22*d*) tolerate few mismatches, essentially only those coming from allophonic variation, such as the one shown in (22).

Such violability of the combinatorial properties of affixes would follow if the latter were types of entailments derivable from the REH (9), since we have seen that those

are generally violable. This view can in fact be upheld straightforwardly, as we see by considering (23).

(23)	LEXICON:	Entailments I	Entailments II	Entailments III
	preventive	ive ⇒ /prevent___	prevent ⇒ V	ive ⇒ /V ___
	generative	ive ⇒ /generat___	generate ⇒ V	ive ⇒ /V ___
	Summation:	noise	___	**ive ⇒ /V ___**

Under the REH, a morphologically complex item like *prevent-ive* will have, as part of its entailment structure, some entailments to the effect that, if there is a sequence *ive*, there must be a sequence *prevent* preceding it, and similarly for *generat-ive*, as indicated in column (I). By themselves, entailments such as those of column (I) would be largely discordant, resulting in no generalization when they are conjoined. Consider, however, that besides containing the items *preventive* and *generative*, the mental lexicon also contains *prevent* and *generate*, both of which are verbs. The specific internal structure of *prevent* will now contain the entailment of column (II), namely: if there is a sound structure [prɪvent], the associated category is 'verb', where lexical category is presumably part of semantic representation. In combination with the entailment from column (I), this new entailment will result in the entailment of column (III) by transitivity, namely: if *ive* entails *prevent*, and if *prevent* is a verb, then *ive* entails a verb. Similarly for *generat-ive*, despite the difference between the stem *gener*[ə]*t-* and the verb *gener*[*ey*]*te*: the two structures are still sufficiently similar to ensure that most of the entailments abbreviated by the single conditional of column (I) will be relevant to column (II). The entailment of column (III) will thus still hold substantially in the case of *generative*, though presumably in a somewhat weakened form. Now, unlike the entailments of column (I), those of column (III) are concurrent, thus yielding a higher-order entailment or generalization by summation over the lexicon. Such higher-order entailments are effectively subcategorization frames for affixes, expressing their general selectional or combinatorial properties. I will henceforth refer to them as 'SELECT' entailments. They differ ontologically from the subcategorization frames or the word-formation rules of previous approaches to morphology by not being part of an extrinsically given grammar, but being rather emergent properties of the lexicon itself—a reflex of the self-organization resulting from the REH. Furthermore, they differ from the machinery of those approaches also empirically, by virtue of being violable—a crucial difference as we see below.

We are now in a position to provide a full analysis of the OT notion of OO-F, the driving force behind PU. The latter is the result of the combination of two different types of entailments, those due to the GA effect, and those due to the SELECT effect, in turn both reflexes of the REH. The respective contributions of these two effects are underscored by the fact that, on the one hand, SELECT effects are not always present. For instance, in the French cases in (5) and (8), the liaison form is not related to the feminine allomorph via affixation, hence no SELECT entailments are expected. The case PU over affixes, where affixal rather than stem material is kept invariant, also does not depend on the presence of further affixes, as we see in the next section. In those

cases, GA will go it alone. On the other hand, when SELECT effects are present, they have the ability to substantially modulate the resulting OO-F/PU effect. We have noted with regard to (22) above that with the Level 1 affixes the OO-F imposed on the stems is generally weak, to the point that there need not be an independent base at all: *inevit-able*/**inevit*, in contrast to the Level 2 affixes that impose a strong identity requirement. The GA effect strictly depends on global distance, and it is thus only the SELECT entailments that could be responsible for the difference. There are two different ways in which SELECT entailments can obtain different strengths. One is numerical. If such entailments sum up over instances as in (23.III), then the number of instances in the lexicon (roughly, what in psycholinguistics is referred to as the 'type frequency', i.e. the lexical frequency of that type of morphological formation) will be a determinant of strength. The other factor is what we may call 'seeded idiosyncrasy'. If a set of morphologically derived forms comes into the language with a high degree of idiosyncrasy, as is surely the case for the English Latinate (Level 1) lexicon in general, then it will give rise to weaker SELECT entailments. For instance a form like *inevit-able* will not contribute—but will in fact contribute negatively—to the entailment that *-able* must select a verb. Intermediately idiosyncratic forms will correspondingly contribute to SELECT entailments only weakly. On the present approach, morphological regularity has a self-feeding character since regular forms recruit others into regularity via the SELECT entailments they generate. The English Level 1 lexicon could thus be in the current equilibrium through a combination of relatively low numbers and historical accident, remaining below a threshold that may lead to the almost complete regularity of the Level 2 lexicon once crossed.

While an outgrowth of the approach independently developed in *PES*, the present perspective converges substantially with Bybee's (1988, 1995). Bybee's notion that there exists a network of surface-to-surface connections whose strength depends on the degree of similarity between items has a direct counterpart in the present claim that representations influence each other by means of the entailments, an influence that is expected to depend on their degree of similarity (GA). Bybee's further notion that sets of parallel connections form 'schemas' that can generate new items finds an analog in the SELECT entailments of (23.III), and the contribution of 'type frequency' to the strength of a schema postulated by Bybee is captured here by the summation of such entailments over the lexicon.

Returning to (22), given that the SELECT entailments (23.III) and the resulting OO-F effect are violated with the Level 1 affixes, the question is what exactly compels those violations. The phonological alternations in (22.II, III), as in *adm*[ɪ]*r-able*, *permi*[ɹ]*-ing*, etc., are obviously due to markedness constraints of the phonology. As for the morphological and semantic irregularities of (22.I), as in *compUlS-ive*, etc., I have argued in Burzio (2000*a*, 2002*a*) that the OT notion of Input–Output faithfulness (IO-F) is relevant in this connection. Once morphological relations cease to be expressed by a common input in UR and are rather handled by OO-F (= SELECT and GA), the notion of input can then be utilized with morphologically complex words to express their idiosyncrasies, just as with morphologically simple ones. Morphological

irregularity is then controlled by the relative ranking of OO-F and IO-F. Dominant OO-F yields morphological regularity—the Level 2 syndrome exemplified by affix *-hing* of (22*d*), while dominant IO-F enables irregularity to exist: the Level 1 syndrome of (22*a–c*), as shown in (24).

(24) *a.* Morphological irregularity (Level 1 Affixes)

Input: /compUlS-ive/ Base: */compell/*	IO-F	OO-F
a. compell-ive	*	
b. ☞ compuls-ive		*

b. Morphological regularity (Level 2 Affixes)

Input: /compUlS-ing/ Base: */compell/*	OO-F	IO-F
a. ☞ compell-ing		*
b. compuls-ing	*	

On this approach, morphological irregularity receives an analysis within OT which parallels that of phonological markedness. The latter is also a form of 'irregularity', arising when markedness constraints—the source of regularity—are dominated by IO-F, as shown in (25).

(25) *a.* Phonological irregularity (Marked value, English)

Input: /cri:me/	IO-F (V-length)	*V:
a. crime	*	
b. ☞ cri:me		*

b. Phonological regularity (Unmarked value, Italian)

Input: /cri:mine/	*V:	IO-F (V-length)
a. cri:mine	*	
b. ☞ crimine		*

On this analysis, IO-F constraints are thus the sole source of irregularity, both phonological and morphological.

Note that in the present approach, in which there are only surface forms, the 'input' can no longer be thought of as constituting a separate representation like the traditional UR and thus needs to be reconceptualized. This consequence obtains in significant ways in Prince and Smolensky's (1993) formulation as well. Under their notion

of 'Lexicon Optimization', in morphologically simple items like [əmérɪkə] input and output representations would be identical, a claim equivalently statable as the present claim that there is only a surface form and no UR. Under Lexicon Optimization, such a single representation must be thought of as having the internal partition depicted in (26).

(26) Surface forms under 'Lexicon Optimization' (for morphologically simple words)

P_I (dominant IO-F) →

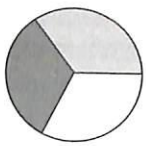
← P_M (dominant Markedness)

That is to say, in the single surface form [əmérɪkə], some properties, like the specific choice of segments, will result from input information. Let us call these 'P_I'. Other properties on the other hand, like the stress and syllabification, will result from Markedness constraints. Let us call these properties 'P_M'. In the Lexicon Optimization approach, the P_I/P_M distinction is not reflected by respective distributions over different levels, but rather by constraint ranking. For any aspect of the representation α, P_I^{α} obtains iff IO-F(α) ≫ Markedness(α), while P_M^{α} obtains iff the opposite ranking holds. Hence, the long vowel of *cri:me* (25) above is a type of P_I, while the short one of Italian *crimine* (25*b*) is a type of P_M. So, while under Lexicon Optimization the input to [əmérɪkə] is the same full-fledged [əmérɪkə], only some if this input—what I have called P_I, is effectively 'active' in determining the output. The rest of it is passive, under the sway of the Markedness constraints—what I have called P_M.

Prince and Smolensky do not extend Lexicon Optimization to morphologically complex words, because they adopt the traditional approach to allomorphy, in which morphemes have a unique form in UR. Since these forms cannot be simultaneously identical to all of their surface reflexes, Lexicon Optimization becomes inapplicable. The present approach rejects UR and by doing so can maintain the arguably more natural claim that Lexicon Optimization is always true. The only difference between simple and complex words in the present perspective is that the latter invoke one additional set of constraints: OO-F constraints, hence turning the bi-partition of (26) into the tri-partition of (27).

(27) Surface forms in the Output–Output Faithfulness approach

P_I (dominant IO-F) →



← P_O (dominant OO-F)

← P_M (dominant Markedness)

The new set of properties P_O will represent the aspects of the representation that are determined by OO-F constraints by virtue of the constraint ranking. As in the case of [əmérɪkə], the input to *compulsive* can be taken to be the exact same form, but only

some of it will be effectively active. In (24) above this part of the input was given in caps. The rest of the input conforms with OO-F, just as the stress of [əmérɪkə] conforms with Markedness. I return below to the question of why only a fragment of the input */compUlSI* (the portions in caps), is active in this fashion (i.e. corresponds to P_1 of (27)). Now in (24*b*) above the same input *compUlS-* was given for the form *compelling*, but in that case it was only as a hypothetical input, to show that it would have no effect.

In sum, we take surface forms to be evaluated by three classes of constraints simultaneously. IO-F and Markedness constraints work much as in standard OT, while OO-F constraints take over the role of traditional morphology, resulting in a fully integrated parallel system. This move makes UR dispensable not only with underived items (as with Prince and Smolensky's 'Lexicon Optimization'), but altogether. In turn, this makes it possible for IO-F constraints to be utilized to express morphological irregularity similarly to the way in which they standardly express phonological markedness.

4.6.2 Interfacing phonology and morphology

This approach, in which affixes are instrumental in modulating the rank of OO-F constraints as in (24) above, has the unique ability to account for the important correlation that exists between morphological and phonological regularities. As noted in connection with (22) above and discussed in Burzio (2002*a*) those regularities correlate inversely. The sectors of the lexicon that abound in morphological irregularity, such as that Level 1 affixes are rich in certain phonological alternations which are perfectly well-expected and hence 'regular', such as re-stressing: *párent/parént-al*, re-syllabification: *in.ter.ve:.nø/in.ter.ven.tio.nø* (continuing to assume the final null nuclei of the *PES* analysis), and vowel shortening: *parasitic*, *criminal*, etc. In contrast, items featuring Level 2 affixes, which are exempt from morphological irregularity, exhibit few of the expected phonological alternations, and are in this respect highly irregular phonologically, e.g. *éffortless* (exceptional stress), *discreetness* (exceptional syllable *creet*), *cri:melessness* (no shortening). The same is true for the much debated English past tense. The morphologically regular class is systematically irregular phonologically, by way of otherwise unattested syllables like that of [biyp.tø] (beeped), while the morphologically irregular class features straightforward phonological regularities, like the expected shortening of [kiy.pø]/[kep.tø]. Essentially, the only kind of phonological alternation that obtains within the Level 2 lexicon is the allophonic kind, as in (Am. English) *permit /permi*[ɾ]*ing* (22*d*) above.[4]

[4] Note that there is sense of 'regularity' in which Level 2 items may seem the more regular, contrary to the text claim. That sense is conveyed by the observation that the phonological processes that do apply to the Level 2 lexicon, like the allophonic one just noted are exceptionless and/or 'automatic', in contrast to Level 1 processes like shortening, which have 'lexical' exceptions, like *obe:sity*. I will return to this fact below.

The correlation in question follows from considering the rankings that were invoked to deal with the distribution of morphological regularity, repeated in (28).

(28) Morphologically motivated rankings:

		Morphological Irregularity: (24*a*)		
OO-F2	≫	IO-F	≫	OO-F1
Morphological Regularity: (24*b*)				

Given these rankings, any Markedness constraint of the Phonology will always have a greater chance of being dominated by OO-F_2 than by OO-F_1, and never the opposite. Since Markedness constraints are the source of phonological regularity, it follows that Level 2 items, controlled by OO-F_2, will be phonologically less regular than Level 1 items, whence the noted correlation.

The rankings in (28) also predict two additional effects related to the relative ranking of IO-F constraints. One is that Level 2 items (controlled by OO-F_2) should tend to be phonologically more marked than underived items (controlled by IO-F), as is surely correct: there are no syllables like that of [biyp.tø] among underived items, or stress patterns like that of *éffortless(ness)*. The second expected effect is that phonological processes that may work over Level 1 items (controlled by OO-F_1) may 'block' in underived items (controlled by IO-F). This is exactly the case with vowel shortening, as shown in (29) in contrast to (25) *cri:me* above.

(29) Vowel shortening with Level 1 items

Base: cri:me			*V:	OO-F1 (V-length)
a.	cri:minal	cri:min-al	*	
b. ☞	criminal	crimin-al		*

Note that the case in (29) is not just an instance of the 'tri-syllabic' shortening, which would naturally not be applicable to *cri:me*. As argued in Burzio (1993, 2000*a*) and *PES*, tri-syllabic shortening is a purely descriptive category, of no theoretical significance. Shortening applies to all positions within Level 1 items in principle (including final syllables: *infinite*), the constraint at work being the maximally general one of (29), which is independently motivated by the difference between English and Italian of (25) above. The English Level 1 lexicon is in fact very much like the whole of Italian, favoring short vowels. Vowels may turn up long under certain circumstances, however, in stressed open penultimates in particular, as in *desí:rous*. This is because of competition with the Metrical Consistency (MC) of (2) above. Given a prohibition on stressed light penultimates (the foot *(L σ), banned by the Metrical Well-Formedness of (2)), shortening would force a violation of MC, as in *desí:rel *désirous*. Even this effect is paralleled by Italian, where stressed open penultimates undergo lengthening: *sapó:ne*

'soap'. The only difference is the source of stress: IO-F in Italian, OO-F in the English case. However, it turns out that MC does not quite dominate *V:, but in fact essentially ties with it, hence permitting variation as in *desí:rous* (MC satisfied)/*blásphemous* (*V: satisfied. Cf. *blasphé:me*), a point to which I return below. On the other hand, antepenultimate stress does not require a heavy syllable (witness *ámerica*, etc.), whence the much greater regularity of shortening in the 'tri-syllabic' cases—*criminal*, *divinity*, *natural*, etc.—in which both MC and *V: are satisfied simultaneously. Other environments also yield regular shortening for comparable reasons, as argued in the references.

Consider further that, under the ranking in (28), any process that blocks in morphologically underived items, like vowel shortening, is guaranteed also to block in Level 2 items for the obvious reason that any Markedness constraint dominated by IO-F, will also be dominated by OO-F_2 by transitivity. We have already seen that this prediction is indeed correct for shortening: *cri:melessness*, etc.

In sum, the complex distribution of English vowel shortening, given in (30), follows from the postulated constraint rankings.

(30)	Domain	English Vowel shortening		Ranking responsible
a.	Underived	no:	parasi:te; di:nosaur; cri:me	IO-F ≫ *V:
b.	Level 1 (some contexts)	yes:	parasit-ic; crimin-al, divin-ity	*V: ≫ OO-F1
b′.	Level 1 (other contexts)	yes/no:	blasphem-ous, desi:r-ous	*V: ≫/<< MC
c.	Level 2 derived	no:	cri:me-less-ness	OO-F2 ≫ *V:

I will return below to the fact that while OO-F_1 (V-length) is outranked by *V: as in (30*b*), MC, also a form of OO-F_1, only ties with *V:, as in (30*b′*).

In contrast to vowel shortening, consider now allophonic variation, in general attributed in OT to the schema in (31) (Kirchner 1997, Burzio 2000*a*) which is applied here to the case of American English flapping of (22.III) above.

(31) Allophonic variation in OT (e.g. Amer. English Alveolar Flap):

Markedness, contextual	≫	Markedness, general	≫	IO-F
*t, d/V́___V	≫	*Flap	≫	Ident (flap)

Putting aside the exact featural analysis of the alveolar flap ([+sonorant] according to Kenstowicz 1994: 68), this sound is in this analysis excluded in general by an a-contextual markedness constraint that dominates IO-F, but is then mandated by a contextual markedness constraint that dominates the rest of the schema in turn, as in (31). The latter constraint excludes regular coronal stops *t*, *d*, between a stressed vowel and another vowel (further ranking among faithfulness constraints, not given in (31),

would also be required to exclude alternative repairs). Consider now what our ranking in (28) above predicts in this connection. With morphologically derived items, OO-F will replace IO-F in (31) in the present system. With $OO\text{-}F_1$, it is clear that the schema will remain unaffected (since $OO\text{-}F_1$ is dominated by IO-F). This correctly guarantees that allophonic variation will never distinguish between underived and Level 1-derived items as vowel shortening does. Hence, the well-recognized fact that allophonic variation never 'blocks' in underived environments (Kiparsky 1993 and refs.) follows automatically from this analysis. With $OO\text{-}F_2$, however, there is no guarantee that the allophonic variation of underived items will still obtain, but there is also no guarantee that it will disappear. All we know is that the ranking '$OO\text{-}F_2 \gg$ IO-F' holds, so that $OO\text{-}F_2$ may or may not end up dominated in (31). Thus the observed fact that flapping has the same distribution with Level 2 items as with all other items is consistent with the present analysis. Other types of allophonic variation do disappear at Level 2, however, as is also consistent with the analysis. For instance, the variation between syllabic [l̩] and non-syllabic [l] of *cyc* [l̩] (cycle)/*cyc*[*l*]*ic* blocks with Level 2 *-ing* in some dialects, that thus feature OO-Faithful *cyc*[*l̩*]*ing*, rather than the regularly allophonic *cyc*[*l*]*ing* of other dialects (see also Burzio 2002*a*).

In sum, the relative ranking of IO-F and OO-F constraints motivated by morphological generalizations automatically yields significant phonological generalizations as well. Specifically, it yields an implicational hierarchy 'Level 2 $\Rightarrow$ Underived $\Rightarrow$ Level 1', such that any process applying to Level 2 items will also apply everywhere else as well, while any process applying to underived items will also apply to Level 1 items, though not necessarily the other way around in each case. It also correctly derives the well-established generalization that if a process is allophonic it will never block in underived environments.

4.6.3 Finer-grained interactions

The foregoing discussion has argued that OO-F reduces to the more primitive notion of Representational Entailment. It appears that the OT notion of IO-F can also be so reduced. On the present view, an input representation is a bundle of entailments potentially in competition with other entailments coming from the rest of the lexicon. In particular, entailments that are multiply instantiated and thus represent cross-lexical generalizations will constitute a source of pressure on an input representation. Such cross-lexical generalizations come in two basic classes in the present system: phonological markedness constraints, and OO-F. To address the interaction between such generalizations and an individual representation, we must consider that generalizations are by definition always lower-dimensional than individual representations. Thus compare a hypothetical generalization 'A $\Rightarrow$ B' with a representation 'A, $\neg$B, C, D' which violates it, as in (32).

(32)

	Representation: A, $\neg$B, C, D	Generalization:
Entailments (B/$\neg$B):	A $\Rightarrow \neg$B; C $\Rightarrow \neg$B; D $\Rightarrow \neg$B	A $\Rightarrow$ B

In (32), there is one entailment favoring B due to the generalization, but there are three entailments favoring ¬B from the representation itself, a clear advantage. Higher dimensionality is thus the edge that enables input representations to output faithfully, violating existing generalizations. The generalization in (32) could be, for example, the SELECT entailment 'IVE ⇒ /V ___' of (23) above. We know that this entailment is shorthand for a larger set of entailments, since *ive* is a complex representation. However, a still larger set is represented by the competing entailment 'ive ⇒ /compuls___' (effectively: 'when combining with the meaning 'compel', *-ive* is preceded by *compuls*). This is because the structure *compuls* with its associated meaning is more complex than just 'V'. At the same time, some of the entailments corresponding to *compulsive* also benefit from other similar representations, like *repulsive*, *impulsive*, etc., further assisting the irregular outcome, and accounting for the noted fact that lexical irregularities are often subregularities.

While lower-dimensional, generalizations have their own numerical advantage, however, in the fact just noted they are multiply instantiated. Hence, in principle, the competition between the general and the specific can have either outcome, the former benefiting from its multiple instantiations, while the latter benefits from its higher dimensionality which results in a larger number of entailments validating each component of the representation. Individual outcomes will depend on the actual numbers that enter into the entailment calculus. While we are not in a position to deal with such actual numbers at present, we have reviewed some of the statistical reasons that would lead to different outcomes in the Level 1 and the Level 2 domains, with individual idiosyncrasy prevailing in the former, while regularity prevails in the latter.

Our discussion has thus far not accounted for the fact that morphological idiosyncrasy within the Level 1 lexicon is not totally unrestrained. If the dominance relation 'IO-F ≫OO-F' postulated above were categorical, with an input representation always prevailing over the demands of morphology, then OO-F should have no effect at all, and any resemblance between affixed stems and independent lexical items should arise only by sheer accident, which is surely incorrect. The remedy to this is in the fact that the basis to OO-F constraints is not only in the SELECT entailments, which may have a fixed rank, but also in the GA effect, which modulates the overall OO-F effect in accordance with representational distance. The revised prediction is, then, that for regions that are very close to an attractor, OO-F may well dominate IO-F even in the Level 1 domain, while for more remote regions the opposite ranking would hold as assumed above. While we are also not in a position to verify this prediction with a high level of precision, several facts seem consistent with it. One is the pervasive phenomenon of bound stems, e.g. *arbore-al*. In such cases there is no attractor, except semantically ('arbore' ≈ 'tree'), and OO-F is predicted to be at its minimum, hence interfering minimally with this type of form. In contrast to this and other idiosyncrasies, there are virtually no exceptions to the metrical consistency effects of (2*a'*, *b'*) above in the analysis of *PES*. This would follow from the fact that metrical consistency obtains under substantial segmental identity almost by definition. Specifically, cases of bound stems are irrelevant to metrical consistency, precisely because of the absence of

a base with which to be consistent. Hence, metrical consistency occurs under conditions of maximal attraction. Recall that where segmental identity is breached, metrical consistency tends to fail as well as expected: *lárynx/larýnges*. This then accounts for the relatively higher rank given to MC in (30) compared with the more general OO-F_1. The distribution of vowel shortening is of further relevance in this connection. In the analysis of *PES*, the shortening constraint *V: of (30) is expected to be always satisfied in 'tri-syllabic' environments like *divínity* (30*b*) because it does not interfere with MC, but it is expected to be satisfied only variably in penultimate syllables as in (30*b*'), where it does interfere. However, the tri-syllabic shortening does have exceptions, like *obé:sity*. In Burzio (2000*a*), I argued that these exceptions arise in the same way as the morphological irregularity of *compulsive*, namely by input specification of vowel length. With IO-F dominating *V:, the long vowel of *obé:sity* would be on a par with that of underived *obe:se*. However, exceptions to trisyllabic shortening appear to have particularly transparent semantic relations with their bases, so that an alternative might be to postulate a higher-ranked OO-F relation due to the tight semantic connection. Similarly, the variability of the shortening of penultimate syllables was handled in Burzio (2000*a*) by IO-F. With MC and *V: in a tie, IO-F would be the natural tie breaker. However, once again, non-shortening cases like *desi:rous* seem more palpably related to *desi:re* than shortening *blasphemous* is to *blasphe:me*, making an account in terms of variably-ranked OO-F_1 also a possibility. Ultimately, this issue may be of the 'chicken-and-egg' kind, since we take semantic non-transparency to be a form of IO-F, like morphological non-transparency.[5] In conclusion, the ranking 'IO-F ≫ OO-F' does not seem to obtain categorically within the Level 1 domain. In this regard the earlier discussion was thus simplified, but remains none the less valid to the extent that the latter ranking finds substantial opportunity within Level 1, and none within Level 2.

Summing up this section, the REH (9) enables us to construe the combinatorial principles of morphology as types of entailments and, as such, as violable. Violability is established by two classes of violations—phonologically driven allomorphy and morphological irregularity—both resulting in disuniform paradigms. Paradigm uniformity results when the combinatorial principles (the SELECT entailments of (23. III)) are high-ranked, yielding high-ranked OO-F. Approaches that take morphology to consist of inviolable rules or principles need to deal with the violations in terms of sequential ordering of operations. On those approaches, morphology would be enforced first, yielding an 'underlying' level of representation, and then other machinery would apply next, distorting the morphology's outcome. This traditional conception is falsified by two considerations. First, there is never any good reason from this point of view why morphology should ever be irregular. The various 'readjustment

[5] With regard to the issue of lexical exceptions of n. 4, we are correctly expecting their exclusion with allophonic processes applying to Level 2 items. On the IO-F account, such exceptions will be banned by the fact that IO-F is totally dominated in the allophonic schema (31). On the alternative OO-F account, lexical variation will be excluded by the uniformly high semantic correspondence between Level 2 items and their bases.

rules', 'filters', and other devices proposed in this connection merely state the problem, in so far as they fail to reduce irregularity to independent properties of the system. This contrasts with the present approach, which seeks to reduce irregularity to the interplay of IO-F and OO-F, both independent components of the system. The second consideration is that there will never be any good reason why the two types of phenomena that occur downstream from the morphology—irregularity and phonologically based allomorphy—should have substantially the same distribution: both are present with Level 1 affixation, and both are absent with Level 2 affixation. However one chooses to capture morphological irregularity in that perspective, the relevant subsystem will end up being independent of the one responsible for the phonology. It is only in the present approach, in which morphological generalizations, i.e. OO-F relations, can either be weak or strong, that those two subsystems can have something in common: they are both competing with the morphology. When morphology is weak (low-ranked OO-F), competitors are let in. When it is strong, they are shut out, together. By contrast, in Kiparsky's (1982*a*, *b*) 'Lexical Phonology', for instance, there is a partial account of morphological irregularity in the assumption that the output of each level must be a lexical item. So, while one can have Level 1 *ARBORE-al*, one cannot have Level 2 **ARBORE-less* (but only *tree-less*). The reason is that the input to Level 2 must be the output of Level 1 and hence necessarily a lexical item, excluding a bound stem like *ARBORE-*. Aside from being stipulatory, this assumption is unrelated to the one needed to ensure that most phonological alternations are excluded at Level 2, which is handled by assuming that Level 2 phonology has fewer rules. Recasting matters into the contemporary OT framework will not help. Level 2 phonology would now presumably be characterized by higher-ranked IO-Faithfulness, but that will remain unrelated to the fact that its input cannot be a bound stem. Note, too, that it is not faithfulness in general that is high-ranked at Level 2, but rather just the faithfulness that concerns stems. Thus, contrastive variation is banned quite generally within Level 2 stems, but not within Level 2 affixes, witness *cat*[*s*]/*dog*[*z*], *bribe*[*d*]/*peep*[*t*], etc. The Level 2 generalization does therefore not concern phonology per se, but rather stems, and consists of the requirement that they must equal words, as argued in *PES* (274) to which the present discussion is a sequel.

Phonology is also intractable within the 'dual route' approach to morphology of Pinker (1991) and others. That approach maintains that regular morphology consists of a system of symbolic rules, while irregular morphology reflects an analogical/associational system. There is no reason from that point of view why the phonology should be unrestricted within the analogical route, while within the symbolic route it is highly restricted if it applies to stems, but unrestricted again if it applies to affixes. There seems no principled way to characterize the curious distribution of phonological regularity from that perspective. The notion that there is one 'route' specifically dedicated to regularity only predicts that regular phonology (e.g. regular stress, regular syllabification) should coexist with regular morphology. As we have seen, the actual facts are exactly the opposite (see also Burzio 2002*a*).

4.7 TYPES OF PU

4.7.1 Uniformity of stems

Beside GA, which is triggered by overall similarity, we have seen that the selectional properties of affixes are also an important source of PU. The general schema in (33) is the one instantiated by (23).

(33) Uniformity with a bare base

	Entailments I	Entailments II	Entailments III
X-A	A ⇒ /X___	X ⇒ V	A ⇒ /X___ where X=V
prevent-ive	ive ⇒ /prevent___	prevent ⇒ V	

In (33), X is a stem like *prevent* and A an affix like *-ive*. As we saw above, the entailment (I) will result in the 'SELECT' entailment (III) when combined with the entailment (II) contributed by the lexical item *prevent* itself. This result follows from the transitivity of the entailments. The SELECT entailment is given here in a slightly different format than in (23) for expository reasons. From this perspective, a comparable SELECT effect will also be expected when the 'base' is not unaffixed as in (33), but is rather itself an affixed member of the same paradigm, as in the Spanish case in (1) above. This is shown in (34).

(34) Uniformity with an affixed base

	Entailments I	Entailments II	Entailments III
X-A1	A1 ⇒ /X___		A1 ⇒ /X___
X-A_2		X ⇒ /___ A_2	where X= /___ A_2
amá-ba	ba ⇒ /amá___		
amá-bamos		amá ⇒ /___bamos	

In (34), two affixes A_1 and A_2 share the same stem X. Each affix then entails the presence of the stem next to it just as in (33) and as indicated in column (34.I) for A_1. Entailments generated by co-occurrence are symmetrical, however, so that entailments in the other direction will also obtain, as indicated in column (II) for A_2. The combination of entailment (I) and entailment (II) will yield entailment (III) by transitivity just as before. That is, the X that A_1 is associated with is also systematically associated with A_2, so that A_1 will come to select as its stem a general object whose property it is that A_2 also occurs attached to it (column III). This contextual property of X in (34) is parallel to the property of X in (33) of being a verb when it occurs by itself. The attained strength or rank of the SELECT entailments in both (33) and (34) is expected to depend on class size, that is on how many times entailments (I) and (II) are simultaneously instantiated through the lexicon (roughly related to the notion of 'type

frequency'). This sheds light on the fact that in the overwhelmingly larger Romance first conjugation (It. *-are*; Fr. *-er*; Sp. *-ar*), PU is much more prevalent than in the other, smaller conjugations. For instance the Italian *-ere* conjugation exhibits palatalization alternations like *vo*[*l*]*-ete*/*vo*[ʎʎ]*-iamo* 'want 2PL/1PL', but the *-are* conjugation correspondingly finds only non-alternating *vo*[*l*]*-ate*/*vo*[*l*]*-iamo* 'fly 2PL/1PL'. All other segmental alternations of Italian are correspondingly also missing from the *-are* conjugation. Similarly, in French, the *-ir* conjugation has alternations like *je*[*paʁ*]/*nous* [*paʁt-õ*] 'I leave/we leave' in which the stem final consonant appears only in a 'linking' capacity, similarly to the liaison cases of (5) and (8) above. The *-er* conjugation, however, correspondingly only has *je*[*pɔʁt*]/*nous* [*pɔʁt-õ*] 'I bring/we bring' where the consonant is uniformly present (see Burzio 1997, forthcoming, for additional Italian material; Tranel 1996 for further discussion of French linking consonants; Bullock 2004 for discussion of the role of PU in the history of French).

The situation characterized in (34) is obviously symmetrical, A_1 and A_2 being interchangeable in the discussion.[6] But if (34) is parallel to (33) as argued, we may expect the situation in (33) to then be symmetrical as well, and therefore that bare stems should also be influenced by their derivatives, not just the other way around. We consider this in (35).

(35) Uniformity with one's derivative (hypothetical)

	Entailments I	Entailments II	Entailments III
X-A	V ⇒ X		V ⇒ X
X=V		X ⇒ /___ A	where X= /___ A

The tabulation in (35) is like the one in (34) except that, instead of alternating between different affixal contexts, X alternates between one affixal and one categorial context in which it takes on the category V. For instance, for the verb *prevent*, entailment (I) would read: 'if something is a verb, it has the sound structure [prɪvent]' For the adjective *preventive*, entailment (II) would then read: 'if something has the sound structure [prɪvent], it must occur followed by *-ive*'. The combination of (I) and (II), and summation over instances of X (*prevent*, *generate*, *assert*, etc.) yields (III): 'if something is a verb, it has the same sound structure as also occurs before *-ive*'. This would predict that morphologically derived forms should be able to influence their bases just as well as the opposite. I have argued in *PES* (245), Burzio (1994*b*) that this is true, the exceptional stress pattern of verbs, e.g. *prevént*, being reducible in this manner to the regular one of their affixed forms, like *prevént-ing*, regularly stressed on a heavy penultimate. Similarly, *PES* (246) also takes the exceptional stress pattern of *-ic* adjectives to reduce to the regular one of their *-ical* counterparts, e.g. *académic*/*académical*. This 'back-copying' does not seem to hold more generally, though (as also confirmed by Kuryłowicz's (1949) 'second law of analogy', which privileges simpler over more

[6] This abstracts away from the possibility that either affixed form in (34) may be more prominent than the other perhaps for semantic reasons, or for reasons related to frequency of use (Bybee 1995).

complex forms). For instance, the stress of *párent* is regular rather than being like that of *parént-al*, despite the fact that the stress pattern of **parént*, while a bit rare, is attested for nouns: *cemént*, etc. The factors that militate against this kind of backcopying can in fact be traced to the entailments (II) of (35). With unproductive affixes, including *-ive*, such entailments are often false. For instance there is a verb *presént* (=X of (35)), but it is not the case that it occurs in the context of *-ive*: **presént-ive*. We may presume that such falsifications tend to cancel the force of the entailment (II) in (35), and hence the force of the entailment (III) that depends on it. This is in contrast with the entailment (II) of (33), which is not comparably falsified. That is, it is much more generally true that if a sequence precedes *-ive*, it occurs as a verb.[7] Hence, at least with unproductive morphology, the SELECT effect and hence the pressure for uniformity is asymmetrical, prevalently from base to derivative.[8] The case of affixed bases in (34) also presupposes productivity for the same reasons, namely the need for entailments (II) to be true with consistency—a criterion that is met by the Spanish example. Note that with Level 1 morphology, even the entailments (II) of (33) incur occasional falsifications, e.g. *compUlS-ive*, *FURT-ive*, *INEVIT-able*, where the base is either non-existent or partially different. But that was in fact the reason given for the relative weakness of PU effects at Level 1 altogether, where virtually the only such effects are the 'metrical' ones in (2) above, parasitic on segmental identity as argued in the previous section.[9]

With productive morphology, however, uniformity of a base with its derivatives will be expected. For instance, it is always the case that if something is a verb it will occur with *-ing*. Hence, entailments (II) in (35) are not falsified in that case, and hence it remains plausible to suppose that bare verbs copy the stress of their *-ing* counterparts. In addition to productivity, it seems also likely that the distinction between inflection and derivation may modulate the transitivity effect that yields entailments (III). The reason relates to the identity of the two Xs in each of (33)–(36). Such identity is—strictly speaking—not true. For example, the 'prevent' of $[prevent]_V$ and that of *prevent-ive* are not completely identical. The first is part of an entailment structure

[7] This backcopying would be different from the phenomenon known as back-formation, like the formation of *construct* from *construction* in the history of English. The latter leads to satisfaction of the conditional: 'if there is a form *X-ion*, then X exists as a verb'. This is essentially the SELECT entailment for *-ion*, and an OO-F effect. In the case of backcopying, on the other hand, the conditional is in the opposite direction: 'if there is a bare verb, it must have the same form (stress) as when it appears affixed with *-ing*'. On the one hand, the text analysis can treat back-formation simply as a case of OO-F. On the other, it is designed freely to permit bound stems (i.e. absence of back-formation) with Level 1 affixes. Further refinements would therefore be required to deal with back-formation.

[8] This account of the asymmetry supersedes the one attempted in Burzio (2002*b*), based on the notion that a bare base always violates some of the entailments produced by the derivative, like 'parent → /___al' due to *parental* but violated by *parent*. This consideration was invoked to argue that a bare base will always be subject to a weaker attraction by a derivative, but it is challenged by the fact that, at least with category-changing morphology, there is a comparable entailment violation by the derivative; e.g. the entailment 'parent → N', due to the base is violated by the derivative *parental*, as noted in the text just below.

[9] Segmental PU effects are rare with Level 1 morphology. They include *levy/levia-ble* of (4) above, where they are parasitic on stress identity as argued. Another rare case is *apha:sia* (regular *CiV*-lengthening)/*apha:s-ic*.

that includes the property of being a verb, while the second is part of an entailment structure that includes the presence of *-ive* and its associated semantics.[10] This misidentity, which is bound to attenuate the force of the entailment (III) in each of (33)–(36) depends on the magnitude of the contribution made by the affix, both positive, in adding to meaning and sound structure, and negative, in suppressing the former lexical category. Since the affixes generally referred to as 'inflectional' are the ones that contribute more modestly, leaving most of the properties of their bases unaffected, most notably their argument structure, one would expect stronger PU effects with inflection than with derivation. In this respect, the *-ic*/*-ical* cases will classify with inflection, since *-al* makes virtually no contribution and there is a good degree of consistency with which *-ic* adjectives have *-ical* counterparts. The plausibility that these cases also constitute a true instance of PU thus remains as well.

In sum, besides reflecting a general 'Gradient Attraction' effect, PU effects on stems are modulated by the selectional properties of affixes. We have seen that a bare stem is expected to affect its derivatives. The opposite effect, however, is expected only for productive and—we presume—inflectional morphology, the same restrictions holding for stems alternating between different affixes.

4.7.2 Uniformity of affixes

To the extent that all representations are taken to yield entailments, the REH predicts no fundamental difference in PU between stems and affixes. This seems correct, as I argue below. In *PES* I contended that stems and affixes are often in competition for uniform expression of their metrical properties, yielding for instance the two different outcomes of (36).

(36) *a.* -ist: Stem consistency prevails
a(mérica)(n-ìstø); (pròpa)(gánd-is)tø

b. -ic: Affix consistency prevails
ti(tán-icø); *(títan-i)cø

In the *PES* analysis, the suffix of (36*a*) exploits the special status of the final 'weak' syllable, which may or may not remain extrametrical, and treats that syllable in accordance with the stem stress. In (36*b*), however, the suffix does not avail itself of the same opportunity, but rather manifests a unique choice as regards the final syllable, thus forcing the stem to re-stress (cf. *títan*). The choice made by *-ic* in parsing the weak syllable is one that ensures the uniformity of *-ic*/*-ical* paradigms as noted above, as in *aca(dém-icø)*/*aca(dém-ica)lø*, but here there is no **?titán-ical* to be consistent with and that could thus not be the immediate reason. Rather, one must conclude that in

[10] It would then be technically incorrect to say that *-ive* 'subcategorizes' for a verb, since its stem is not an actual one. One could argue that in *prevent-ive*, the stem is still a verb, the adjective meaning roughly: 'able to prevent'. While obviously some of the verbal meaning is still present, the text reasoning is that major properties of the verb are lost in the adjective, which is true at least for the argument structure.

(36*b*) *-ic* is simply consistent as an affix, parsing in *ti(tán-icø)* just as it does in *aca(démicø)*, and winning the consistency competition with the stem, unlike *-ist* of (36*a*). It is argued in *PES* that the basis for such competition is essentially statistical. Those affixes that are not able to accommodate the stem stress in all cases choose to satisfy their own uniformity instead. This is the case with *-ic*. While the stress of *títan* could in principle be accommodated as indicated in (36), with items like *al(lérg-icø)*/ *al(lérg-í)cø*, neither option for the final 'weak' syllable enables the stress to duplicate that of *(áller)gy* (whose final syllable is also weak: *PES*, sect. 3.6). This is in contrast with *(áller)* (*g-ìs.tø*), where *-istø* has enough material to form its own separate foot, unlike **(áller)(g-ì.cø)* where the final foot headed by a light syllable is ill formed. We can characterize this situation by postulating that the uniformity of affixes prevails over that of stems in the more general case as in (36*b*), but that matters reverse when the uniformity of stems can be achieved consistently over all stems, as in (36*a*). The reversal can be understood in terms of entailment summation. With affixes like *-ist*, the SELECT entailment 'IST ⇒ /N__' can always be satisfied with respect to stress (recall that a SELECT entailment is an abbreviation for a cluster of entailments that pertain to specific pieces of the representation), and it can thus undergo summation over the whole class of stems, resulting in a stronger requirement of metrical consistency for its stems: 'IST ⇒ /$N_{(stress)}$__' . In the case of *-ic*, however, the class of cases like *allérg-ic* constitute inevitable falsifications of the corresponding SELECT entailment, resulting in an ultimately weaker requirement that the stem be metrically consistent: 'IC ⇒ /$N_{(stress)}$__'. This weaker requirement is evidently outranked by the consistency of the affix itself. It would appear then that affixes are endowed with greater inherent stability than stems, only subject to the reversal just described. The vowel shortening facts in (37) confirm this conclusion.

(37) Vowel shortening in affixes:

a.	satir-i:ze, organ- i:ze	*a'.*	organ-iz-ation
b.	gener-a:te	*b'.*	gener-at-ive
c.	altern-a:te	*c'.*	altern-ateN

As argued above, vowel shortening is a rather general phenomenon of Level 1 formations. It occurs for example in the stem *satir-* of (37*a*) (cf. *sati:re*). Yet affixes are generally spared as shown in (37*a*, *b*, *c*), except under the conditions of (37*a'*, *b'*, *c'*). We interpret these facts as follows. Each occurrence of an affix is in a Gradient Attraction relation to other occurrences of the same affix through the lexicon. Hence the *-i:ze* of *satir-i:ze* in (37*a*) is attracted to that of *organ-i:ze* and vice versa, the *-a:te* of (37*b*) is attracted to that of (37*c*) and vice versa, etc. More accurately, each occurrence of an affix is conditioned by the set of entailments collectively generated by all other occurrences—just the consequence of the REH. We then take lexical category to be a major component of semantic representation, consistently with some of the previous discussion. When an affix bears categorial information, as in (37*a*, *b*, *c*), attraction is maximal because their representation is maximally similar to that of their attractor. The usual considerations of limited productivity of Level 1 morphology enable us to iden-

tify as the main attractor the long-vowel allomorph, rather than the short vowel ones of (37*a'*, *b'*, *c'*)—the same considerations that enabled us to dismiss a Level 1 derivative as a significant attractor to its base (discussion of (35)). In contrast, stems such as *satir-* of (37*a*) are in a weaker relation to their attractors precisely because they have lost their lexical category. We know that the affix in (37*a*) acts as an enhancer of the stem's attraction, by way of its SELECT entailment, but we also know that, with such Level 1 affixes, the enhancement effect is relatively weak. Hence we take the stronger attraction affecting the affixes in (37*a*, *b*, *c*) as sufficient to prevail over the vowel shortening constraint, while the weaker attraction of Level 1 stems will itself be prevailed upon. When affixes themselves lose their canonical lexical category in Level 1 processes, however, their behavior will be expected to change to the more general one of Level 1 stems, undergoing shortening, indeed as in (37*a'*, *b'*, *c'*). The case in (*c'*) shows that it is not further affixation per se that is necessary, but rather just the switch away from the category of the main attractor.

The above interpretation is further confirmed by the Italian facts in (38), which closely parallel the English ones in (37).

(38) Italian (Burzio 1998)

a.	vénd-ere 'sell-INF.'	*a'.*	vínc-ere 'win-INF.'
b.	vend-út-o 'sold-PART.'	*b'.*	vín-t-o 'won-PART.'
c.	*vénd-ut-o/-it-o	*c'.*	*vínc-ut-o/-it-o
d.	vènd-it-óre 'seller'	*d'.*	vìnc-it-óre 'winner'
d.	vènd-it-a 'sale'	*e'.*	vìnc-it-a 'win-NOM.'

The participle in (*b*) shows that the participial suffix *-út-* imposes its own stress, overriding the metrical consistency of the stem *vénd-* of the infinitive in (*a*) (adjacent stresses being excluded). The final *-o/-a* in (38) are merely inflectional endings that will not affect the discussion. De-stressing of the participial suffix, as in (*c*) or similarly (*c'*), appears not to be an option, and this is the general predominance of affixes noted above as in (36*b*). The case in (38*b'*) shows further, however, that the Metrical Consistency/PU of the stem can be attained here by resorting to a syncopated form of the participial affix, which, lacking vowels, will thus lack stress. The syncopated affix, which is either *-t-* as in (38*b'*), or *-s-* as in *pre-s-o 'taken'* depending on the verb, is obviously disuniform from its canonical counterpart *-út-*. In Burzio (1998) I suggested that the metrical consistency for the affix dominates its segmental consistency, syncope violating only the latter, while the forms in (*c*, *c'*) would violate the former. In the present context, we can reinterpret this account in terms of Gradient Attraction, however. Segmentally faithful *-ut-* is excluded when unstressed because too similar to the main attractor *-út-*, with which it would thus neutralize. This is the outcome in (38*b*). The participial allomorphs *-t-* and *-s-* are then to be thought of as suppletive forms—a type of morphological irregularity similar to that of (22*a*) above *compulsive*. As such, they are due to some 'input' provided to the surface form of the participle directly. This accounts for the unpredictability of the forms themselves, in particular the *-s-* allomorph, as well as the general unpredictability of which verbs utilize which

form. As with other irregularities, these are shared by multiple items, however, resulting in certain level of entailment satisfaction across each class. The segmental divergence between primary allomorph *-út-* and suppletives *-t-* and *-s-* can now be taken to weaken the GA effect sufficiently to permit the latter's existence, which will thus result in satisfaction of the stem's metrical consistency as in (38*b'*). The oscillation of (38*b*, *b'*) can be viewed as a roughly even tension between the attraction by main allomorph *-út-* (weakened for the suppletives), which prevails in (38*b*), and the stem's metrical consistency, which prevails in (38*b'*). While this account is unable to predict which verbs within this conjugation will have a syncopated participle and which will not (as this seems truly unpredictable (Burzio 1998)), it does correctly predict that only verbs of this conjugation will have syncopated participles. The reason is that other conjugations have affix-stressed infinitives, and hence stems with unstressed final syllables, e.g. *sap-ére* 'know'. Since the claim is that participial syncope (a violation of affixal PU) is needed to maintain the stress of the stem-final syllable (stem PU), no syncope will be expected when there is no such stress to maintain, correctly, as there is no case like *sap-ére*/ **sát-t-o*, 'know/known' but only *sap-út-o* 'known' (see also Burzio 2003).

Turning to the nominalizations in (38*d*, *d'*, *e*, *e'*), they mirror the English cases in (37*a'*, *b'*, *c'*): when an affix is no longer associated with the same lexical category as its attractor, it is free to undergo further changes (weaker attraction). Hence the unstressed and yet un-syncopated participial allomorph *-it-* which was excluded in (*b*, *b'*) now becomes possible and is utilized to satisfy other types of stress uniformity—both that of the stem, and that of the nominal suffix *-óre*—avoiding a stress clash between them by means of the intervening syllable. An interesting further twist lies in the fact that the unstressed allomorph is in fact *-it-*, rather than the expected *-ut-*. The latter is borrowed from another conjugation, as shown by (39).[11]

(39)	I. Conjugation/ infinitive	II. Participle (canonical)	III. Departicipial Nominal
a.	**-ere**: vénd-ere 'sell'	**-út-**: vend-út-o	**-it-**: vènd-it-óre
b.	**-ére**: conten-ére 'contain'	**-út-**: cònten-út-o	**-it-**: cònten-it-óre
c.	**-íre**: sped-íre 'send'	**-ít-**: sped-íto	**-it-**: spèd-it-óre

What (39) shows is that the *-út-*/*ít-* contrast of the participles disappears when the stress is removed, as in the nominals. We can deal with this by thinking of the distance between *-ut-* and *-it-* as being perceptually enhanced by the presence of stress—the same assumption needed to account for the structure of vowel inventories under the Dispersion Theory of s. 4.3 above. When stress is removed, distance is reduced, promoting neutralization under GA. Hence this is another OO-F effect. The choice of uniform *-it-* over uniform *-ut-* may perhaps be due to the less marked status of *i* com-

[11] Beside the three conjugations in (39), Italian has the already mentioned *-áre* conjugation, which is totally regular and hence also immune from the syncretisms of (39): *port-áre*/*port-át-o*/*pòrt-at-óre* 'carry/carried-PART./carrier'.

pared with *u*. Note that alongside of the syncretism of column (39.III), there is a partial one in column (II) as well. By analogy with the infinitive, the participle in (39*a*) should be unstressed rather than stressed *-út-*. We may thus assume as already suggested that the segmental identity of the stressed and unstressed variants will bring about their neutralization to *-út-* (yet another OO-F effect). The asymmetry between participles and their infinitives in (39*a*, *b*), which do not undergo a comparable neutralization, is likely to reduce to a generalization to be reviewed in the next section, to the effect that the more 'basic' categories are able to maintain the most distinctions. The facts in (39) also shed light on the fact that in the competition among *-út-*, *-t-*, and *-s-* for the expression of participial morphology in the conjugation of (39*a*) (shown in (38*b*, *b'*)), it is *-út-* that acts as the main competitor or attractor. The reason for this is that *-út-* is more regularly instantiated through the lexicon, being present not only in the conjugation of (39*a*), but in that of (39*b*) as well..

In sum, English derivational affixes exhibit PU effects in their resistance to the otherwise general phenomenon of V shortening. This is paralleled by the fact that in Italian affixes do not undergo stress allomorphy. The superior stability of affixes appears tied to their role as exponents of the word's lexical category. When such role is revoked, affixes undergo allomorphy just like stems—a typical clustering effect in the present analysis. Such PU properties of affixes cannot be expressed by more traditional means. Without surface-to-surface relations, affixes are only expected to be uniform by virtue of a common UR, an expectation that is falsified by the critical role played by surface properties. Thus, the difference between *-ic* and *-ist* in (36) would require associating *-ic* but not *-ist* with a fixed metrical structure in UR. However, this would miss the fact that the different behaviors are ultimately predictable from the ability of *-ist*, but not of *-ic*, to accommodate the stem stress over the whole class of stems, a global property over surface structures. Correspondingly, the disuniformity of the paradigm *-út-/-t-/-s-* in the expression of the participial affix in the *-ere* conjugation would have to result from different URs, given both the segmental unpredictability of the *-s-* allomorph and the unpredictable distribution of allomorphs within that conjugation. But this would miss the fact that such allomorphy is none the less motivated by the metrical consistency of the stem—again a surface property. Finally, the uniform *-it-* in (39.III) would require either underlying *-it-* rather than *-ut-* in (39*a*), or a rule turning *-ut-* to *-it-* in unstressed position. The first account would miss the fact that the *-ut-* to *-it-* switch is controlled by surface properties—absence of stress—while the second would miss the fact that the resulting *-it-* exists independently in another conjugation (a natural attractor in the present account). The stem metrical consistency in the *-ere* conjugation, which drives the participial syncope, is itself also inexpressible derivationally. In order to transfer the stress from the infinitive to the participle, a derivation would rather implausibly have first to form the infinitive and calculate its metrical structure, then remove the infinitival affix (the one responsible for the stem stress) and finally form the participle from the same—now stressed—stem (see Burzio 2003) for more detailed discussion). The present conclusion is thus that the patterns of allomorphy reviewed above do not reflect the contingencies of derivations, but rather the

general character of the lexicon as being 'conservative' in the sense of Steriade (1997*a*, 1999*a*), namely in making maximal use of independently existing material, a property that we have reduced to the REH and Gradient Attraction. When the forces of allomorphy, either phonology or idiosyncrasy (= IO-F) move a representation away from its main attractor, other neighboring attractors stand at the ready.

4.7.3 Morphological leveling

Another well-known instance of PU is provided by morphological syncretism, sometimes also referred to as 'morphological leveling' (Hock 1986: s. 9.1.4.). The Old English (OE) paradigms in (40) provide one illustration.

(40) Old English weak conjugation: *trymman* 'strengthen'

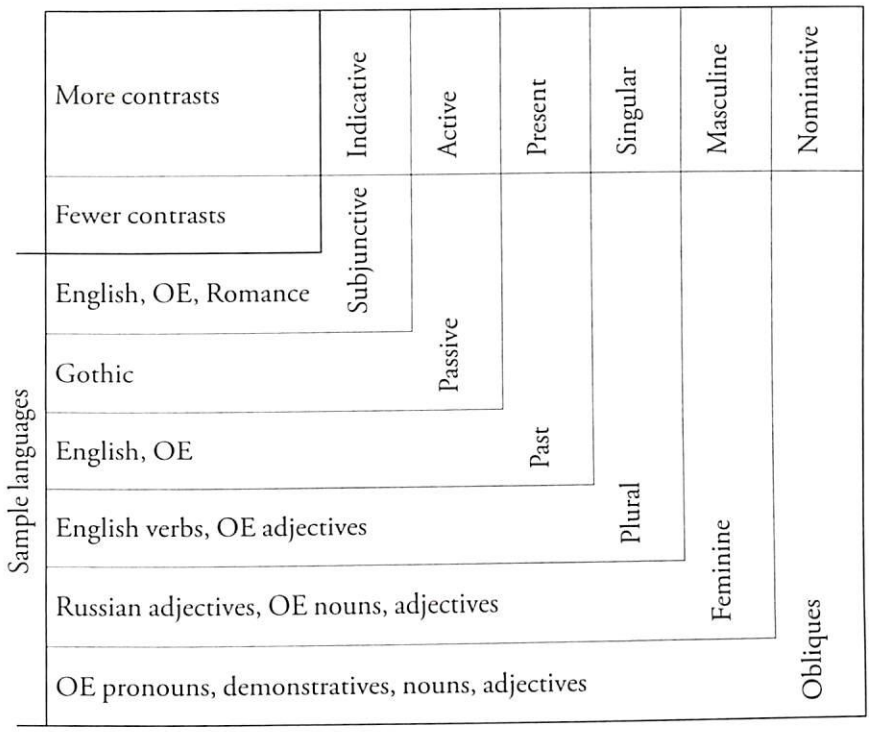
	Present		Past	
	Indic.	Subjunct.	Indic.	Subjunct.
1sg	trymme	trymme	trymede	trymede
2sg	trymest		trymedest	
3sg	trymeð		trymede	
1pl	trymmað	trymmen	trymedon	trymeden
2pl				
3pl				

In OE, the inflectional paradigm of the subjunctive was considerably more leveled than that of the indicative. That of the past indicative was also more leveled than that of the present (the past first and second person singular have identical inflections), and in the indicative the plural was more leveled than the singular. The same three asymmetries are in fact also detectable in Modern English, as the *-s* of the third person singular indicative has no counterpart either in the past, the subjunctive, or the plural. These tendencies as well as the further ones in (41) appear to hold cross-linguistically, pointing to the generalization in (42). The difference noted for Italian (39.I, II) above suggests that 'infinitive' is also a less marked category than 'participle'.

The framework of Distributed Morphology (DM: Halle and Marantz 1993, 1994; Noyer 1992; Embick and Noyer 2001; Bonet 1991) provides a way to capture such effects by means of special rules of 'Impoverishment'. For example, the case of contemporary English subjunctive, *It is necessary that he walk*, could be handled in the manner of (43).

In both examples in (43), the inflection node in the syntactic structure would initially bear the feature [3SG], reflecting its agreement with the subject. In the indicative case, vocabulary insertion would then apply directly, choosing vocabulary item *-s* over

(41)



(42) Syncretism Generalization: More marked categories exhibit fewer internal contrasts (more syncretism/leveling) than less marked ones.

(43)

Example	Vocabulary	Impoverishment rule	Voc. Insertion
He walk-s	[3sg] ⇔ [-s]	[3sg] ⇒ ø /governed by	[-s]
	⇔ -ø	Subjunct.	
. . . that he walk-ø			ø

-ø (null), because it best matches the features of its insertion node. In the subjunctive, however, DM would postulate the Impoverishment rule shown, applying prior to vocabulary insertion. With the featural content of the insertion node thus 'impoverished', vocabulary insertion will now find the null affix as the best-fitting candidate.

While descriptively adequate, impoverishment rules provide little explanation since they add complexity to the formalism for no reason that the formalism itself makes evident. In contrast, morphological syncretism is directly expected under the REH, modulo the assumption in (44).

(44) The unmarked categories—Indicative, Present, Singular, Masculine, Nominative—are unspecified.

According to (44), for example, 'Indicative' would not consist of a specific value for the feature 'Mood', but rather of the absence of such a feature altogether. This assumption

seems natural and has been independently advanced by work such as Jensen (1990) and Noyer (1993). Of particular relevance is the fact that it would also be required by the DM approach. If unmarked categories are unspecified, then impoverishment rules simply cannot refer to them. Hence, for instance, the rule in (43) could not refer to the indicative, incorrectly yielding leveling in the indicative rather than the subjunctive, because 'indicative' does not exist—the mere absence of a feature.[12]

Returning to the REH, if (44) is correct, then marked categories always add a dimension or component to the representation. This will have the effect of enhancing the leveling pressure among the members of the paradigm, in the way illustrated in (45).

(45) Morphological leveling

		Indicative		Subjunctive	
a.	*walk*, basic sound and meaning	A B	A B	A B	A B
b.	Added component: [Subjunctive]			C	C
c.	Alternating component: [3sg] ⇔ [-s] / other ⇔ -ø	S	¬S	S	¬S
d.	Entailments	A ⇒ S B ⇒ S	A ⇒ ¬S B ⇒ ¬S	A ⇒ S B ⇒ S C ⇒ S	A ⇒ ¬S B ⇒ ¬S C ⇒ ¬S
e.	Leveling pressure (contradictory entailments)	2		3	

In (45), we make again the simplifying assumption that the basic sound and meaning of a verb like *walk* can be given a simple two-component representation A, B, as in (*a*). Then we take the two affixes *-s* and -ø to instantiate entailments from sound to meaning and vice versa as in (45*c*), consistently with both the DM view in (43) and the REH. In the table, we refer to these two complementary choices of affixes simply as S (for affix *-s*) and ¬S. Unlike in DM, the entailments provided by the REH will not only run internally to each morpheme, however, but across different morphemes as well. For instance, in the representation *walk-s*, the subpart *walk* entails *-s*, as would be expressed by the two entailments A ⇒ S, B ⇒ S. But the representation *walk*, as in 'I walk' will generate contradictory entailments from *walk* to absence of *-s*, namely: A ⇒ ¬S, B ⇒ ¬S. Hence the Indicative alternation is disharmonic as indicated in (45*d*, *e*) by way of two violated entailments, which thus measure the extant pressure towards

[12] The unspecification of (44) is not to be confused with the underspecification approach to syncretism (see especially Wunderlich 1995). In the latter approach, for example the English past tense morpheme *-ed* would lack specifications for person or number (a kind of underlying impoverishment). The difference is clear in that in (44) it is the non-syncretic categories (e.g. present) that are unspecified. The general effect that needs to be captured is that there appears to be a cap on featural specifications: if a form is specified for tense or mood it is less likely to be also specified for person. One can well imagine the underspecification approach postulating such a condition. The present approach is the only one, however, in which the cap follows from the inherent properties of the system, as shown below.

morphological leveling. Recall that in the present approach *all* morphological alternations are disharmonic (as in (23.I) above, for example) and hence occur only as compelled. Specifically, failing to combine *walk* with *-s* as in 'He walks' would violate the entailment '[3SG] ⇒ [-s]'of (45*c*), which is evidently able to stem the leveling pressure over the indicative paradigm. But when we move to the subjunctive, the leveling pressure will be greater due to the added component—the abstract semantic feature [subjunctive]: C of (45*b*). The new pressure consists of three entailments, as indicated in (45*d*, *e*), and this is evidently able to turn things around, now overriding the entailment '[3SG] ⇒ [-s]'of (45*c*) and resulting in the fully uniform paradigm of the subjunctive. This discussion has so far not provided a way to choose between *-s* and -ø as the leveling outcome. We presume that the choice of -ø is due to the fact that the latter is independently predominant in the paradigm, so that extending it rather than *-s* constitutes a more minimal change.

In sum, while it would be hard to predict exactly when syncretism will occur, we are correctly predicting the general circumstances under which it may occur: those of (41), (42). As with other types of PU, more traditional resources shed little light on the phenomenon of syncretism, by requiring additional machinery to express it. In contrast, on the present approach, the latter is reducible to the general neutralization effect predicted by the REH: GA. The reason that more marked morpho-semantic categories are more highly syncretic is that they correspond to representations that are independently more similar by sharing more features. As usual, similarity promotes further similarity, hence the leveling. Obviously, syncretism is yet another type of PU that the notion of cyclicity has no way of capturing. The fact that the subjunctive inflection of '. . . that he walk-*ø*' is the same as that of '. . . that I walk-*ø*' is plainly not because they are related by a cyclic derivation.

4.8 CONCLUSION

Paradigm Uniformity can be dealt with in terms of the OT notion of Output–Output faithfulness: members of a paradigm are required to be faithful to other members of the same paradigm. Yet this view raises a number of fundamental questions, listed in (3) above, which the present chapter has sought to answer.

One question (3*a*) was what criteria identify the sources of OO-F for any form under calculation. If being members of the same paradigm is at least a necessary condition, then the question is what defines the notion of paradigm. Here, I have argued first that the problem of allomorphy demands a theory able to express the notion of attraction over distance, a respect in which this problem intersects significantly with problems in both phonology proper and perception, suggesting a single overarching property of mental representations. I have then argued that such a theory needs to incorporate the Representational Entailments Hypothesis, under which representations become analyzable as sets of mutual entailments or dependencies among their components,

similarly to the distributed representations of neural nets. In part, OO-F would then consist of attraction by neighboring representations, and 'paradigms' would be definable as neighborhoods. I have argued further, however, that regular morphological relations modulate the attraction effects substantially. For instance, the attraction of *parént* by *paréntal* is insignificant, the underlying reason being that the relation of nouns to *-al* adjectives is not sufficiently regular or general, witness **studental*, while attraction in the opposite direction is significant, as the relation of *-al* adjectives to nouns is fairly regular. Even this regularity is moderate for Level 1 formations, however: *arboreal*/**arbore*, in contrast to the much greater regularity of Level 2 formations. In further contrast, in cases of fully productive morphology, such as *prevént*/*prevénting* or the Spanish paradigm in (1), pairwise relations are regular in both directions, and attraction is thus expected—at least in principle—to be symmetrical. From this point of view, the notion of 'paradigm' would then also require reference to the degree of regularity of morphological relations. It is ill-suited to express the asymmetries, however (e.g. the one of *parént*/*paréntal*, since there is no sense in which one item could form a paradigm with the other but not vice versa) and hence the term 'paradigm' remains of limited theoretical relevance. In the present analysis, the descriptive notion of PU is thus replaced by the OT notion of OO-F, in turn replaced by the combination of GA entailments, expressing attraction over distance, together with SELECT entailments, expressing the role of morphological relations.

Another question (3*b*) concerned the relation between morphology and OO-F and the potential redundancy between them in generating surface-similarity patterns. In answer to this question, I have argued that morphology *is* OO-F, in the form of the SELECT entailments, and hence that there is no redundancy. For instance the entailment 'IST ⇒ /N __' simultaneously validates both the segmental structure of *américanist* (2*a*′) above, formerly the result of word-formation rules of the morphology, and at the same time its accentual relation to *américan*, a type of OO-F. Even aside from their redundancy with OO-F, traditional approaches to morphology have been argued to be inadequate because of their reliance on inviolable formalisms, a characteristic that makes it impossible to integrate the morphology with the phonology, in particular to express the inverse distributional relation of morphological and phonological regularities of (22) above.

Finally, a third question (3*c*) was whether there was any residual validity to the traditional notion of the cycle, the answer to which was negative. Cyclicity would at best be applicable to only a small portion of the PU phenomena reviewed: those where an affixed derivative is uniform with an unaffixed base, like (2*a*′) and possibly (2*b*′). It is inapplicable to the cases where there is no affixation, like the French liaison cases in (5); it is also inapplicable to cases where the base has its own affix, such as the Spanish cases in (1) and the Italian ones in (38*a*′, *b*′) (i.e. to the general schema in (34)); it is inapplicable as well to the uniformity of affixes of s. 4.7.2, and to the morphological leveling/syncretism cases of s. 4.7.3.

The present discussion has thus shown that no legitimate derivational basis remains to either phonology or morphology.