Phonologically Conditioned Syncretism

Luigi Burzio
Johns Hopkins University

1. Introduction

The mainstream generative tradition had postulated discrete modules that feed one another in a cascading arrangement: Morphology would feed Phonology which would then feed Phonetics. This hypothesis makes grossly incorrect predictions about the range of possible interactions. It predicts that Phonology could not be driven by Phonetics except perhaps indirectly via evolutionary effects that weed out phonetically ill-suited phonologies, and it predicts that Phonology may not have any effect on Morphology. The incorrectness of the first prediction has been forcefully underscored by a very productive line of work of recent years aimed to show how perceptual cues and perceptual distances are behind phenomena that have been the traditional bread and butter of phonological work. See, e.g. Hayes et al. (2004). The present article addresses the incorrectness of the second prediction, by considering syncretism –an eminently morphological phenomenon, which is nonetheless controlled by phonological factors in certain cases. I will argue in particular that several cases of syncretism, including the well known 'spurious *se*' case of Spanish are responses to phonological demands. These cases and their distributions will be shown to be inexplicable in terms of purely morphological approaches to syncretism. While adopting a constraint-based approach to Phonology as in Prince & Smolensky's (1993) Optimality Theory, I will argue for a constraint based approach to Morphology as well. In this larger constraint-based framework, Morphology and Phonology are able to work in parallel in the determination of the optimal output, hence accounting for the interactions.

The article is organized as follows. The next section introduces the present model of Morphology, and section 3 characterizes the nature of syncretism within that model. Section 4 considers the widespread replacement of a dative clitic when followed by an accusative clitic in the Romance languages, including the notorious 'spurious *se*' case of Spanish, arguing against a purely morphological approach. Section 5 reviews the complex pattern of syncretism in Campidanian Sardinian, and considers that, while underspecification is key to accounting for the direction of syncretic spreading, the latter notion is correctly definable only over surface forms and not over underlying entries, an issue highly germane to the issue of parallel versus serial architecture. Section 6 considers cases where the Phonology is not only the cause of syncretism, but is also responsible for subverting the direction of the syncretic extension. Section 7 ties the arguments based on syncretism with other known arguments for parallel Morpho-phonology. Section 8 concludes.

2. Morphology and Representational Entailments

The present approach to Morphology consists of postulating a set of relations among surface forms enforced by appropriate constraints. Such constraints aim to express two major effects that can be labeled 'selection' and 'attraction', respectively. The most typical selection effects are expressible as in (1a, b), each of which is to be construed as a surface or output constraint.

(1) SELECTION constraints
a. **Affixation to a base**.   E.g. English *-al*:   / __ -al = N
b. **Affixation within a paradigm**.   E.g. Italian verb inflection: / __ -iamo = / __ -iate

The constraints in (1a, b) are effectively subcategorization frames for affixes. The one in (1a) states that *-al* must be preceded by a noun, as in *parental*, the noun being *parent*, and the one in (1b) states that 1st plu-

ral subjunctive and the 2nd plural subjunctive affixes *-iamo, -iate* must attach to the same stem, as in *portiamo* 'we carry', *portiate* 'you carry'. The constraints in (1) differ from the subcategorization frames of other work in morphology, however, both because they are enforced over surface forms, and because they are violable. For example, the one in (1a) demands identity not with the underlying form /pærɛnt/, but rather with the surface form [pǽrənt]. That identity is obviously violated by the stem [pərɛ́nt] of *parental*. Application of (1a, b) to surface forms is what accounts for the fact that, in certain cases, surface properties such as stress are indeed transferred, unlike in the *parent/ parental* case. For instance, in a case like *Américanist*, the pre-antepenultimate main stress is only explicable if inherited from its base *Américan*, where it is regular (cf. also the regular stress of *antágonist*, where there is no base **antagon)*. Surface application of constraints like (1b) accounts for example for the Italian form *cono[šš]-uto* 'known' where palatalization uncharacteristically occurs before a back vowel (cf. *conos[k]-o* 'I know'), by taking participial stems to stand, in the manner of (1b), in a (high-ranked) relation with their infinitival counterparts, here as in *cono[šš]-ere 'to know'*, where the palatalization is regular before the front vowel. The constraints in (1a, b) each correspond to a class of Output-Output faithfulness (OO-Faith) constraints of recent Optimality Theory (OT) literature. The one in (1a) correspond to what is referred to as 'Base Identity', demanding that a stem be faithful to its base, while the one in (1b) corresponds to the 'Paradigm Uniformity' or 'Optimal Paradigms' constraints, demanding that a stem exhibit an invariant form through an inflectional paradigm (see in particular McCarthy, 2005). The constraints in (1) differ from the conception of OO-Faith constraints in other work, however, in that they are not taken to exist in the Phonology independently, in a way that happens to mirror properties of the Morphology, but they are taken instead to constitute the Morphology proper, which is enforced directly over the phonological output via parallel interaction, as argued in Burzio (2002a, b, 2005a). The conception of Morphology that utilizes constraints like (1a, b) above is consistent with the tradition that is referred to as 'Word and Paradigm' or 'Word-based' Morphology (Matthews, 1972; 1991; Bybee, 1988; Bochner, 1993; Becker, 1993, Aronoff, 1994; Blevins, 2005. See also Pirrelli & Battista, 2000), which also postulates a set of relations from words to other words, in contrast to the mainstream Generative tradition, that endorses the 'Item and Arrangement' approach, in which relations go from constituent morphemes to words. However, the system in (1) is more directly grounded in the independent 'Representational Entailments Hypothesis' (REH) of Burzio (2002a, b, 2005a, b), stated in (2).

(2) **Representational Entailments Hypothesis.** Mental representations are sets of entailments, a representation A, B, C corresponding to the set of entailments {A $\Rightarrow$ B; B $\Rightarrow$ A; A $\Rightarrow$ C; C $\Rightarrow$ A; B $\Rightarrow$ C ; C $\Rightarrow$ B} (if A then B, etc.).

In (2), components A, B, C could be construed for instance as the constituent features within a phoneme, or as the constituent segments within a morpheme, or as the morphemes within a word, etc. The REH (2) embodies the claim that mental representation takes cooccurrence of any two entities A, B, to be necessary, such that any later occurrence of A will yield an expectation of B and vice-versa. This claim is independently entertained in neuroscience, where it is known as 'Hebbian learning', after Hebb (1949). The hypothesis in (2) yields both of the effects cited above as key to a characterization of Morphology. The first is selection, representable in the form of constraints like (1a, b) above. The second is attraction, which will be taken as the basis of syncretism. The latter effect is predicted to obtain between neighboring representations. To see this, consider a representation R_1 consisting of A, B, C, referred to the 'attractor' and the set of entailments it generates. Consider as well a neighboring representation R_2 = A, B, -C (where -C: 'minus C', or 'not C'). Occurrence of -C in R_2 will violate two of R_1's entailments, namely A $\Rightarrow$ C; B $\Rightarrow$ C. However, in a more distant representation R_3 = A, -B, -C (where two-components now differ from those in the attractor), the same '-C' will now violate only one entailment, namely A $\Rightarrow$ C (while B $\Rightarrow$ C will now be satisfied). While the reader is referred to Burzio (2002a, b, 2005a) for more detailed discussion, it is easy to see that the number of entailments pressuring any '-C' of an R_i to assume the same value C as in the attractor R_1 equals the number of components that R_1 and R_i have in common, namely their similarity or proximity in space, whence the label 'attraction' (the shorter the distance the stronger the force).

Attraction can be seen as the factor behind a wide range of phenomena, including morphological syncretism. Consider in this regard that members of an inflectional paradigm are rather similar to one-another by definition, normally sharing the same stem as imposed by constraints like (1b) and differing only by the phonological form and the morpho-syntactic features of the affix. Affixes with similar feature matrices, i.e.

sharing specifications for number and/ or person, mood, and so on, will result in stem-affix combinations that are highly similar to one-another, hence prone to neutralizing under attraction (see Burzio & Tantalou, 2007 for detailed discussion). The list in (3) is to give a general sense of the theoretical potential of the REH and the attraction effects it predicts.

(3) **Scope of attraction in Phonology and Morphology**
   a. **'Dispersion'** effects in segmental inventories; Neutralization of weak segmental contrasts (e.g. Coda devoicing; reduction of vowel inventories in unstressed positions) in the spirit of Steriade (1995).
   b. **Non Derived Environment Blocking** (= attraction to an input; see below); Lexical Conservatism (= attraction to a neighboring output; see below).
   c. **Morphological blocking** (E.g. *arrivation* blocked via neutralization with *arrival*); **Morphological syncretism**.

   Some of the phenomena in (3) will be touched upon below in addition to syncretism. For further discussion, see Burzio (2002a, b, 2005a, b).
   Along with attraction, selectional constraints like those in (1a, b) are also reducible to the REH (2) as argued in Burzio (2005a). In essence, a complex word like *parent-al* will yield a complex set of entailments, a subset of which can be summarized as: 'al $\Rightarrow$ / [pərént] __' (the sequence *al* entails the left hand context [pərént]). In a similar vein, the noun *parent* will yield the entailment '[pǽrənt] $\Rightarrow$ N' (the sound structure [pǽrənt] entails the semantic category 'Noun'). Transitivity of entailments will then result in the entailment 'al $\Rightarrow$ / N __' (*al* entails a noun as its left-hand context), which is equivalent to (1a) above. This is true despite the fact that the noun [pǽrənt] and the stem [pərént] are not identical. Their similarity will be sufficient to ensure some degree of transitivity. Note that meaning is also part of the representation and entailment structure, also contributing to the similarity relation. In the present framework, the strength or rank of constraints like those in (1) will in fact depend on the degree of identity holding over such pairs, as well as the number of contributing pairs (like *nature/ natur-al; territory/ territori-al*, etc.). A cluster of differences holding over the behavior of level 1 affixes like *-al, -ic, -ous, -ity*, and that of level 2 ones like *-ness, -less, -ful,* is known to reduce to the different ranks of the respective selection constraints, as I note further below and in Burzio (2002b).
   Attraction and selection can superimpose on one-another, as for instance in *rem[íy]di-able*, versus *l[é]vi-able* (as in 'levy a tax'). In the first case, general constraints on stress require re-stressing of the stem compared with its base *rémedy*. This difference makes the base a weaker attractor, which in turn opens the way for a regular lengthening process, the same one that also applies in *Can[ey]dian, Mong[ow]lia*, etc. In the second case, where re-stressing is not analogously forced, the base *l[é]vy* will be a stronger attractor of the stem, blocking the lengthening process. This is an instance of the Non Derived Environment Blocking (NDEB) of (3a) above, as lack of re-stressing in *l[é]vi-able* provides an environment that is not 'derived', i.e. not affected by changes, correlating with the blocked lengthening. In the present framework, NDEB effects result from the fact that faithfulness relations do not have a fixed rank as in mainstream OT, but rather have a rank characterized by attraction, hence dependent on distance. In the cases in question, the output is faithful to its input. However, that 'input' is provided in each case by the surface forms of the respective bases *remedy* and *levy*. Hence, the relation involved is effectively an OO-Faith relation as discussed in connection with (1). Thus, the introduction of attraction updates the character of OO-Faith relations compared with (1). The latter relations are now viewed as the combined effects of selection and attraction. Selection is affix-based, affixes varying in the strength of their selection requirements. The contrast *rem[íy]di-able/ l[é]vi-able,* however, cannot be due to selection since the same affix is involved. Attraction, on the other hand, is based on the distance between the allomorphs. The overall OO-Faith effect is then the result of summation of the entailments corresponding to each individual effect.
   Summing up, the phenomena to be addressed below will require that Morphology and Phonology interact in parallel. This interaction is made possible by taking not only the Phonology to be a system of violable constraints as in OT, but the Morphology as well. A system of constraints enforcing appropriate morphological relations can be constructed on the basis of the REH in (2). Within the latter, selectional properties of affixes as in (1a, b) arise via summation of like entailments across the lexicon. In addition, similarity of representations yields pressure towards neutralization. This will provide a basis for analyzing syncretism as

we see below.

3. The roles of similarity and markedness in syncretism

In this section, I review certain general tendencies exhibited by syncretism cross-linguistically, which seem predictable from the present approach. One of these tendencies is for syncretism to affect adjacent cells in the paradigm. From this point of view, the Italian singular present subjunctive of regular verbs in (4a) would be typical, while the case in (4b), illustrating the verb 'be' in the present indicative, also in Italian, would be a-typical.

(4) a. **Typical** 'carry' PRES. SUBJ.

	SG	PL
1		port-iamo
2	port-i	port-iate
3		port-ino

b. **Atypical** 'be' PRES. INDIC.

	SG	PL
1	sono	siamo
2	sei	siete
3	è	sono

From the present perspective, the tendency in (4a) results from the notion of attraction. Adjacent cells represent maximally similar sets of morpho-syntactic features by definition, differing by exactly one feature (the feature [person] in (4a)). Hence, the affixal forms that represent those features would be maximally similar in that regard. In addition, the word forms containing those affixes would share identical (or highly similar) stems, hence contributing to the overall similarity of those forms. I will suggest that superficially atypical cases like (4b) may result from phonological similarity playing a comparable role to morpho-syntactic similarity, the Latin etyma in this case being *sum* 'I am', versus *sunt* 'they are', yielding *sum/ sun* after likely sound changes.[1]

While the entailment framework based on (2) is numerical in character like connectionist models, thus not abiding by the strict-domination hypothesis of OT, the OT machinery can still be recruited in regard to the latter, at least for general expository purposes. Specifically, we can assume a scale of attraction constraints as in (5), where the highest ranked constraint corresponds to the closest (i.e. most similar) attractor, and the lowest ranked one to the farthest attractor.

(5) $\text{ATTR}_1 >> \dots \text{ATTR}_i >> \dots \text{ATTR}_z$ (1 = closest; z = farthest)

Attraction constraints demand identity with the attractor, hence they are interpretable as types of faithfulness constraints in OT, whether attraction combines with selection as in the case of base-to-stem relations

---

[1] An anonymous referee raises the important question of the relation between synchrony and diachrony and to which of these notions the present accounts may pertain. As the referee notes, there is one case where this distinction seems important and would need to be drawn explicitly. This is the case where syncretism is driven by phonological similarity. Where such similarity is accidental (hence not reconstructable from the rest of the system), as is the case of Latin *sum/ sunt* of the text, a synchronic grammar could obviously not take advantage of it, since the original data are no longer available to learners. In any such case, the account must remain purely diachronic (or synchronic for the earlier stage of the language at which the relevant data were available). The contemporary synchronic grammar must therefore treat the resulting identity as an accident, unless synchronically available data also support it. In regard to (4b), the referee notes that the rest of the paradigm might point to either 1SG *s-o* or 3PL *son-ono*, suggesting then plausibly that the former may be dispreferred due to its synonymy with 'I know', while the latter violates the OCP, hence providing some synchronic justification for the two identical *sono*. It would remain to be seen, under either the purely diachronic account or the partially synchronic one suggested by the referee, how the Italian case is to be distinguished from that of either Rumanian or -as the referee notes- Ancient Greek, where the 1SG / 3PL syncretism is not limited to 'be' but affects entire conjugations.

In all other cases, however, I would see no need to draw the synchronic/ diachronic distinction. For instance, if the OCP is what led to the transformation 'le lo ⇒ glielo' historically, then the same OCP will favor *glielo* over *le lo* in the acquisition of the new synchronic grammar, despite the fact that such sequences are no doubt lexicalized, hence potentially idiosyncratic.

discussed above, or not, as in the present case. This practice of 'grounding' the character and rank of OT constraints in some external and independently understood domain is commonplace. In many cases, the grounding refers to Phonetics (perception or articulation. See Archangeli & Pulleyblank, 1994; Steriade, 1999a, Hayes et al., 2004, and others). In this case, the grounding refers to the hypothesis in (2). An OT-style calculation will then result in syncretism if a member of the hierarchy in (5) dominates faithfulness to the input (IO-Faith), that input specifying different forms for different cells, while lack of syncretism would result when IO-Faith is undominated. As noted above, IO-Faith would also be a form of attraction in the present approach, the attractor being the input representation.[2] Given the hierarchy in (5), when syncretism obtains, it should be as demanded by the nearest attractor all else being equal, hence accounting for its noted tendency to affect adjacent cells.

A second tendency or generalization is for syncretism to affect marked categories like subjunctive or plural. The overall distribution of syncretism can in fact be described as in (6).

(6) SUBJ > INDIC; PASS > ACTIVE; PAST > PRES; PL > SG; FEM > MASC; OTHER CASES > NOM

For more extensive discussion of this point and evidence from specific languages, see Burzio & Tantalou (2007). From this point of view, the subjunctive paradigm in (4a) would now be atypical, since the singular cells are affected, while the plural ones are not. A comparison of (4a) with the indicative counterpart in (7) may be revealing in this regard.

(7)

'carry' PRES. INDIC.

	SG	PL
1	pórt-o	port-iámo
2	pórt-i	port-áte
3	pórt-a	pórt-ano

The lack of syncretism in (7) compared with (4a) is consistent with the generalizations in (6). However, in addition, assuming that the indicative paradigm serves, in some fashion, as the 'model' for the subjunctive one, we see that the singular forms are highly similar to one-another, differing by a single unstressed vowel, while the plural ones are more heterogeneous (the stresses are not reported in (4a), but they follow the same pattern as in the indicative forms in (7)). The above suggestion that phonological similarity also plays a role in syncretism and is thus able to override the purely morphological tendencies in some cases is thus confirmed by this case.

An account of (6) from the present perspective will rely on underspecification, namely the assumption that the unmarked pole of each of the oppositions in (6) lacks the feature altogether, so that indicative would represent zero mood; singular zero number; and so on. From this point of view, the generalizations in (6) would follow from the fact that members of marked categories are relatively more similar to one-another in their morpho-syntactic make up, in the way schematically illustrated in (8).

(8) **Effect of Underspecification**

Less similar		More similar	
α	β	α	β
~~INDIC~~	~~INDIC~~	SUBJUN.	SUBJUN

---

[2] With the ranking of both syncretic Faith and IO-Faith pre-determined by the relevant similarity relations, one may wonder how languages may ever differ, or how language change may occur. The answer is that the nature of the input is more complex than so far suggested. Each of the entailments that constitute the input may have different strengths depending on its degree of satisfaction or violation through the lexicon. An entailment that is violated will be easier to violate further (see discussion of English *are* below). For this reason, the overall dynamical system will find multiple points of equilibrium, whence cross-linguistic differences. As for how it could move from one point of equilibrium to another as in language change, one will need to postulate that the system is inherently noisy. This is true in any model.

What (8) aims to suggest is that, whatever degree of similarity two forms of the indicative paradigm α and β may have, their subjunctive counterparts will automatically have a greater degree of similarity by also sharing the feature [subjunctive] in addition to whatever α and β share. By contrast, indicative mood cannot provide a comparable enhancement, by the underspecification hypothesis.

Finally, a third generalization concerns the direction of syncretism. The tendency is for a form that is already relatively general by occurring in multiple cells to be generalized further. An alternative characterization of this tendency would maintain that it is the 'unmarked' member of an opposition that is consistently extended. This tendency is illustrated, in either variant, by the Italian cases in (9).

(9) **Extension of the general or unmarked**

	SG, alone	SG/ ___ ACC	PL (colloq.)
MASC	gli	gli(e) lo	gli
FEM	le		

As shown in (9), in Italian the masculine singular dative clitic *gli* replaces feminine *le* in the context of an accusative clitic, so that a sequence *glielo* will be ambiguous ('it to him/ it to her'). At the same time, the same form *gli* also replaces non colloquial plural *loro* in colloquial styles. Prior to either extension in (9), clitic *gli* is already more general than either *le* or *loro* for having undergone one other extension. Being both masculine and singular, it is also less marked than the forms it replaces. Extension of the unmarked is predictable in the present framework in the way illustrated in (10) below.

(10)		Two forms	**gli**	**le**
	a.	similarity:	{3, DAT, clitic}	{3, DAT, clitic}
			~~MASC~~	FEM
	b.	Attraction:	*{3, DAT, ..} ⇒ le	*{3, DAT,..} ⇒ gli
	c.	One form	**le**: * le ⇒ FEM	
			gli: *(FEM ⇒ le)	

The two forms *gli* and *le* bear a similarity relation to each other for sharing the properties listed in the shaded line of (10a). This similarity is the source of the mutual attraction that could potentially result in syncretism, replacing either form with the other. More specifically, attraction takes the form of the entailments listed in (10b). The ones in the *gli* column are those generated by any occurrence of *le* and violated by *gli*, and conversely for the *le* column. Syncretism in either direction would remove such violations. Underspecification of the unmarked now provides the source of directional asymmetry. As indicated in row (c), replacement of *gli* (masculine) by *le* (formerly feminine) would violate the entailment 'le ⇒ FEM' (if the form is *le*, the gender is feminine) that other uses of *le* generate. By contrast, the symmetrical counterpart to that entailment, 'gli ⇒ MASC' does not exist, since by the underspecification hypothesis 'masculine' is only a descriptor and not a real feature, as shown in the 3rd row of (10a). In this regard, *gli* can then extend to feminine uses without a violation (extension of the unmarked). As indicated in the second line of (10c), however, extension of *gli* to feminine still violates the entailment FEM ⇒ *le* (feminine gender entails the form *le*). This expectation will not revoke the just noted asymmetry, however. The reason is that the two entailments in (10c) have a sharply different status in terms of their degree of satisfaction through the lexicon, which is the source of weight or ranking in the present system. The entailment 'le ⇒ FEM' is never violated outside of the case under consideration, while the entailment 'FEM ⇒ le' is massively violated, by any (nominal or pronominal) expression that is feminine but has a different form.[3] In the present framework, which is numerical and summational as noted analogously to the functioning of a neural network and the calculation of levels of activations in it, satisfaction adds to an entailment's strength, while violation reduces its strength. This is the notion of 'Hebbian learning' embedded in the hypothesis in (2) above: all associations are automatically learned, violations thus inducing negative learning. The differential status of the two

---

[3] This entailment is never violated in the context of 3rd singular dative clitics, but this will not challenge the asymmetry.

entailments in (10c) is expressed by the parentheses around the second. While the compromised status of the latter entailment thus preserves the needed asymmetry, with masculine *gli* taking over the function of *le* and not the other way round, that same entailment is nonetheless critically needed. Without it, there would be nothing preventing *gli* from incorrectly taking over the function of *le* altogether rather than only in certain contexts. We can take that entailment, and in fact both of the entailments in (10c), to represent the Input-Output faithfulness (IO-Faith) in OT. That is, we take it to be a property of the input that there exists a bidirectional association between feminine gender and the form *le* (along with the properties 'dative', 'clitic', etc.). The distribution of *le* can then be accounted for by assuming that IO-Faith relative to this form is undominated in general, *le* thus surfacing faithfully, but dominated by a contextually determined syncretic pressure to be investigated below, which will result in input /le/ yielding output [gli] in the appropriate contexts. This characterization is unaffected by the fact that in the present framework, faithfulness to an 'input' is in fact faithfulness to outputs, namely the surface occurrences of *le,* via the entailments they generate. The takeover by *gli* not only of feminine *le* but also of plural *loro* as in (9) will be amenable to a comparable explanation, based on the assumption that 'singular' is underspecified number, thus making *gli* but not *loro* prone to extension.

Along with extendability of the unmarked, the present approach also predicts extendability of a more general form, by making underspecification a consequence not only of markedness, but also of surface distribution. Consider in this regard the present tense paradigm of English *be*.

(11)   **Underspecification by aggregate surface distribution**

	SG	PL
1	am	
2	are	
3	is	

violated entailments:
are $\Rightarrow$ ~~1~~; $\Rightarrow$ ~~2~~; $\Rightarrow$ ~~3~~; $\Rightarrow$ ~~PL~~; $\Rightarrow$ ~~SG~~

Whatever the historical reasons behind the syncretism of the shaded area in (11), the present framework will automatically ascribe an underspecified status to the form *are* on account of the fact that each of the entailments it generates by occurring in some cell is violated by *are*'s occurrence in some other cell, as schematically indicated on the left in (11). Entailment strength, which is taken to be enhanced by validation but diminished by contradiction, thus provides a formal characterization of *are* as bearing no strong association with either any person or any number. In turn, this is then able to predict, based on the discussion of the asymmetry in (10), that the same form should be prone to further extensions, as is in fact the case with interrogative *aren't I?*, which will be briefly discussed below. One might suggest that *are*'s underspecification is the *cause*, rather than the *result* of the syncretism in (11), but we will see below that this interpretation is not sufficient in general, as it appears that in some cases an item's underspecification cannot be determined once and for all, but must rather be assumed to change as its aggregate surface distribution changes. The markedness-based underspecification is sufficient for the case in (10), however, as we have seen.

Because a more marked or more fully specified form is thus more stable under mutual attraction given its entailment structure, it seems appropriate, when characterizing syncretism as the result of attraction, to build the asymmetry into the notion of attraction, with the unmarked or more general form yielding as if drawn by a stronger attraction, as shown for the *gli/ le* pair in (12).

(12)

gli/ le	ATTR-1	ATTR-i
Extension of the less general/ more marked (*le* replaces *gli*)	*	
☞ Extension of the more general/ less marked (*gli* replaces *le*)		*

To summarize, the present framework predicts that syncretism should conform with the three generalizations in (13).

(13) a. Extension to an adjacent cell.
     b. More marked categories (subjunctive, plural, ...) are more affected.
     c. The more general or less marked is extended.

The correctness of (13a, b, c) is supported by cross-linguistic evidence as argued in Burzio and Tanta-

lou (2007). The same generalizations are also largely predicted by the influential 'Distributed Morphology' framework (Halle & Marantz, 1993; Harley & Noyer, 2003). The present interest in these generalizations, however, is rather the extent to which they are violated and the role of phonological constraints in those violations. Consider in particular, a situation characterizable by the ranking schema in (14), which I will refer to as 'Model A'.

(14) **Model A**

	PHONO	ATTR$_1$	ATTR$_i$	IO-FAITH
Ext. to nearest/ by more general or less marked	*		*	*
☞ Ext. to farther/ by less general or more marked		*		*

This would be a situation in which syncretism violates one or both of the generalizations (13a, c). It would follow from the fact that the syncretism ranking schema, represented by the last three columns in (14), which in itself would yield those generalizations, is dominated by phonological constraints, which impose a different choice. Another situation predicted by the present approach is characterized by 'Model B' in (15).

(15) **Model B**

	PHONO	IO-FAITH	ATTR$_1$	ATTR$_i$
No syncretism	*		*	*
☞ Ext. to nearest/ by more general or less marked		*		*

This is a situation in which syncretism *does* conform with the generalizations in (13), but in fact *only* obtains under phonological compulsion, being absent otherwise due to high-ranked IO-Faith. Finally, present predictions also include those of Model C in (16).

(16) **Model C**

	PHONO	IO-FAITH	ATTR$_1$	ATTR$_i$
No syncretism	*		*	*
☞ Ext. to farther/ by less general or more marked		*	*	

This model simply combines the properties of Models A and B. It characterizes a situation in which syncretism only obtains under phonological compulsion (Model B), and at the same time also violates either or both of (13a, c) (Model A).

We now turn to the first of the cases that call for such interaction with the Phonology and the inability of other approaches to deal with it.

4. Distributed Morphology and the 'Spurious *se*' rule.

In the framework of Distributed Morphology (DM; Halle & Marantz, 1993; Harley & Noyer, 2003) the principal device accounting for syncretism are the rules of 'Impoverishment' (RI), which manipulate morphosyntactic information prior the insertion of Vocabulary Items (morphemes). With one major caveat noted below, the RIs can successfully deal with the generalizations in (13), as briefly discussed next.

With regard to the adjacent-cell generalization in (13a), RIs can produce that effect in virtue of the fact that adjacent cells in a paradigm differ by a single feature. Impoverishing the morpho-syntactic representation by that feature will then result in adjacent cells necessarily receiving the same vocabulary item. For instance, the Italian case in (4a) could be handled by the rule in (17a) and the vocabulary in (17b) (based on (7)).

(17) a. Impoverishment: person $\Rightarrow$ Ø / ___ [SUBJ, SG]
     b. Vocabulary: -o $\Leftrightarrow$ 1;   -i $\Leftrightarrow$ Ø;   -a $\Leftrightarrow$ 3

After impoverishment, only the default form *-i* will be eligible for insertion, the other two items bearing incompatible person specifications. The restriction to adjacent cells comes from the structural description of the rule, identifying the subjunctive singular block of cells. The caveat about DM's ability to capture the

generalization in full can be revealed by reviewing the violating case in (4b) above, where 1st singular and 3rd plural are syncretic. One way to account for the latter syncretism would be to take the form *sono* to be a default form, like *-i* of (17b), and postulate two independent RIs: $1 \Rightarrow \emptyset / \_\_\_$ [SG]; $3 \Rightarrow \emptyset / \_\_\_$ [PL], each in the further context of the present indicative of 'be'. If this were the only means, the generalization would be captured, since syncretism over non-adjacent cells would now require more machinery (two RIs). There are, however, additional means. In DM, vocabulary could simply be defective, lacking, for instance, a 1SG item, thus allowing a default form *sono* to occupy the 1st singular cell without any RIs. A single RI could then extend that item to 3rd plural or for that matter to any cell, in violation of the generalization. In fact, no RI is needed at all, as the vocabulary could simply lack both a 1SG and a 3PL form, hence resorting to a default form in both cases. I will argue below that this ability of DM to bypass the adjacent-cell generalization is unwanted. As suggested above, violations such as those in (4b) are attributable to phonological similarity and should not be cause for relaxing the Morphology. This unwanted ability of DM, which comes from its freedom in setting up an underlying vocabulary, does not extend to the present framework, for the simple reason that the latter does not have underlying representations and hence cannot have an abstract 'underlying' vocabulary. For instance, in the case of Italian *sono*, once this form is utilized in the role of 1SG., it will generate entailments to that effect. And even though 'singular' is being taken as the lack of specification, potentially permitting extension to the plural, '1st person' is not, thus still preventing extension to the 3rd. At the same time, phonological similarity can play no role in DM, where Vocabulary Insertion occurs upstream of the Phonology.

Moving to the other generalizations, the one in (13b) that syncretism affects marked categories seems consistent with the DM machinery, under the same underspecification hypothesis of the present account. For instance, the RI in (17a) could have no counterpart in the indicative if the latter is just absence of mood. Reference to 'singular' in (17a) would be problematic in the face of underspecification, however, in contrast to the appeal to phonological similarity of the present analysis. Finally, concerning the generalization in (13c), DM will correctly predict that the more general form should extend, because being more general entails default status, which in turn is required for further extensions. If underspecification is adopted, then an unmarked form will also be a default form, hence correctly predicting extension of the unmarked. In sum, DM can correctly predict the generalizations in (13), except for the adjacent-cell generalization of (13a), which can be bypassed by postulating a defective underlying vocabulary. We will see that DM is also defective with respect to the notion of underspecification that was proposed for (11).

We are now finally in a position to consider phonologically induced violations of the generalizations in (13). Our first case is that of dative-accusative clitic clusters in the Romance languages. This is the context of the well-known 'Spurious *se*' rule of Spanish, informally given in (18).

(18)   Spanish DAT-ACC sequences: *le(s) $\Rightarrow$ se / ___ lo(s)/ la(s)

Within the DM approach, this has been analyzed as in (19) (Halle & Marantz, 1994).

(19)   Impoverishment: DAT $\Rightarrow \emptyset$ / ___ACC;   Vocabulary: *se* = Caseless

After removal of the Case specification by the RI, only reflexive *se* would remain eligible for insertion under the 'Caseless' characterization in (19). We will see that this account sheds no light on the several other repairs that are attested across Romance, however, including the one of Italian in (9), repeated in (20).

(20)   Italian DAT-ACC sequences: *le $\Rightarrow$ gli / ___ lo/ li/ la/ le

In Italian, only feminine clitic *le* is repaired in this context. Masculine *gli* is unaffected, yielding *glielo* [ʎelo] just the same (I will ignore the appearance of *e* in *glielo*). One could of course formulate a RI for Italian that would delete [GENDER] in this context, forcing insertion of unmarked/ genderless, but still dative *gli*. But this would establish no relationship between the two phenomena, leaving the virtual identity of the two input strings in both languages as a remarkable accident. Rather, what this comparison suggests is that the repair is induced by the OCP. Each language would then simply replace *le* with the nearest item in the paradigm (adjacent- cell generalization), as indicated in (21).

(21)

3rd person clitics		Spanish		Italian	
		MASC	FEM	MASC	FEM
ACC	SG	lo	la	lo	la
	PL	los	las	li	le
DAT	SG	le		☞ gli	le
	PL	les		loro	
LOC/ GEN				ci/ ne	
REFL		☞ se		si	

The adjacent-cell relation holds transparently in Italian, and holds in Spanish as well due to the absence in this language of either a locative or a genitive clitic, a point to which I return. The OCP-based account explains immediately why there is no repair of first or second person datives in either language, e.g. Italian/ Spanish *me lo*, although this is also consistent with the DM analysis, which regards not only *se* as Caseless, but also the 1st and 2nd person clitics. These are therefore also insertable after application of the RI in (19), though only in contexts that are specified for [person], where *se,* which is both Caseless and personless, is excluded. Note that the suggested violation of the OCP would not implicate adjacent segments as the two *l*'s are separated by one or more segments in the various cases. But this is consistent with other known OCP effects, e.g. the Latin liquid dissimilation case of *lun-aris* 'lunar' vs. *mort-alis* 'mortal' (Kenstowicz, 1994: 509).

A review of the rest of Romance, to which I turn below, confirms the OCP as the driving force, as virtually no language repairs this type of sequence unless the two clitics utilize the same consonant[4]. An account of this will then require Model B in (15) above, and seems precluded in a theory in which the Phonology as a whole is ordered after the Morphology like DM.[5] The present conclusion that the OCP is involved in the phenomenon in question concurs with Blasco (1984), Maiden (2000), Pescarini (2005), who, however, do not dwell on the theoretical consequences of interest here.

Turning then to a broader sample of languages, three different repairs are instantiated beside those in (21), as illustrated in (22) with examples from Italian dialects (variously borrowed from Calabrese, 1995; Maiden, 2000; Pescarini, 2005).[6]

(22) a. DAT ⇒ LOC (overwhelmingly more frequent)

(22) a.	DAT ⇒ LOC (overwhelmingly more frequent)	Napoli	*le lo	⇒ nče lo
		Arce (Frosinone)	*glie glie	⇒ če glie
		Logudoro (Sardinia)	* li lu	⇒ bi lu (bi = LOC)
b.	DAT ⇒ GEN (Partitive)	Grottaglie (Taranto)	*li lo	⇒ ni lo
c.	DAT ⇒ REFL	Campidanian (Sardinia)	*ddi ddu	⇒ si ddu

Note that in all the languages in (22) feminine and masculine dative clitics are independently syncretic (i.e. they lack the *le/ gli* distinction of Italian), hence switching to the masculine form as in Italian will not help. With the apparent exception of (22c), all cases in (22) seem to conform with the generalizations in (13a, c) in that they both instantiate syncretism of adjacent cells, and arguably also extend the less marked form. In this regard, we may consider the plausible feature decomposition in (23), based on Calabrese (1995).

---

[4] One exception is the Italian dative-genitive sequence *gliene*, where the dative is gender-ambiguous just as before an accusative. See Maiden (2000) for discussion of this case.

[5] J. Bobaljik (p.c.) points to the possibility of recruiting phonological constraints to act as filters on the operation of 'vocabulary insertion= in DM . This would make DM consistent with the text argument by effectively adopting some measure of parallelism. Other arguments presented below will continue to pose challenges to DM, however.

[6] As noted in Maiden (2000), in some languages the OCP violation is repaired not by syncretism but by allomorphy, i.e. by using a dissimilated variant of the same dative clitic. In a small group of languages, the OCP violation is tolerated.

(23)

	GEN	DAT	LOC	REFL
Possessor	+	+		
Location		+	+	

From this point of view, both genitive and locative would be equidistant from dative, accounting for the variation in (22a, b), and would also be less marked than the latter. The Spanish case in (21) would also resort to a less marked form and conform with the adjacent-cell generalization given the noted absence of either genitive or locative clitics. The surprising case is that of Campidanian in (22c), which replaces the dative with a reflexive despite the availability of both the genitive and the locative clitics. I discuss this case in the next section, where I argue that an independent pattern of syncretism brings the reflexive clitic within reach, so that the adjacent-cell generalization is in fact obeyed even by this case.

In contrast to the languages in (22), in languages in which dative and accusative forms are relatively different from each other in compliance with the OCP, no repair is attested. In most of these languages, however, and many Italian dialects, dative clitics are independently syncretic with either a locative or a genitive clitic regardless of the presence of an accusative. Examples involving an accusative clitic to illustrate compliance with the OCP are given in (24).

(24) a. ~~DAT~~ LOC

Piemontese	y lu	Campi	nči lu
Bari	nǧə lə	Brindisi	nči lu

  b. ~~DAT~~ GEN | Otranto | nde lu |

Like the cases in (22), those in (24) also conform with the adjacent-cell and the extension-of-the unmarked generalizations, for the same reasons.

In sum, in the Romance languages a dative clitic is replaced by some other clitic before an accusative only when a violation of the OCP would be incurred, demanding that Phonology participate in the calculation in the manner of Model B in (15). In addition, and pending a discussion of the Campidanian case in (22c), all replacements, whether just before an accusative or wholesale, conform with the adjacent-cell generalization consistently with the present analysis. Note that, while Model A in (14) above has the ability to induce syncretism across distant cells, this would be expected for the OCP cases under consideration only if none of the surrounding cells contained material satisfying the OCP –a fortuitous situation that has not been encountered.

5. Underspecified but not Underlying

The case of Campidanian in (22c) (and other neighboring dialects, according to Pescarini, 2005) raises the question of why the reflexive clitic is chosen instead of the locative or the genitive, which would seem closer (these are *nči/ nke; ndi*, respectively, according to Calabrese, 1995 and references). It turns out that the adjacent-cell generalization holds in this language as well, due to the fact that other instances of syncretism have extended the reflexive form to 1st and 2nd person plural both accusative and dative, as shown in (25).

(25) Campidanian plural clitics
    (from Calabrese, 1995; Blasco, 1986)

PL	3	2	1
ACC	ddus...		
DAT	ddis	si	

This means that, at least when the third person dative to be replaced is plural *ddis* of (25), the formerly reflexive *si* is indeed present in an adjacent cell. But once *si* is thus further extended to 3rd plural dative, it will once again be present in an adjacent cell when the dative to be replaced is singular *ddi* as in (22c). The question is then what mechanism would extends *si* to 1st and 2nd persons.

As noted in Blasco (1984), Calabrese (1995), Maiden (2000), Pescarini (2005), use of *si/se* as a 1st person non-reflexive clitic is quite common across Romance. As argued in Calabrese (1995), this is most plausibly attributable to the fact that the interpretation of this element bifurcates into reflexive and impersonal (Manzini, 1986; Burzio, 1992) and that the impersonal interpretation ('one') is rather close to 1st person ('we'). On this point, see also Blasco (1984), and for comparable 1st plural/ impersonal equivalence in other language families, see Baerman, 2005: fn. 4 and refs.). As argued in Burzio (1992), the scope of the reflexive/ impersonal divide can be captured by the simple preference in (26) in conjunction with some version of the 'Binding Theory'.

(26)   Reflexive ≻ Impersonal (Reflexive interpretation is better than Impersonal interpretation)

The preference in (26) correctly excludes an impersonal interpretation of *si* in all positions except those where a reflexive would be disallowed by the Binding Theory, namely as the subject of a tensed clause (cf. **He thought that himself would win*, etc.). Once the impersonal use of *si* is recognized, then the complex pattern of syncretism in Campidanian is again reducible to the adjacent-cell generalization, which obtains iteratively as shown in (27).

(27) **Campidanian syncretism**

(Shading: OCP-induced)

	3	2	1	0
ACC	PL / SG	⇐ / PL	⇐ PL /	⇐ IMP (=REFL)
DAT	PL ⇐ / ⇒ SG	⇐ / SG	/ SG	
LOC/ GEN				
REFL				

In particular, in the impersonal (zero person) use, *si* can be regarded as adjacent to 1st plural in virtue of the semantic affinity just noted, which is further established by the coreference possibility in the Italian example in (28), from Burzio (1992).

(28) Si_i è contenti quando ci_i scrivono   'One is happy when they write to us'
     IMP. is happy-PL when   to-us   they-write

In (28) *si* is the subject of the main clause, while the 1st person pronoun that refers back to it is the object of the lower clause. The felicitousness of such coreference is evidence that 1st person *ci* is semantically the form closest to impersonal *si*, another instance of *si* being predictably excluded in that position. In its impersonal reading, such object *si* would be blocked by (26), while in its reflexive/ reciprocal reading it would be allowed but only as referring to the local subject (clitics are never 'long distance' anaphors), *quando si scrivono* thus only meaning 'when they write to themselves/ each other'. The proximity of impersonal to 1st plural can be expressed by way of the decomposition in (29) based on Calabrese (1995), where the two would differ by just one feature.

(29) Semantic features	[speaker inclusive]	[participant]	[person]
Formal features 1	+	+	+
2		+	+
3			+
Impersonal *si*	+	+	

Once in the 1st plural cell in (27), *si* will in turn be positioned to take over 2nd plural, an extension attested in fewer languages but always entailing the first as per the adjacent-cell generalization. Note that the accusative-dative syncretism of 1st and 2nd person both singular and plural shown in (27) is general to the whole of Romance and thus pre-exists the extension of *si* to those cells, which just inherits it. Finally, once in the 1st and 2nd plural dative cells, *si* finds itself next to the 3rd plural cell, and from there next to the 3rd singular cell, hence extending to those two cells when required by the OCP as in (22c) above.

The fact that in (27) *si* would spread to an adjacent cell at each step, consistently with the generalization in (13a), follows from the entailment-based approach as argued above, as elements that occupy adjacent cells express maximally similar feature sets and are thus in a relation of maximal attraction, measurable as entailment violations. For example, occurrence of a distinct 1st person plural pronoun (descendant of Latin *nos*) will violate the set of entailments in (30) that are generated by impersonal *si* given the feature decomposition in (29), along with the fact that the impersonal reading is plural, at least in the languages in question.

(30) Violated by *nos:*   [+ speaker inclusive] $\Rightarrow$ *si*; [+ participant] $\Rightarrow$ *si*; [+plural] $\Rightarrow$ *si*

The reason for such violations is that *si* and *nos* overlap in each of the features on the left of the entailment sign. Once *si* takes over the function of *nos* as in (27) all such entailments will be satisfied. Conversely, use of impersonal *si* will also violate a comparable set of entailments generated by *nos*. Such violations can in fact make sense of the otherwise curious preference in (26). On the present analysis, impersonal *si* is a quasi-allomorph of the 1st person plural pronoun, and allomorphy is always disharmonic in the present framework due to entailment violations/ attraction. The neutralization of the *nos/ si* distinction in Campidanian and other languages resolves the attraction. The reflexive use of *si* mandated by (26) avoids it.

The fact that in (27) it is consistently *si* that spread to other cells in conformity with the generalization in (13c) is also in line with the present predictions. The table in (31) reviews the entailments that each extension of *si* violates, to consider whether these may be violated independently.

		Extension of *si* to:	Entailments violated
(31)	a.	1PL	$si \Rightarrow$ [- person]
	b.	2PL	$si \Rightarrow$ [+ speaker inclusive]
	c.	3PL	$si \Rightarrow$ [+ participant]
	d.	3SG	$si \Rightarrow$ [+ plural]

The entailments in (31) are those that *si* generates prior to each extension and that the extension contradicts. Putting aside the shaded rows for a moment, the entailments in (31a, b) are independently violated by the use of *si* as a 3rd person singular or plural reflexive element. By contrast, the forms *si* replaces (Latin *nos*, *vos*) generate entailments that are not violated independent of their competition with *si*. We are thus appealing here to the same notion of underspecification or weakness of specification that was characterized for the case of English *are* in (11) above, a notion based on aggregate surface distribution. We can characterize the direction of syncretic extensions in OT terms by considering that extending an underspecified elements will not violate that element's input specifications, while extending a fully specified one will. However, as I discuss below, what is critical to a characterization of the pattern in (27) is that the present notion of 'input' is non-distinct from the set of outputs that generate relevant entailments, so that, in turn, the notion of underspecification will not pertain to some 'underlying representation', which the present system does not have, but rather only to the aggregate surface distribution of the item in question. This conclusion is not changed by the fact that the entailment in (21a) may in fact not hold if [- person] is just lack of person specifications. Underspecification would be automatic in this case, rather than due to violation of the entailment, but the entailment in (31b) is not amenable to this reinterpretation, still requiring the surface-based notion of underspecification. Turning to the entailments in (31c,d), these are also violated by the reflexive uses of *si*, but the issue of underspecification as controlling directionality is moot in these cases, since here the OCP is at work. It is only by extending *si* to these cells that the OCP will be satisfied. Extending the OCP violators to other cells will not help.

In contrast to the above account of the iterative syncretism in (37), consider a serial framework in which underspecification is necessarily a property of underlying forms, like DM. I noted above that the DM machinery has the capability to extend a default form to any cell by postulating either some appropriate RI or an appropriately defective vocabulary (discussion of (4b)in sect. 4). That capability will be unwanted in (37), where the adjacent-cell generalization holds instead. In order to ensure that the first extension of *si* is to 1st plural consistently with the cross-linguistic evidence, this vocabulary item will have to bear some pre-specification, like '[+ speaker inclusive]; [+ participant]; [+ plural]' that restrict its compatibility to 1st plural

only (the RI could then be '[person] $\Rightarrow$ Ø/ ...' However, just the prespecification required to ensure the correctness of the first extension in (37), will automatically make all other extensions impossible. In particular, the specification '[+ speaker inclusive]' will prevent *si* from extending to 2nd and 3rd person; the specification [+ participant] will prevent it from extending to 3rd person; and [+ plural] will prevent it from extending to 3rd singular, no matter what the RIs may be. What would be needed in this system is thus for the underlying specifications of *si* to keep changing at each step. In particular, the specification [+ speaker inclusive] would need to be dropped just after the extension to 1st plural; the specification [+ participant] would have to be dropped just after the extension to 2nd plural, and [+ plural] would have to be dropped just after the extension to 3rd plural. These updates, that track the violations in (31), would be completely ad-hoc from this point of view, hence yielding no explanation for the resulting distribution. As we have seen, in the present system all updates are automatic given the surface distribution. Hence each of the specifications in (31) are already weak prior to each extension due to *si*'s use as a reflexive and hence prone to further violation. In addition, once an entailment like (31b) '*si* $\Rightarrow$ [+ speaker inclusive]' is violated by *si*'s extension to 2nd plural, it will no longer militate against *si*'s extension to 3rd plural. At the same time, all entailments in (31) nonetheless initially hold given *si*'s use as an impersonal element, making it extendable to 1st plural and from there on to other cells as discussed, always by minimal entailment violation at each step. Hence only surface distribution and not underlying representations can yield the correct underspecification at work in syncretism. Note as well that ascribing the 'impersonal' set of features of (30) or (31) to the underlying form of *si* would in any event be incorrect for other reasons. Impersonal *si* is an allomorph of reflexive *si*, which surely cannot share those features. Hence impersonal features must be assigned by a general procedure that applies only where reflexive interpretation fails, as determined by (26) and the Binding Theory. This explains why the same impersonal features are present in other elements as well, including Italian possessive *proprio*, and 'Arbitrary' PRO (Burzio, 1992) –a further accident, if these features were just 'underlying'.

In sum, the complex pattern of syncretism of Campidanian Sardinian was shown to reduce to the usual generalizations, which are satisfied at each iteration. This iterative character requires that the properties associated with the extending element change at each step. This is true by definition if all associations are defined over surface forms, but not if they are defined once and for all over some 'underlying' vocabulary. It is easy to see that the hypothesis that there exist such underlying forms is a mere contingency of the serial architecture, to which the parallel one being advocated here thus proves again superior.

6. Retreat *from* the general. Italian *cisi*, French *bel ami*, Armenian possessives.

We have seen that replacement of dative clitics before accusatives in the Romance languages combines phonological and morphological effects. The Phonology steps in asserting the OCP and the Morphology provides alternatives, ranked in order of distance and underspecification. The case of English *aren't I?* discussed in Bresnan (2001) can be analyzed in the same vein. Whereas Bresnan postulates a 'lexical gap' for the absence of **amn't I?*, it seems plausible to simply invoke the Phonology as excluding the cluster **mnt*, as independently proposed in Dixon (1982). On this view, this case would also invoke Model B in (15), like those of the previous sections. By contrast, the next few cases will require Model C in (16), in that they violate the extension-of-the general, or the 'retreat to the general' pattern as it is referred to in DM work (Noyer, 1998).

One such case is that of Italian *cisi* , illustrated in (32).

(32) a. Si$_i$ sperava di vedersi$_i$
　　　　IMP. hoped   to see-each-other　　'We hoped to see each other'
　　b. Ci$_i$ si$_i$ vede   (*si si)
　　　　us   IMP.   see　　　　　　'We see each other'

The case in (32a) shows that impersonal and reflexive *si* can co-occur and be coreferential, while (32b) shows that they cannot cooccur adjacent to one-another. We take this to be another OCP effect, and the repair to be akin to the one in (28) above, where 1st plural *ci* also replaces *si*. This repair is exactly in the opposite direction as the one in (27), however, where *si* replaces 1st plural. Note that the adjacent-cell generalization continues to hold, since a reflexive coreferential with impersonal *si* would carry the latter's fea-

tures, which we have seen place it next to 1st plural. However, the conflicting directions will challenge the DM machinery. The first step in (27) could be accounted for with a RI as we have seen, but then its mirror-image in (32) will now require the opposite kind of rule, namely a rule of 'enrichment', inserting 1st plural features. This is indeed the solution proposed in Bonet (1991: 149) within DM, but it provides little under-standing of such departures from the usual pattern, characterizing them as arbitrary. In contrast, Model C in (16) above attributes them to independent phonological principles, under appropriate ranking. Note that, in addition to (32b), the case in (28) would also require a rule of enrichment, but a different one given the dif-ferent context. In the present system, the two cases result from the different factors interacting with the mor-phology –the OCP in (32b), and the Binding Theory in (28), both established independently. Note as well that while the Binding Theory is satisfied in (28), causing the syncretism, it is violated in (32b) due to the locally bound pronoun, indicating that the latter is also a system of violable constraints, as argued in Burzio (1991, 1998), like the Phonology and the Morphology.

Steriade (1999b) refers to the phenomenon illustrated for French in (33) as 'Lexical Conservatism'.

(33)   */bo ami/ $\Rightarrow$ [bɛl ami]   'beautiful friend-MASC.'

According to Steriade, in (33), the feminine allomorph [bɛl] is pressed into action in a semantically masculine context so as to avoid a vowel hiatus.[7] The Phonology is thus being 'lexically conservative' be-cause, rather than creating new structures via epenthesis or deletion, it recruits an independently existing allomorph. It is easy to see that Steriade's 'Lexical Conservatism' is closely equivalent to the present de-scription of phonologically controlled syncretism, according to which the Morphology supplies the candi-dates and the Phonology chooses. A purely morphological account of (33) will in fact face exactly the same issues as one of (32b), as feminine gender is clearly more marked than masculine.

Like the cases in (32b) and (33), the Armenian case in (34) also involves extending the use of a more specific or marked morpheme.

(34)   / __ -eɾ (PL)   = #σ

		cow	cows	cat	cats
/ __ -neɾ (PL) = σ σ	X	gov	gov-eɾ	gadu	gadu-neɾ
/ __ -ni (POSS) = σ σ	our-X	gov-eɾ-ni-s		gadu-ni-s	gadu-neɾ-ni-s

As discussed in Vaux (2003), Armenian has two plural allomorphs, -eɾ, used with monosyllabic stems, and -neɾ, used with plurisyllabic stems, as indicated in the left side of (34). In addition, it also has a posses-sive suffix -ni, historically related to -neɾ and sharing with it the requirement of a plurisyllabic stems, as also indicated. The fact of interest is that when -ni is combined with a stem that would otherwise be monosyl-labic, the latter receives an augment -eɾ to satisfy -ni's plurisyllabicity requirement. The latter augment is of course just the plural morpheme, which is thus used regardless of singular meaning, as shown in the shaded area in (34). This lexical conservatism is parallel to that of the French case in (33). Once again, a marked form, here normally plural -eɾ, takes over the function of a less marked form, namely the singular, normally zero. The only difference is that in the Armenian case the condition behind the syncretism is effectively morphological, in the form of the plurisyllabicity requirement, while in the French case it was purely phono-logical, as hiatus avoidance, or the OT constraint ONSET. Nonetheless, the plurisyllabicity requirement of (34) still employs a phonological vocabulary (reference to syllables), hence still requiring that the Phonol-ogy participate in making morphological choices, as in the other cases discussed so far.

In sum, we have seen cases that subvert the retreat-to-the general pattern, under conditions that are clearly phonological or phonologically definable in each case, requiring parallel interaction of Morphology

---

[7] A referee points to data like *j'en ai un bel ‖ , éléphant* 'I have a good one (pause) elephant' as indicating that such adjective allomorphy is lexicalized (*éléphant* selects *bel* lexically), with phonology only providing the general basis for the lexicalization. This is an important qualification, but this type of dynamics is in fact consistent with the present approach, in which higher-order units such as plurimorphemic words and -by the same token- phrases can be lexical-ized without denying a role to the grammar (parallel relation of lexicon and grammar). Under the REH in (2), a repre-sentation like *bel éléphant* 'nice elephant' favored by the phonology (grammar) would automatically result in an asso-ciation between the noun *éléphant* and the adjectival allomorph *bel* (lexicon). The latter association can then account for the case the referee notes.

and Phonology again. While earlier cases had invoked Model B of (15) above, the present cases invoke Model C of (16) by which the Phonology both causes the syncretism and at the same time reverses the retreat-to-the general pattern. I am not aware of cases that would instantiate Model A above, by which the Phonology would merely change the pattern of some syncretism that occurs independently, but the properties of Model A are in any event subsumed by those of Model C as was noted earlier.

7. The bigger picture

The phenomenon of phonologically conditioned syncretism discussed above joins a gallery of other phenomena discussed in Burzio (2005b, 2006), that also demonstrate the parallel character of Morphophonology. One of these is morphological selection. Like syncretism, the latter can also be phonologically conditioned. We have seen this to be the case with Armenian plurals in (34) above. In this regard consider in particular the alternation in (35).

(35)   mɑnəɾ '*small thing*'   mɑnɾ-əɾ '*small things*'   (from Vaux, 2003)

To account for the fact that an epenthetic [ə] is inserted in the singular but not in the plural form in (35), Phonology must apply to the output of the Morphology once the latter has attached -*eɾ*. Yet, since plural -*eɾ* is only attached to monosyllables as indicated in (34), the Morphology must at the same time apply to the output of the Phonology, which adds a syllable only in the singular. However, both of these dependencies can only hold if both Phonology and Morphology calculate the output form simultaneously. Another case of this type is that of English suffix -*en* as in *sof(t)-en,* which attaches only to stems that end in a single obstruent (cf. **lax-en*). Deletion of the stem-final *t* is thus necessary to satisfy morphological selection, and yet that deletion is clearly phonological and requires the presence of the suffix to take place, cf. the adjective [sɔft]. Booij (2002: 227) reaches similar conclusions in his analysis of Dutch morphology, stating that '...it is crucially the set of ranked output conditions of the phonological module ...that makes the correct [morphological (LB)] choice.'

Just as the output of the Phonology can control morphological selection, so morphological selection can control the output of the Phonology. Consider the rampant morphological irregularity of the English Latinate/ 'Level 1' lexicon, as in *crimiN-al*, compUlS-*ive*, *ARBORE-al* where the capitalized portions do not correspond to base material. As argued in Burzio (2002b), such morphological irregularity can be ascribed to a relatively low rank of the relevant selection constraints like the one in (1a) above, hence allowing for misidentities between stem and base. However, the same affixes that tolerate such irregularities also accept regular phonological processes like re-stressing, vowel shortening and velar softening (e.g. *parént-al*; *cr[ɪ]min-al*; *criti[s]-ism*), while the affixes of the Germanic/ 'Level 2' class both enforce morphological regularity and block regular phonological processes (e.g. *éffortless-ness*; *cr[ay]me-less*; *froli[k]-ing*). Short of duplication of resources, this only follows if morphological selection, which evidently has different degrees of rigidity for the two classes, is enforced in parallel with the constraints of the Phonology.

The present claim that Morphology operates in parallel with the Phonology entails that the Morphology itself is a parallel, constraint-based system, leading to the expectation of some purely morphological evidence to that effect. Aranovich (2005) observes in fact the following pattern in Erza Mordvin.

(36) Pattern of syncretism in Erza Mordvin (Uralic).

	OBJ. SG.	OBJ. PL
SUBJ. SG.		
SUBJ. PL.		

 (from Aranovich, 2005)

This language has affixes that express features of the subject and of the object simultaneously as suggested by the structure of the table in (36), although each of the cells in that table stands for a block of cells that still differ by the person of both subject and object. What is of relevance here is the fact that the three blocks of cells identified by the shading are identical. Hence object number is syncretic in the presence of a plural subject, as is subject number in the presence of a plural object. Consider then the RIs that would be required within in the DM framework:

(37) **RI 1**: Subject-PL $\Rightarrow$ Ø / __ Object-PL;   **RI 2**: Object-PL $\Rightarrow$ Ø / __ Subject-PL

As Aranovich argues, it is easy to see that the rules in (37) cannot be ordered, as they destroy each-other's context of application. Such interaction provides a textbook argument for parallelism, in this case applying to the theoretical resources involved in the characterization of syncretism even aside from their interaction with Phonology. In the present framework, the pattern in (36) is straightforwardly expected, since the cells involved are adjacent. In particular, following the model of (10) above in which less marked *gli* replaces more marked *le,* we will expect the material in the corner cell in (36) to come to be replaced by the material in the cell just above it, and at the same time also by the material in the cell to its left, resulting in the same material occupying all three cells. Finer-grained considerations will determine which of the two morphemes competing for the take-over of the corner cell will prevail, but that is beyond present concerns.

The same conclusion that Morphology has an internally parallel organization is also suggested by Booij's (this volume) observations concerning cases like (38a, b).

(38) a.   de-caffein-ate (*de-caffein; *caffein-ate)   de-V + [N-ate]$_V$ = de-N-ate
     b.   de-moral-ize (*de-moral; *moral-ize)   de-V + [A-ize]$_V$ = de-A-ize

In (38), while both *decaffeinate* and *demoralize* can be attributed to regular morphological processes, they cannot be produced by applying one process at a time, as for instance neither **decaffein* nor **caffeinate* exist. Rather, in each case, both the schema that yields verbs by affixing *-ate* or *-ize*, and the one that produces a verb from another verb by prefixing *de-* must apply simultaneously, as if the two joined forces, inconsistently with sequential application. Space limitations preclude an explicit account within the present framework, but Booij's discussion provides one within an alternative parallel system.

In sum, the phenomenon of phonologically conditioned syncretism that forms the topic of the present article appears to be the tip of a larger iceberg of evidence that Phonology and Morphology interact in paral-lel. In order for this to be possible, Morphology itself needs to have a parallel architecture, for which we have seen there is independent evidence.

8. Conclusions

I have argued that, within the 'Representational Entailments Hypothesis' of Burzio (2002a, b), morpho-logical syncretism is analogous to neutralization of phonemic contrasts, both cases being characterizable in terms of attraction between neighboring representations. This characterization appears to correctly predict certain tendencies observable cross-linguistically. At the same time, because of its parallel architecture, the same approach can also account for the fact that those tendencies can be subverted when phonological con-straints make competing demands. Evidence outside the realm of syncretism also supports the same architec-ture.

The present approach to syncretism is similar to that of recent work within OT, such as Grimshaw (2001), Bresnan (2001), Aranovich (2005). It differs from those works in taxing similarity rather than mark-edness. While the markedness approach may seem directly supported by the noted fact that marked forms are replaced by unmarked ones, we have seen that the latter generalization is also captured by the present approach via underspecification and its effect on entailment violations. An argument in favor of the present approach, which space prevents me from developing in detail, is that phonological similarity seems to be behind certain cases of syncretism, as suggested for (4a, b) above. In addition to (4a, b), phonological simi-larity also seems at play in the historical development of Romance clitics, aspects of which were discussed above. In contrast to the case in (27), in some dialects 1st plural is taken over by genitive clitic *ne/ nde*. While Calabrese (1995) regards this as an unsolved problem, both Maiden (2000) and Pescarini (2005) point to the similarity of this clitic to original 1st plural *nos*, no language extending *ne/ nde* to sharply dis-similar 2nd plural *vos*. On the other hand, *vos* is taken over by locative *ibi* $\Rightarrow$ *(i)vi* in Italian and other lan-guages, consistently with their relative similarity. Extension of the alternative locative *(i)nci* $\Rightarrow$ *ci* to 1st plu-ral *nos* in Italian also finds some role for similarity, as does, according to Blasco (1984), the replacement of *nos* by *se* in languages in which the latter form reduces to *os*. Markedness will shed no light on any of these cases, while similarity, either morphological or phonological, thus seems applicable to all cases.

References

Aranovich, Raúl. 2005. *Optimizing Verbal Agreement in Mordvin.* ms. UC Davis http://linguistics.ucdavis.edu/FacultyPages/raranovich/

Archangeli, Diana and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge: MIT Press.

Aronoff, Mark. 1994. *Morphology by Itself*. Cambridge: MIT Press.

Baerman, Matthew. 2005. Typology and the formal modelling of syncretism. In G. Booij, J. van Marle, eds. *Yearbook of Morphology 2004.* Dordrecht: Kluwer. 41-72.

Becker, Thomas. 1993. Back formation, cross-formation and 'bracketing paradoxes' in paradigmatic morphology. In G. Booij, J. van Marle, eds. *Yearbook of Morphology 1993*. Dordrecht: Kluwer. 1-25.

Blasco, Eduardo. 1984. L'estensione del pronome riflessivo SE in Sardo e nelle lingue romanze. *Revue de linguistique romane* 43:12-35.

Blasco, Eduardo. 1986. *La lingua sarda contemporanea*. Cagliari: La Torre.

Blevins, James P. 2005. Word-based declensions in Estonian. In G. Booij and J. van Marle, eds. *Yearbook of Morphology 2005*. Dordrecht: Kluwer. 1-25.

Bochner, Harry. 1993. *Simplicity in Generative Morphology*. New York: Mouton de Gruyter.

Bonet, Eulàlia. 1991. *Morphology after Syntax*. Ph.D. Dissertation. Cambridge: MIT.

Booij, Geert. 2002. *The Morphology of Dutch*. Oxford: Oxford University Press.

Booij, Geert. this volume. Construction morphology and the lexicon.

Bresnan, Joan. 2001. Explaining morphosyntactic competition. In M. Baltin and C. Collins, eds. *Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell. 11-44.

Burzio, Luigi. 1991. The morphological basis of anaphora. *Journal of Linguistics* 27:81-105.

Burzio, Luigi. 1992. On the morphology of reflexives and impersonals. In C. Lauefer and T. Morgan, eds. *Theoretical Analyses in Romance Linguistics (LSRL XIX)*. Amsterdam: John Benjamins. 399-414.

Burzio, Luigi. 1998. Anaphora and soft constraints. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, D. Pesetsky, eds. *Is the Best Good Enough? Optimality and Competition in Syntax*Cambridge: MIT Press and MIT Working Papers in Linguistics.

Burzio, Luigi. 2002a. Surface-to-surface morphology: when your representations turn into constraints. In P. Boucher, ed. *Many Morphologies*. Somerville: Cascadilla Press. 142-177.

Burzio, Luigi. 2002b. Missing Players: Phonology and the Past-tense Debate. *Lingua* 112:157-199.

Burzio, Luigi. 2005a. Sources of paradigm uniformity. In L. J. Downing, T. A. Hall, R. Raffelsiefen, eds. *Paradigms in Phonological Theory*. Oxford: Oxford University Press. 65-106.

Burzio, Luigi. 2005b. Lexicon and grammar: unequal but inseparable. ms. JHU. http://www.cog.jhu.edu/faculty/burzio/LexGramPaper05.pdf

Burzio, Luigi. 2006. Lexicon versus grammar in English morphophonology: Modularity revisited *Korean Journal of English Language and Linguistics* 6-3:437-464.

Burzio, Luigi and Niki Tantalou. 2007. Modern Greek accent and faithfulness constraints in OT. *Lingua* 117/6:1080-1124.

Bybee, Joan. L. 1988. Morphology as lexical organization. In M.Hammond and M. Noonan, eds. *Theoretical Morphology*, San Diego: Academic Press. 119-141.

Calabrese, Andrea. 1995. Syncretism phenomena in the clitic systems of Italian and Sardinian dialects and the notion of morphological change. *Proceedings of the North East Linguistic Society 25*, vol. 2. Amherst: GSLA.151-174.

Dixon, Robert M.W. 1982. Semantic neutralization for phonological reasons. In R.M.W. Dixon. *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Berlin: Mouton Publishers. 235-238.

Grimshaw, Jane. 2001. Optimal clitic position and the lexicon in Romance clitic systems. In G. Legendre, J. Grimshaw, S. Vikner, eds. *Optimality-Theoretic Syntax*. Cambridge: MIT Press. 205-240

Halle, Morris and Alec Marantz. 1993. Distributed Morphology. In K. Hale, S. J. Keyser, eds. *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge: MIT Press; 111-176.

Halle, Morris and Alec Marantz. 1994. Some key features of distributed morphology. In A.Carnie, H. Harley, eds. *MIT Working Papers in Linguistics 21*. Cambridge: MIT Press. 275-288.

Harley, Heidi and Rolf Noyer. 2003. Distributed Morphology. In L. Cheng, R. Sybesma, eds. *The Second GLOT International State-of-the-Article Book*. New York: Mouton de Gruyter. 463-496.

Hayes, Bruce, Robert Kirchner and Donca Steriade, eds. (2004). *Phonetically Based Phonology*. Cambridge: Cambridge University Press.

Hebb, Donald O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley and Sons.

Kenstowicz, Michael J. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Maiden, Martin. 2000. Phonological dissimilation and clitic morphology in Italo-Romance. In L. Repetti, ed. *Phonological Theory and the Dialects of Italy*. Amsterdam: John Benjamins. 169-189.

Manzini, Maria Rita. 1986. On Italian *Si*. In H. Borer, ed. *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. New York: Academic Press. 241-262.

Matthews, Peter H. 1972. *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.

Matthews, Peter H. 1991. *Morphology*. Cambridge: Cambridge University Press.

McCarthy, John J. 2005. Optimal paradigms. In L. J. Downing, T. A. Hall, R. Raffelsiefen, eds. *Paradigms in Phonological Theory*. Oxford: Oxford University Press. 170-210.

Noyer, Rolf. 1998. Impoverishment theory and morphosyntactic markedness. In S.G. Lapointe, D.K. Brentari and P.M. Farrell, eds. *Morphology and Its Relation to Phonology and Syntax*. Stanford: CSLI. 264-285.

Pescarini, Diego. 2005. Types of syncretism in the clitic systems of Romance. Paper presented at BIDE '05. Bilbao-Deusto University. *International Journal of Basque Linguistics and Philology*, in press.

Pirrelli, Vito and Marco Battista. 2000. The paradigmatic dimension of stem allomorphy in Italian verb inflection. *Rivista di Linguistica*. 12.2:307-380.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Report no. RuCCS-TR-2. New Brunswick: Rutgers University Center for Cognitive Science. Published (2004). Oxford: Blackwell.

Steriade, Donca. 1995. Positional neutralization. ms. UCLA

Steriade, Donca. 1999a. Phonetics in phonology: the case of laryngeal neutralization. In M. Gordon, ed. *Papers in Phonology 3* (UCLA Working Papers in Linguistics 2). Los Angeles: Department of Linguistics, UCLA.

Steriade, Donca. 1999b. Lexical conservatism in French adjectival liaison. In B.Bullock, M. Authier, L. Reed, eds. *Formal Perspectives in Romance Linguistics*. Amsterdam: John Benjamins. 243-270.

Vaux, Bert 2003. Syllabification in Armenian, Universal Grammar and the lexicon. *Linguistic Inquiry* 34.1:91-125.

Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse

edited by Fabio Montermini, Gilles Boyé, and Nabil Hathout

Cascadilla Proceedings Project   Somerville, MA   2007

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Burzio, Luigi. 2007. Phonologically Conditioned Syncretism. In *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, ed. Fabio Montermini, Gilles Boyé, and Nabil Hathout, 1-19. Somerville, MA: Cascadilla Proceedings Project.

or:

Burzio, Luigi. 2007. Phonologically Conditioned Syncretism. In *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, ed. Fabio Montermini, Gilles Boyé, and Nabil Hathout, 1-19. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1611.