11

On the distribution of stem alternants

Separation and its limits

DAVID EMBICK

11.1 Introduction

A primary goal of grammatical theory is to account for the systematicity of form/meaning connections. It is because of this that deviations from 'one-to-one' connections between syntax/semantics and phonology, of the type found with allomorphy and syncretism, are of central interest to morphological theory. In the domain of allomorphy broadly construed, a number of questions surround the analysis of stem alternations. As part of a more general theory of alternations in the grammar, the questions raised by stems are, broadly speaking, of two types. The first (Q1) is *the question of the relation between alternants*: in particular, whether different stem forms are derived from a single underlying phonological representation, or whether they exist individually in memory as suppletive alternants. Since mechanisms that effect both phonological alternations and suppletive allomorphy have independent theoretical motivation, stem changes could in principle be treated in either way. The second (Q2) is *the question of how stem distributions are accounted for*. In one type of theory, stem alternations are determined *contextually*, such that a given stem alternant is employed only when its local context (defined in terms of morphemes, or phonological representations) derives that alternant, or conditions its insertion. In a second type of theory, stem alternants have *paradigmatic* distributions, such that particular stems that are derived or stored 'offline' are associated with particular sets of features (=paradigmatic cells) by an independent mechanism of stem distribution.

Both (Q1) and (Q2) implicate the general question of how systematic (or unsystematic) relations between form and meaning may be. An important idea that has been incorporated into many different theories of morphology is that form/meaning connections must be relaxed in ways that depart from a 'one form, one meaning' ideal. In terminology that is associated with Beard (1966) and related work (with

The Morphome Debate. First edition. Ana R. Luís and Ricardo Bermúdez-Otero (eds)

further connections with Chomsky (1965) and Matthews (1965)), many theories of grammar adopt the *Separation Hypothesis*: the idea that the sound (or, more generally, the formal) component of the morpheme is separated from its syntactic and semantic (shorthand: synsem) components. For Separationist theories, formal and synsem features do not originate in a single primitive object. Rather, phonological features are added to collections of synsem features—i.e., to morphemes, in the view adopted here—so that the synsem features are said to be *realized* phonologically. For this reason, theories with Separation are also called *realizational* theories.

A crucial question for Separationist theories is exactly how much separation is required. A large amount of work associated with Distributed Morphology explores different ways of constraining Separation. The theory implements Separation in the form of a Vocabulary Insertion process that adds phonological content to morphemes. The Vocabulary Items that accomplish this may potentially be underspecified with respect to the morphemes that they apply to. Underspecified Vocabulary Items are motivated empirically: they provide a way of accounting for syncretisms systematically (e.g. Bobaljik (2002), Embick and Noyer (2007) for overview discussion). Allowing underspecification allows for form/meaning connections that are less transparent than those that are allowed in theories with 'traditional' morphemes, in which sound and meaning are present from the beginning.

Admitting underspecification into the theory, while an important move architecturally, is not an end in itself; rather, it is the starting point for a further set of questions. These questions fall under two broad headings. The first is whether the move to Vocabulary Insertion is sufficient to explain patterns of syncretism, or whether other mechanisms (e.g. Impoverishment; Noyer (1992,1998) and related work) are required. The second (more general) question is how to allow Separation while maintaining a theory in which form/meaning connections are as transparent as possible in the normal case. The latter question is at the center of a research programme that investigates the mechanisms responsible for different types of 'mismatches' between the synsem and morphological and phonological parts of grammar; see e.g. Embick and Noyer (2007) for an overview.

The idea that languages employ *morphomic stems*, which is advanced in Aronoff (1994) and related work, is directly relevant to Separation and its limits. Aronoff argues that certain patterns of stem distributions—*morphomic distributions*—require a theory that allows for a further type of separation between form and meaning: in particular, a type in which the derivation of stem alternants is separated from the procedure that accounts for the distribution of stems. I will address two aspects of this position below, and argue that the separation of stem derivation from stem distribution is both undesirable and unnecessary. In particular:

Point 1: Separating stem derivation from stem distribution is problematic because it precludes any straightforward account of contextual locality effects (trigger/target relations) in the theory of stems.

Point 2: Separating stem derivation from stem distribution is unnecessary, because it is possible to account for the relevant patterns without making this move.

After the theoretical significance of separating stem derivation from stem distribution is outlined in section 11.2, section 11.3 sets out the limits to Separation that are posited in a current version of Distributed Morphology, and reviews some current hypotheses on the locality conditions on allomorphy. Articulating Point 1, which is the main focus of the paper, section 11.4 examines stem alternations of two types: one type that involves suppletive contextual allomorphy, and one that involves a (morpho)phonological derivation of one alternant from another. The argument is that these alternations are constrained by contextual locality conditions that follow from a theory in which stem derivation and stem distribution are accounted for together, but not from a theory in which derivation is separated from distribution. On Point 2, section 11.5 argues that even if truly morphomic distributions of stems do exist, there is still no need to separate derivation from distribution in order to account for their properties. Section 11.6 presents some general conclusions.

11.2 Separation and stem distribution: two approaches

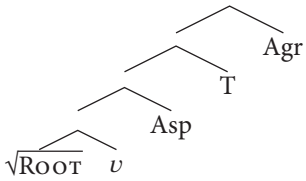
In the theory of Distributed Morphology, which is piece-based and realizational, the syntax manipulates morphemes that consist of semantic and syntactic (synsem) features, with no underlying phonological content (see Halle and Marantz 1993, Embick and Noyer 2007). In the PF component of the grammar, the operation of Vocabulary Insertion provides morphemes with phonological content. So, for example, in English the plural morpheme [+pl] has the Vocabulary Item [+pl] ↔ /z/ applied to it, such that it receives the phonological form /z/. The phonological form of a Vocabulary Item—/z/ in this example—is called a phonological *exponent*.

As mentioned in the first section, adopting Separation in the particular form just outlined allows Vocabulary Items to be *underspecified* with respect to the morphemes that they apply to. With underspecification, a single Vocabulary Item can potentially apply to a number of different morphemes. When this occurs, morphemes that are distinct in terms of synsem feature content are realized with the same phonological exponent. This important departure from simple, one-to-one form/meaning connections is motivated by the need to analyse *syncretism* in a systematic way (see the references cited immediately above and in section 11.1 for overview discussion).

For the analysis of stem alternations, the theory makes available two options: Vocabulary Insertion, which is responsible for suppletive contextual allomorphy, and, in addition morphologically-conditioned phonological rules. Thus, in principle each of the two answers to (Q1) above could be used to analyse the relation between stem forms. To be more specific about these options, it is convenient to make reference to the way in which morphemes are realized formally. This occurs after

they have been composed into complex objects. For illustration, (1) shows a schematized verb, consisting of a Root along with a verbalizing (v), Asp(ect), T(ense), and Agr(eement) morphemes:[1]

(1) A 'verb'



It is in terms of such structured complexes of morphemes—and in terms of the linear representations derived from them in the phonological component—that the different options for treating stem allomorphy must be assessed. A main theme of much work, which is reviewed in section 11.3 below, is that suppletive contextual allomorphy (=Vocabulary Insertion) and morphophonology (=morphologically conditioned phonological rules) might be subject to different locality conditions. Crucially, though, both types of alternation must respect *some* set of locality relations between the trigger and the target of the effect. Speaking generally, this type of theory thus holds that stem alternations occur when a Root occurs in the context of local synsem features that trigger contextual allomorphy or a morphophonological rule. For this reason, I will refer to all theories in which distribution is determined in this contextual, 'online' way as Derivation/Distribution (DD) theories:

(2) Derivation/Distribution Theory: Roots appear in structures that are derived by composing them with other morphemes. Stem alternants are derived in such structures, in ways that are restricted by locality relations between the features/morphemes that are the triggers of allomorphy, and the Root that is the target of the change.

An important aspect of (2) is that it makes crucial reference to morphemes. The motivation for this restriction will become clear later in this section. While I will be assuming a specific morpheme-based DD theory in this paper, it is important to emphasize that the theories that fall under (2) can differ in terms of a number of points; as long as they do *not* separate the mechanisms that relate stem alternants

[1] See e.g. Embick and Marantz (2008) for the approach to Roots and their categorization that is assumed here.

The terminology used in the discussion of (1), particularly with respect to the use of *Root*, is potentially complicated. In particular, I will not examine the question of whether Roots are subject to the Vocabulary Insertion operation, or whether Vocabulary Insertion applies only to functional morphemes. Although it is of some relevance to the analysis of stems broadly speaking, the question of Root-insertion is not crucial for the point in the main text about contextual conditioning.

from those that distribute stems, they count as DD theories for my purposes (see the end of this section for some additional discussion).

In a paradigmatic theory of the type advanced in Aronoff (1994) and related work, on the other hand, derivation and distribution are not linked in the way defined in (2). Instead, a particular stem alternant that is represented 'offline' is associated with a particular distribution that is defined in terms of paradigmatic cells. Schematically, this type of theory has two components. The first accounts for relations between stem alternants:

(3) Paradigmatic Theory, Component 1 ('Stem System'): For a Root with a stem X, a stem alternant X′ can be either:
   a. derived by rule from X; or
   b. stored as an unanalysed object, such that the Root may be said to possess the *stored stems* X and X′.

The stem alternants that are associated with a Root by (2) have a status in the grammar that is independent of their use in a particular derivation. Thus, in order to distribute stems in the appropriate way, the grammar must also contain a system that specifies how specific stems relate to particular combinations of features:[2]

(4) Paradigmatic Theory, Component 2 ('Distribution System').
   a. Feature system defines a set of paradigmatic cells (i.e. a 'paradigm space')
   b. Individual stem forms are associated with sets of cells; e.g., 'stem X′ is used in cells with the following feature combinations:…'

This type of theory employs two types of Separation. First, form is separated from meaning, as it is in realizational theories in general. In addition, though, the type of theory just outlined holds that the mechanisms for deriving/storing forms are separated from the contexts in which those forms are distributed. With this in mind, the designation PSD for 'Paradigmatic Stem Distribution' will be employed for theories of this type.

The DD and PSD theories are both in principle capable of saying the same things about (Q1), concerning how stems are related to each other and to Roots, i.e. each type of theory can employ both suppletion and derivation by rule. However, although the two theories are similar on this point, they differ in important ways in terms of how stem forms relate to synsem features. In a DD theory, stem forms are produced when the target of a change (i.e. the Root) is in a local relationship with the trigger of the change. Thus, as emphasized above, stem distribution is predicted to be constrained in ways that reflect morphological, phonological, and syntactic locality constraints.

[2] Aronoff (1994, 2012); see also Maiden (1992, 2005) and Stump (2001).

In contrast, a PSD theory employs stem alternants that are created in ways that are independent of any relation with a particular trigger (set of synsem features). The different stem forms derived in the 'Stem System' are specified to appear in certain paradigmatic cells which, in a theory that does not have morphemes—as is assumed by Aronoff and others—are feature matrices (=paradigmatic cells) consisting of a Root and synsem features. This type of object is shown in (5), where $\pm X$ etc.; in, for example, a verb these stand in for features like aspect, tense, agreement, and so on (recall by way of contrast (1) above):

(5) Representation of a paradigmatic cell

$$\text{Root} \begin{bmatrix} \pm X \\ \pm Y \\ \pm Z \end{bmatrix}$$

Since the derivation of stem forms is independent of the assignment of a particular stem to a set of such matrices, the distribution of stem forms is not expected to be affected by any type of locality relations. Rather, any stem could in principle be specified to occur in any set of paradigmatic cells. Another way of thinking of this is to note that to the extent that there is a substantive PSD theory of stem distributions—i.e. something more than 'anything goes'—it would have to be paradigmatic in nature, not syntagmatic.[3] Thus, if languages show stem distributions that are restricted in ways that are expected syntagmatically—i.e. distributions that always respect locality conditions that derive from syntactic, morphological, and phonological representations—this would be evidence in favour of DD and against PSD. Relatedly, it would be an argument in favour of theories with morphemes and against theories without them (see also Embick (2013)).

11.3 Contextual locality in morphophonology

The key empirical point raised in section 11.2 is whether or not stem alternations show contextual locality effects. Although the focus of this paper is on stems, it is worth noting that the question about contextual locality is much broader in scope. Stem alternations are one component of a larger domain that comprises all morphophonological alternations, which includes the behaviour of non-stem morphemes as well. The question to be posed, then, is whether contextual locality effects play a role in morphophonology more generally. Most research programmes agree that they do.

[3] Carstairs (1987), which argues that affixes and stems behave differently from such paradigmatic constraints, is potentially relevant on this point. For the types of paradigmatic constraints that are proposed for affixes—and what is predicted in theories that do not employ paradigms—see the discussion in Cameron-Faulker and Carstairs-McCarthy (2000) and Halle and Marantz (2008), as well as Carstairs-McCarthy (2010).

Although there are, among these, active debates about everything from large architectural matters to small points of detail, the majority of approaches to morphophonology agree (i) that there are significant restrictions on when triggers and targets of alternations may interact, and (ii) that it is a primary task of the theory to identify these locality conditions and examine their broader implications. This particular focus is by no means novel to recent theories. Rather, it has been a primary concern of research in this area since at least the beginnings of generative grammar.[4]

With respect to particulars, it will be assumed in the discussion to come that the following types of locality are potentially relevant for stems (cf. Embick 2010a,b; 2012):[5,6]

(L1) A type of morphological locality, stated in terms of the concatenation $\frown$ (immediate linear adjacency) of morphemes. For operations respecting this, morphemes X and Y can interact only when $X \frown Y$ or $Y \frown X$. Concatenation is implicated for suppletive contextual allomorphy (Embick 2010a and references cited there), and also for a particular type of morphophonological rule as well (Embick 2010b; 2012; 2013; Calabrese 2012).

(L2) Phonological locality, of the type that is manifested in autosegmental and other types of phonological representations. For this, a trigger X and a target Y of a rule will interact only when they are local in the relevant phonological sense. This kind of locality is hypothesized to be relevant for morphophonological rules in which either the trigger or the target is identified in terms of its phonology alone (Embick 2010b, 2012).

[4] Two early examples are as follows. First, in the discussion of 'affix hopping' and related phenomena in Chomsky (1957) employs morphophonemic rules to relate *take+past* (for the past tense of *take*) to its phonological form *took*. The affixation of the *+past* element is possible only when *take* is local to it in a way that is encoded in the affix hopping rule. It is for this reason that sentences like **John did not took the class* are ungrammatical. It can be seen that this theory holds that the derivation of stem alternants like *took* is restricted to contexts in which the target of the change is local to the trigger of the change, i.e. that, ultimately, this part of morphophonology cannot be understood without reference to locality conditions that derive from syntactic representations and operations.

Second, Halle (1959: 27) presents arguments about *directionality* effects in allomorphy, and considers whether roots determine the form of affixes, or vice versa; his general argument for the former position implicates key questions about the trade-off between memorization of alternants and derivation by rule that continue to be of central importance for grammatical theory.

[5] Both (L1) and (L2) are implicated in morphologically conditioned phonological rules. Terminologically speaking, the term *Readjustment Rules* is used for these in, e.g., Halle and Marantz (1993) and Embick and Halle (2005). For reasons that are discussed in Embick (2010b, 2012, 2013), this term might cover what is actually distinct types of rules, with different locality properties (one type obeys (L1), the other (L2)). To avoid complications related to this I will speak of *morphophonological* rules to refer to this part of the theory.

[6] Embick (2010a) also proposes a third type of locality condition on morphophonology that is defined in terms of phases (cf. Chomsky 2000, 2001; Marantz 2007, 2013a; Embick and Marantz 2008). This component of the theory is not directly relevant to the particular arguments about stems that are advanced in this paper, although it is in principle relevant to many other aspects of stem morphophonology.

The condition (L1) holds that contextual allomorphy (and perhaps a certain type of morphologically conditioned phonological rule) is restricted in a way that is dependent on the concatenation of morphemes. A primary claim of this part of the theory is that a theory that dispenses with morphemes cannot make the correct predictions about when allomorphy does and does not occur.

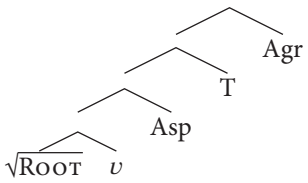
On the face of it, (L2) is primarily phonological in nature. It says that triggers and targets can interact when they are local in terms of a phonological representation. So, for instance, autosegments [F1] and [F2] on the same tier could interact only when there is no intervening feature there; or the target of some operation could be defined in terms of metrical structure; and so on. Because they are defined in terms of phonological representations, operations of this type can potentially skip morphemes. However, although (L2) locality effects are defined in terms of phonological representations, this is only one facet of their behaviour. As will be discussed in section 11.4.3, morphophonological alternations covered by (L2) appear to behave in a way that is defined by the position of a particular morpheme in a complex structure. Thus, the morpheme is crucial to (L2), just as it is to (L1).

One issue of particular importance for stems concerns how morphophonological rules are triggered by synsem features. A working hypothesis in Distributed Morphology is that the same morphophonological rule can be triggered by a disjoint set of features (see Embick and Halle 2005 for discussion). So, for example, the umlaut process of German, which is phonologically a fronting rule, is triggered by a set of morphemes that do not form a natural synsem class: plurals, diminutives, certain agreement features, and so on. This property will be referred to as *Disjoint Conditioning*:

(6) *Disjoint Conditioning:* A single (morpho)phonological rule *R* can be triggered by a set of features $\{F_1 \ldots F_n\}$ that do not form a natural class.

Disjoint Condition is, along with underspecification of Vocabulary Items, another way in which form/meaning connections are not always direct. At the same time, it is crucial to emphasize that a theory that allows Disjoint Conditioning nevertheless makes strong predictions about the locality of triggers and targets, as long as it is centred on morphemes, and locality conditions like (L1) and (L2). For purposes of illustration, consider the verbal structure employed in section 11.2:

(7) verb



A theory with Disjoint Conditioning says that a single phonological rule could be triggered by a set containing Asp[+perf]; T[+past], Agr[+1,-2,+pl]; and so on. However, since it operates in terms of morphemes, a natural hypothesis is that collections of features from different morphemes *may not* serve as triggers:

(8) *Morpheme Trigger Hypothesis:* A morphophonological rule may be triggered only by features on a single morpheme.

So, for instance, a rule could not be triggered by [+perf,+1,+pl], since these are features of different morphemes (Aspect and Agr).[7] The predictions of the Morpheme Trigger Hypothesis are, of course, not statable in a theory that dispenses with morphemes in favor of feature matrices, a point that will be visited at several points in section 11.4.[8]

In summary, the locality conditions (L1) and (L2) in combination with Morpheme Trigger Hypothesis produce a substantive theory of contextual locality for stem alternations, even if Disjoint Conditioning introduces some additional indirectness into this part of the theory by allowing the same phonological rule to be called by features that do not have to form a natural class.

Before moving on to the case studies, it bears repeating that the theory based on (L1) and (L2) is one among many DD theories. The literature contains a number of different proposals concerning the type of information that is available for contextual allomorphy, the structural conditions under which suppletion occurs, the nature of morphophonological locality, and so on (for a small sample of views relevant to (L1) and (L2) in particular, see Carstairs (1987), Carstairs-McCarthy (1992, 2001), Bobaljik (2000, 2012), Lieber, (1987, 1992), Kiparsky (1982, 1996, 2007), Bermúdez-Otero (forthcoming), and Wolf (2006, 2008)). These theories make a number of distinct

[7] Of course, nothing precludes a situation in which, e.g., Asp[+perf] triggers one rule, and Agr[+1,+pl] triggers another, so that they have a 'net' effect on the target. In this scenario, however, each of the two rules should be individually identifiable.

[8] The predictions of the Morpheme Trigger Hypothesis are affected in an interesting way if the Fusion of adjacent nodes is allowed in the theory (for Fusion, see, e.g., Halle and Marantz 1993). Fusion takes two morphemes [+α] and [+β] and creates from them a single morpheme [+α,+β] which is the target of Vocabulary Insertion. This operation is motivated when expected independently occurring exponents of [+α] and [+β] (e.g. /-x/ and /-y/) fail to occur, with a 'portmanteau' /-z/ occurring instead; this kind of situation can be analysed with Fusion, and the Vocabulary Items in (i):

(i) Vocabulary Items

+α,+β	↔	-z
+α	↔	-x
+β	↔	-y

With Fusion, and with reference to (7) for illustration, a bundle of features from two morphemes could trigger a morphophonological rule if those two morphemes are fused (for example, if v and Aspect, or Tense and Agr fused in particular feature combinations). Thus, the prediction of the Morpheme Trigger Hypothesis is that this kind of 'bundled' triggering is possible only when there is clear evidence (e.g. in the form of a portmanteau realization) of Fusion for the morphemes with the trigger features.

(and often conflicting) predictions about when triggers and targets of allomorphy may interact. But whatever differences they have in the details, all of these approaches agree on the central point, which is that morphophonological alternations are subject to locality relations of the type that derive from structured complexes of morphemes.

11.4 Illustrations

This section illustrates suppletive and morphophonological alternations, with reference to the general idea that stem distributions are determined in ways that respect contextual locality conditions. Subsections 11.4.1–11.4.4 offer particular points concerning DD versus PSD on this theme; the general conclusions to be drawn from the case studies are presented in 11.4.5.

11.4.1 Suppletion (contextual allomorphy) in latin

Suppletive contextual allomorphy is found when a single morpheme is realized by more than a single Vocabulary Item; the particular form that is realized is determined by an element in the context of the morpheme undergoing insertion. The alternation between *go* and *went* and English can be treated in this way. On the assumption that these forms are the realizations of a light verb that I will refer to as v_{go}, the Vocabulary Items in (9) account for the stem alternation:[9]

(9) $v_{go} \leftrightarrow$ went/_ ⁀T[+past]
$v_{go} \leftrightarrow$ go

The first of these two items takes precedence over the second when v_{go} (the target) is concatenated with the past tense morpheme T[+past] (the trigger). This is a simple illustration of how the Vocabulary Insertion mechanism can be used to implement stem suppletion. For further illustrations of the role of contextual locality in suppletive allomorphy, more complex examples are required. A suitable one for present purposes is provided by *esse* 'be' in Latin. The indicative forms are shown in (10); the last row in the table is an abstract segmentation (B= 'be'; Asp(ect); T(ense); T/M= Tense/Mood; Agr(eement)):[10]

[9] Minor variants on (9) achieve the same effect; for example, the first Vocabulary Item could be treated with the exponent *wen*, in which case the T[+past] morpheme would be realized by *-t*.

[10] For reasons that are discussed later in this section when the subjunctive forms are introduced, I have glossed the Tense morpheme as 'T' in (10), and as 'T/M' in (11).

Several phonological processes are at play in (10) (and in (11) below). For example, the Asp[+perf] morpheme, which is taken to be *-i*, fails to surface in some forms of the perfect indicative, and is changed to /e/ when it precedes an /r/ (in, e.g., the pluperfect and future perfect). These and other aspects of the morphophonology do not affect the main points in the text.

(10) Indicative: *esse*

	Present	**Imperfect**	**Fut.**	**Perf.**	**Pluperfect**	**Fut. Perfect**
1s	su-m	er-a-m	er-ō	fu-ī	fu-e-ra-m	fu-e-r-ō
2s	es	er-ā-s	er-i-s	fu-istī	fu-e-rā-s	fu-e-ri-s
3s	es-t	er-a-t	er-i-t	fu-i-t	fu-e-ra-t	fu-e-ri-t
1p	su-mus	er-ā-mus	er-i-mus	fu-i-mus	fu-e-rā-mus	fu-e-ri-mus
2p	es-tis	er-ā-tis	er-i-tis	fu-istis	fu-e-rā-tis	fu-e-ri-tis
3p	su-nt	er-a-nt	er-unt	fu-ērunt	fu-e-ra-nt	fu-e-ri-nt
	B-Agr	B-T-Agr	B-T-Agr	B-Asp-Agr	B-Asp-T-Agr	B-Asp-T-Agr

In these forms, there are effectively three allomorphs of *esse* (more precisely, v_{be}) in evidence: *su-*, *es-*, and *fu-*. With regard to the *es-* form I am assuming that Latin has a rule of *Rhotacism* that creates *er-* when the /s/ of *es-* is intervocalic.

The subjunctive forms in (11) show a *sī-* allomorph of v_{be}, in addition to *es-* and *fu-*:

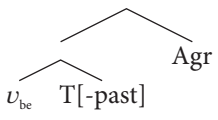
(11) Subjunctive

	pres	**impf**	**perf**	**plpf**
1s	si-m	es-se-m	fu-e-ri-m	fu-i-s-se-m
2s	sī-s	es-sē-s	fu-e-rĭ-s	fu-i-s-sē-s
3s	si-t	es-se-t	fu-e-ri-t	fu-i-s-se-t
1p	sī-mus	es-sē-mus	fu-e-rĭ-mus	fu-i-s-sē-mus
2p	sī-tis	es-sē-tis	fu-e-rĭ-tis	fu-i-s-sē-tis
3p	si-nt	es-se-nt	fu-e-ri-nt	fu-i-s-se-nt
	B-T/M-Agr	B-T/M-Agr	B-Asp-T/M-Agr	B-Asp-T-T/M-Agr

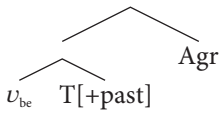
Taken together, the facts in (10)–(11) look complex. In a sense, this is not surprising; after all, suppletion is a kind of worst-case scenario as far as systematic form/meaning connections go, and some aspects of (10)–(11) are irreducible (i.e. they will simply have to be memorized on any approach). At the same time, though, there are important aspects of the distribution of *esse*'s allomorphs that illustrate the workings of the contextual locality effects that were introduced in section 11.3.

A working analysis of (10)–(11) can be constructed along the following lines. A fundamental distinction in the Latin verb (in analyses from Varro onwards) is that between imperfect and perfect tenses. Structurally speaking, the perfects contain an aspectual morpheme Asp[+perf] in addition to Tense, whereas imperfect tenses contain no such morpheme (see Embick 2000). So, for example, the present indicative and imperfect indicative are distinguished by [±past], as shown in (12)–(13); the perfect, which contains Asp[+perf], is shown in (14), along with a pluperfect (15):

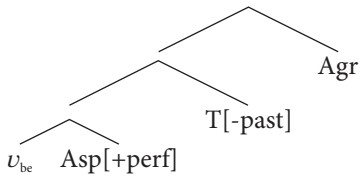
(12) present



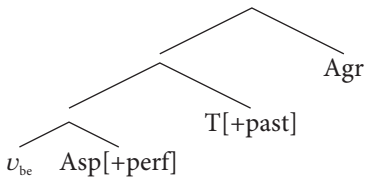
(13) imperfect



(14) perfect

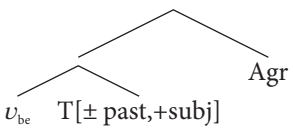


(15) pluperfect



For subjunctives, I will assume that the feature [+subj] is present on tense in addition to whatever other features tense has. So, for example, present ([-past]) and imperfect ([+past]) subjunctives have the structure shown in (16):[11]

(16) present and imperfect subjunctives



Looking now at the forms in (10)–(11) in the light of the structures (12)–(16), a number of generalizations emerge. In particular, (i) the *fu-* allomorph occurs next to Asp[+perf]; (ii) the *sī-* allomorph appears next to T[-past,+subj]; and (iii) the *es-* allomorph appears elsewhere, except in (iv) the 1s, 1p, and 3p present

[11] In effect, this analysis says that there is a single 'Tense/Mood' morpheme in Latin verbs. The only place where two such morphemes are required is in the pluperfect subjunctive, if the segmentation in (11) is maintained. The presence of two Tense/Mood heads in that type of verb fits well with its past irrealis semantics (cf. Iatridou 2000).

indicative, where *su-* is found. These generalizations are accounted for by the Vocabulary Items in (17):

(17) Spell-out of *esse*

v_{be}	↔	su	/__⌒	[+1],[-1,-2,+pl]
v_{be}	↔	sī	/__⌒	T[-past,+subj]
v_{be}	↔	fu	/__⌒	Asp[+perf]
v_{be}	↔	es		

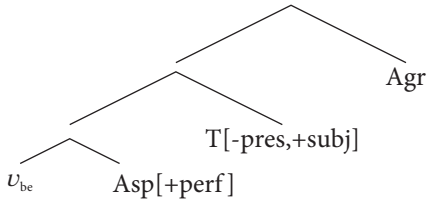
One issue worth further comment in (17) concerns the Vocabulary Item that inserts the *su-* exponent. For this Vocabulary Item to work properly, v_{be} must be concatenated with the Agr node with person/number features. As shown in (12), the present tense verb contains a T[-past] morpheme. This morpheme must not intervene between v_{be} and Agr, if this analysis is to work. As discussed in earlier work on linear locality (Embick 2003, 2010a), it appears that at least some morphemes that have no overt realization are transparent for allomorphic purposes. One way of implementing this idea is by deleting such morphemes. In the case of Latin T[-past], I will assume that a general rule (of 'radical' Impoverishment; e.g. Arregi and Nevins (2007)) deletes this node prior to Vocabulary Insertion as a whole. Thus, when Vocabulary Insertion begins at v_{be}, it is concatenated with Agr.[12] Part of the idea behind this deletion is that while T[-past] is required for syntactic reasons, it has no obvious effects in the morphology of Latin, so that it may be deleted at PF in the way described. The deletion rule is, however, restricted to T[-past]; it does not delete present subjunctive T[-pres,+subj], which is referred to by the second Vocabulary Item in (17).

Two aspects of the analysis developed above speak directly to the theory of stems. The first is that while the forms of v_{be} appear to be quite irregular, their distribution follows a pattern in which the features conditioning suppletive allomorphy are always local to the target, in the way that is expected in the particular DD theory assumed here. The second point is that this kind of local contextual conditioning could only be accidental in PSD theories that associate stem forms with arrays of features.

These points can be seen clearly when we consider that the features in the three contextual conditions in (17) may co-occur with one another. For example, in perfect subjunctives, Asp[+perf] co-occurs with T[-past,+subj]:

[12] Another issue with the *su-* Vocabulary Item is that it has a disjunctive contextual condition: [+1] and [-1,-2,+pl]. Depending on other assumptions, this might be rejected in favour of an analysis without disjunction, for which various alternatives are possible. However, since this is not a question that directly implicates the locality of the trigger and target, I will put it to one side.

(18) perfect subjunctive



Moreover, all of the verb forms in question cooccur with the person number features [+1] and [-1,-2,+pl] that are referred to in the first Vocabulary Item in (17).

The co-occurrence of these contextual features is important because the form of v_{be} must be affected by each independently, as in (17). Thus, when triggering features co-occur, it is in principle possible for any of *su-*, *sī*, or *fu-* to be inserted. Crucially, the allomorphs that are in fact found are those that are conditioned *locally*, by the linearly adjacent head. In (18), for example, v_{be} is concatenated with Asp[+perf]; and it is not concatenated with T[-past,+subj]. For a theory with (L1), it is unsurprising that *fu-* is inserted in this context. Similarly, in a present subjunctive with the structure in (16), v_{be} is concatenated with T[-past,+subj], not with the Agr morpheme. According to the analysis developed above, this is why *sī* is inserted, and not *su-*.

While the attested pattern of allomorphs under 'feature competition' is what is expected in a DD theory, the same is not true for a PSD theory that associates stems with feature matrices. Consider the case of a first person plural perfect subjunctive. In terms of a morphemeless representation, the paradigmatic cell defined by this combination of features can be represented as in (19):

(19) feature matrix for 1pl perfect subjunctive of *be*

$$\begin{bmatrix} \text{BE} \\ +1 \\ -2 \\ +\text{pl} \\ +\text{perf} \\ -\text{past} \\ +\text{subj} \end{bmatrix}$$

Each of the stem allomorphs *su-*, *sī-*, and *fu-* could in principle be associated with this cell, since each of these alternants is associated with features that appear in (19). A paradigmatic theory cannot appeal to contextual locality to resolve this competition, since, in a representation without morphemes like (19), all of the features in the matrix are equally local to *be*. Thus, the fact, e.g., that *fu-* wins out over *sī-* and *su-* must simply be stipulated; and so must the outcome of every other such competition.

11.4.2 Two alternations in italian

Two alternations in the verbal system of Italian provide a further case study for the claim that stem distributions are conditioned by local context. This particular case study is significant for two reasons. First, because it has been described as a case in which 'global' phonological properties determine a suppletive stem alternation (Carstairs 1988, 1990); this is something that is predicted not to happen in localist theories of allomorph selection. And second, because it has been proposed that the distribution of allomorphs reflects paradigmatic structure; e.g. Maiden's (2005) treatment of these facts argues against a phonological approach and for paradigms in the morphomic tradition. In line with the general goals of this section, I will show how the pattern can be analysed with local relations between the triggers and targets of allomorphy (cf. Embick 2010b).

Carstairs (1988,1990) follows traditional discussions of Italian in describing suppletion of *andare* 'go' as being conditioned by surface stress. The facts show that the stem is *va(d)-* when under stress, and *and-* otherwise. The same distributional pattern is found with *-isc* alternating with *-i* in many *-ire* (conjugation III) verbs like *finire*; that is, *-isc* appears exactly where *va(d)* does:

(20) Present forms of *finire* and *andare*

	finire		*andare*	
p/n	**pr. ind.**	**pr. subj.**	**pr. ind.**	**pr. subj.**
1s	fin-**ísc**-o	fin-**ísc**-a	**vádo**	**váda**
2s	fin-**ísc**-i	fin-**ísc**-a	**vái**	**váda**
3s	fin-**ísc**-e	fin-**ísc**-a	**vá**	**váda**
1p	fin-i-ámo	fin-i-ámo	andiámo	andiámo
2p	fin-í-te	fin-i-áte	andáte	andiáte
3p	fin-**ísc**-ono	fin-**ísc**-ano	**vánno**	**vádano**

The relevance of these facts for the theory of allomorphy—particularly because of the putative global phonological conditioning—has been addressed in Kiparsky (1996), Paster (2006), and many other places. Many prior analyses take the conditioning by stress at face value. However, it is important to note that although stress *correlates* with allomorphy in these examples, there is no conclusive evidence that it actually *causes* the allomorphic alternation (see Embick (2010b)). For example, there is no way of moving the stress, such that a single person/number combination 'vacillates' between *and* and *va(d)*. Thus, there is no direct evidence that global phonology must be referred to in the synchronic analysis of this phenomenon (this conclusion is also reached in Maiden 2005).

With this in mind, I will now develop an analysis in which the alternations in (20) are triggered locally. As a first step, a more complete look at the verbs in question is useful. Starting with *andare* (i.e. realizations of v_{go}), while (20) might make it look as though *and* appears when there are person features [+1] or [+2] and [+pl], the

distribution of *and* is not restricted in this way. Rather, the *and* stem also appears in a number of non-finite forms, as well as in a number of other finite tenses:

(21) Forms of *andare*

p/n	pr. ind.	pr. subj.	impf.	pret.	fut.
1s	vádo	váda	andavo	andai	andrò
2s	vái	váda	andavi	andasti	andrai
3s	vá	váda	andava	andò	andrà
1p	andiámo	andiámo	andavamo	andammo	andremo
2p	andáte	andiáte	andavate	andaste	andrete
3p	vánno	vádano	andavano	andarono	andranno

It can be seen from (21) that *and* is clearly not the allomorph with the restricted distribution; it is the default form of v_{go}. The *va(d)* allomorph is special: the only place it appears is in the present tense, whether indicative or subjunctive (and in imperatives, which I assume to be present tense as well). The same distribution is found with *-isc*; it appears only in the present indicative and subjunctive (and in the imperative), while *-i* is found elsewhere:

(22) Forms of *finire*

p/n	pr. ind.	pr. subj.	impf.	pret.	fut.
1s	fin-ísc-o	fin-ísc-a	finivo	finii	finirò
2s	fin-ísc-i	fin-ísc-a	finivi	finisti	finirai
3s	fin-ísc-e	fin-ísc-a	finiva	finì	finirà
1p	fin-i-ámo	fin-i-ámo	finivamo	finimmo	finiremo
2p	fin-í-te	fin-i-áte	finivate	finiste	finirete
3p	fin-ísc-ono	fin-ísc-ano	finivano	finirono	finiranno

Specifying the Vocabulary Items that insert the special items contextually for present tense (here T[-past]) accounts for most of the facts in (21) and (22). In (23), *-isc* is treated as the realization of a conjugation class feature [III] which I have put on *v* in (23b); it could just as well be treated as a Theme node:[13]

(23) Two pairs of Vocabulary Items

a. $v_{go} \leftrightarrow$ va(d)/___⁀T[-past]
   $v_{go} \leftrightarrow$ and

b. *v*[III] ↔ -isc/___⁀T[-past]
   v[III] ↔ -i

[13] Within the Italian third conjugation, some verbs show the *-i*/*-isc* alternation, while others show only *-i*. This complication is not addressed in the main text, but it could be handled by, e.g., adding a further contextual condition listing the verbs in question to the left of *v*[III] in (23):

(i) *v*[III] ↔ -isc-/LIST ___⁀T[pres]

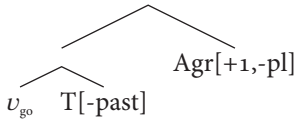
Other options are possible, but I will not dwell on this point here.

As it stands, the analysis produces the special allomorphs in all person/number combinations of the present indicative and present subjunctive. This is only partially correct. What is needed in addition is a way of making the defaults *and* and *-i* appear in the first and second person plurals. This can be done with Impoverishment. The rule in (24) produces the correct results by deleting the feature [-past] in the context of Agreement nodes that are [+part] (i.e. participants [+1] and [+2]) and [+pl]:[14,15]

(24) *Impoverishment:* [-past] →Ø/___ [+part,+pl]

To see how the different components of the analysis based on (23) and (24) fit together, consider the derivation of 1s *vado*, for which the structure is (25):

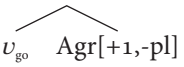
(25) Structure for 1sg *vado*



Vocabulary Insertion inserts *vad-* for υ_{go}, since it is adjacent to T[-past]; the T[-past] node is realized as *-Ø*, and 1s Agr as *-o*.

In the case of e.g. 1pl *andiamo*, the output of the syntax is identical to (25), except with [+1,+pl] instead of [+1,-pl] features on Agr. The Impoverishment rule (24) applies prior to Vocabulary Insertion, and eliminates [-past]:

(26) Structure for 1p *andiamo*

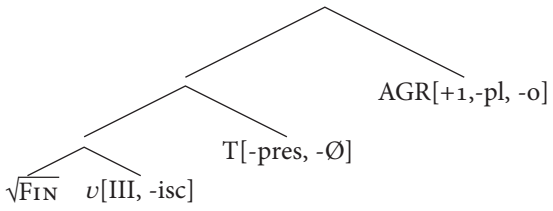


Thus, Vocabulary Insertion at υ_{go} cannot insert *vad*, because the contextual condition on its application is not met. Rather, the default *and* is inserted. The distribution of *-isc* relative to *-i* for the verbs of conjugation III is accounted for in exactly the same way, except that in this case it is a *v* with a conjugation class feature (or a Theme node with such features) that is realized as *-isc* or *-i*:

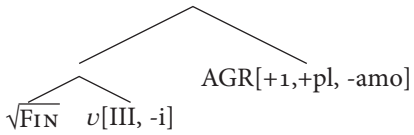
[14] As discussed in detail in Calabrese (2012), the rule (24) might be subsumed under a broader set of generalizations about this combination of features in Italian. Calabrese's discussion contains a number of important points about the relationship between markedness and Impoverishment that are relevant to both the facts in the text and to certain other types of stem distributions in Italian. These connect with the possibility there might be more general motivations for rules like (24). But for the immediate focus of this section—the locality of trigger target relations—(24) suffices.

[15] Note that (24) deletes [-past] on Tense. It does not delete the mood feature [+subj] that is found in present subjunctives. Thus, while present subjunctives will be the same as indicatives with respect to the Vocabulary Items in (23), [+subj] can have an effect on outer morphemes in the subjunctives.

(27) structure for 1s *fin-isc-o*



(28) structure for 1p *fin-i-amo* (T[-pres] deleted by (24))



In the other (non-present) tenses of (21) and (22), the *va(d)-* and *-isc-* alternants are not found. In the analysis developed here, the reason for this is that the structures for these verb forms do not have the [-past] Tense feature. Or, because the [-past] feature is not local to the target; for example, the Future is presumably [-past], but contains a modal head (realized as *-r*) that interrupts the locality relationship between v_{go}/v[III] and the triggering Tense head.

In summary, this analysis accounts for the distribution of stems in a way that does not require reference to global phonology (problematic for reasons discussed in Embick 2010a), or by moving to a PSD theory (problematic for the reasons discussed in this paper).

11.4.3 Morphemes and phonological locality

The alternations examined in sections 11.4.1 and 11.4.2 involve suppletive contextual allomorphy. This type of allomorphy, which is analyzed with the Vocabulary Insertion operation, is expected to obey morphological locality constraints; in the approach adopted here, this means concatenation (L1).

As emphasized in section 11.3 with reference to (L2), suppletive allomorphy comprises only part of the theory of alternations. Languages also employ stem alternations in which the stem forms appear to be related phonologically, as in Spanish diphthongization (e.g. *pienso*/*pensar* for the verb 'to think'), and so on. The DD Theory outlined in section 11.3 holds that many 'morphophonological' alternations of this type are constrained to apply under *phonologically* defined locality conditions. There are, in fact, two components to this part of the theory. The first is, in the terminology of Embick (2013) that such alternations have a *morphological locus*: a morphophonological rule acts in a way that is local to the morpheme that triggers it. The second part is phonological: from that locus, the rule

will affect elements that are local in terms of phonological representations. With respect to the general goals of this section, morphophonological interactions falling under (L2) illustrate a second type of contextual locality that is crucial for the theory of stems.

The phenomenon of *metaphony* that is found in many varieties of Italo-Romance can be used to illustrate the interaction between morphological loci and phonological representations. Phonologically speaking, metaphony raises (or diphthongizes) a stressed vowel, in a way that was originally triggered by a high vowel in the following syllable (e.g. Calabrese 1985, 1999, 2009; Maiden 1991). The process is morphologized in many dialects, where the triggering elements have been reduced or eliminated. In such dialects, the alternation, though a phonological change, is morphologically triggered. For instance, Maiden (1991:159) gives the examples of metaphony from the dialect of Ischia that are shown in (29), where ('standard') Italian is given for comparison:

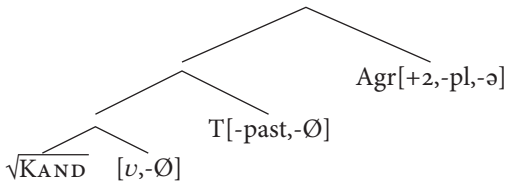
(29) Metaphony triggered by 2s AGR (Maiden 1991:159); *cant/kand* 'sing'

	Italian		*Ischia, Campania*	
	present	*imperfect*	*present*	*imperfect*
1s	canto	cantavo	kandə	kandavə
2s	canti	cantavi	**kɛ̲ndə**	**kandɛ̲və**
3s	canta	cantava	kandə	kandavə

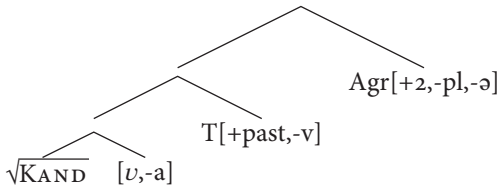
Metaphony affects the underlined vowels in the boldfaced words, raising them to /ɛ/. As mentioned above, the process originated phonologically, triggered by the 2s Agr ending, *-i* (cf. Italian). In the Ischia variety, the 2s Agr morpheme is always realized as *-ə*. It is for this reason that metaphony is morphologically triggered in the synchronic language.

A key observation to be made about (29) is that metaphony affects the vowel that is phonologically local to the 2s Agr morpheme, i.e. it is the phonologically adjacent vowel to the left of this trigger that is affected. The structures underlying the present and imperfect forms are those in (30) and (31) (for convenience the theme *-a* in the latter is shown as a realization of *v*):

(30) 2s present



(31) 2s imperfect



In the present tense, it is the vowel of the Root that is changed, producing a stem change. In the imperfect, though, the phonologically adjacent element is the theme vowel *-a* that is affected. In a theory in which morphophonological rules are triggered by morphemes in the way outlined above, this type of morphophonological locality effect is what is expected.

There are two points to be drawn from this discussion. The first is that metaphony obeys phonological locality (adjacency in a phonological representation), not morphological locality (adjacency of morphemes). This is clear from the fact that, in the imperfect, the morphophonological change affects the theme, even though there is an intervening Tense morpheme that is realized as *-v*. It is because of this kind of effect that the theory of stems must involve both morphological and phonological components, as embodied in (L1) and (L2) of section 11.3.

The second point, which is directly relevant for comparing DD and PSD theories, is that this kind of locality effect is predicted only in a theory that makes use of morphemes, since it is the position of the Agr morpheme that defines the locus of metaphony. In terms of the representations (30) and (31), the fact that metaphony affects the Root and the theme respectively follows straightforwardly. On the other hand, in a theory employing feature matrices for paradigm cells, this type of locality effect is not expected, for the reasons that have been discussed earlier in this section; see also 11.4.5 below.

11.4.4 'Raising' verbs in spanish

One important aspect of the theory of morphological loci is that phonologically defined interactions under (L2) may potentially produce distributions of stems that are difficult or impossible to state in terms of local synsem features. Given the obvious role of historical phonology in the study of stem alternations, this point is directly relevant to understanding why some stem alternations might appear to be 'unnatural' from a paradigmatic (= synsem feature) point of view. If the localist, DD theory defended in this paper is correct, even distributions that are not natural in terms of synsem feature classes must be conditioned locally in one of the ways defined by (L1) and (L2).

A class of Spanish verbs from conjugation III verbs provides a good illustration of the tension between synsem and phonological triggering of stem alternations (this section elaborates on Embick 2010b, 2012). The verb *pedir* 'to ask' is a standard representative of this class; it shows an alternation between a mid- and a high-stem vowel, as shown in (32):

(32) Forms of *pedir*

	1s	2s	3s	1p	2p	3p
pr. ind.	**pido**	**pides**	**pide**	pedimos	pedís	**piden**
pr. subj.	**pida**	**pidas**	**pida**	**pidamos**	**pidáis**	**pidan**
pret.	pedí	pediste	**pidió**	pedimos	pedisteis	**pidieron**
impf.	pedía	pedías	pedía	pedíamos	pedíais	pedían
impf. subj.	**pidiera**	**pidieras**	**pidiera**	**pidiéramos**	**pidierais**	**pidieran**
fut.	pediré	pedirás	pedirá	pediremos	pediréis	pedirán
cond.	pediría	pedirías	pediría	pediríamos	pediríais	pedirían

For historical reasons (the mid-vowel is original), these verbs are referred to as 'Raising' verbs. Spanish has a number of verbs that belong to this class, and all show the same distribution of stem forms as *pedir*. The alternation is a typical 'morphophonological' one, in the sense that it applies to some Roots and not others; that is, some verbs from conjugation III look as though they could alternate in this way (e.g. *sumergir* 'to submerge'; *escribir* 'to write'), but do not.

In terms of synsem features, the stem alternants are distributed as follows:

(33) a. *ped*: 1 and 2 pl present indicatives; non-third person preterites; all imperfects, futures, and conditionals.
b. *pid*: 1, 2, 3 sg, and 3 pl present indicatives; all the present subjunctives; all the imperfect subjunctives; 3s and 3p preterites.

If the trigger for deriving one alternant from the other had to be stated in terms of such feature combinations, it is unclear that much would be left of a restrictive theory of stem distributions. As discussed in section 11.3, there is no problem per se with having a single morphophonological rule triggered by a set of distinct morphemes (Disjoint Conditioning). The stem forms of *pedir*, though, go beyond this, because the changes are associated with 'bundles' of features; for example, 3rd plural does not take the *pid* stem across the board, only in the present tenses, the preterite, and the imperfect subjunctive; and so on. Allowing the theory to make reference to (arbitrary) subsets of bundled features is tantamount to removing all contextual constrains on stem distribution, i.e. accepting that (at least for this kind of allomorphy), the affixless/paradigmatic view is correct.

While stating the distribution of alternants in synsem terms is problematic, there is a straightforward morphophonological analysis of (32). Building on Harris (1969), the alternation can be treated as phonological dissimilation, with an underlying /i/

lowered to /e/ when the nucleus of the following syllable is a high vowel; (34) schematizes this:[16]

(34) Dissimilation: i →e/___(C)i
<for the specified class of Roots>

As noted above, the restriction to a particular class of Roots reflects the fact that (34) is a morphophonological rule; not all potential undergoers are subject to it: for example, $\sqrt{\text{PID}}$ 'ask' is, but $\sqrt{\text{ESCRIB}}$ 'write' is not. It is important to note with reference to the forms in (32) that the orthographic *i* that appears after *pid* in 3s and 3p preterites *pidió* and *pidieron*, and imperfect subjunctive forms like *pidiera*, etc. is a glide, not a nucleus; for this reason, (34) does not lower stem vowels in these contexts. Overall, (34) accounts for the distribution of alternants in a way that employs a local phonological trigger, as is expected in a DD theory.

For PSD, the raising verbs highlight a number of questions. An analysis stated in synsem terms (i.e. in terms of features in paradigmatic cells) is certainly possible for PSD; it would take the form of (33). But such an analysis could be maintained only at the cost of any substantive contextual theory of stem distributions. The only motivation for such a view would be conclusive proof that patterns of stem alternants are distributed in a way that resists analysis in terms of local morphological or phonological triggers; and this is precisely not the case with the verbs considered here, since the point of this section is that there is a straightforward morphophonological analysis of (32).

It is at least conceivable that a PSD theory could analyse the raising verbs in the phonology, and not with the system for deriving and distributing stem alternants. This move would have architectural consequences. Taken at face value, it would amount to the claim that there is one morphophonology that is computed offline (for creating stem alternants to be paradigmatically distributed), and another morphophonology that is computed online (on the phonological representations associated with such cells). So, for example, such a theory might predict that stems with the same phonological change could either be distributed non-locally with respect to a trigger or locally to it, depending on whether the alternants were computed offline or online. In any case, it is difficult to be more concrete about this because the role of morphophonology in PSD theories is unclear.[17]

To summarize, the distribution of stem alternants in the Spanish raising verbs illustrates the importance of connecting stem distributions with a broader theory of local morphophonological interactions. While the distribution of alternants would

[16] A precise version of this rule would be extended to include a few verbs in which /o/ and /u/ alternate in the same way. This complication has no bearing on the point in the text, so I will proceed with the simplified (34).

[17] From another perspective, Anderson (1992: 45–6) touches on the prediction that is discussed in the text, in a way that highlights some of the complications that morphemeless theories have with morphophonological alternations.

require non-local interactions if stated in terms of synsem features, it is straightforward and local in nature if treated morphophonologically.

11.4.5 Interim summary

The main line of argument in this section is that stem distributions are subject to contextual locality conditions that follow from a DD theory, in which stem alternants are distributed in a way that reflects local derivational context. A second point is that these conditions do not follow in a PSD theory, because of its separation of derivation from distribution, and its use of features matrices/paradigmatic cells. The case studies reviewed above look at suppletive contextual allomorphy (11.4.1 and 11.4.2), and at morphophonological alternations (11.4.3 and 11.4.4), two ways in which stems can relate to each other. In each case, the triggers of the alternation are shown to be contextually local to the targets of the change. In the case of suppletive contextual allomorphy, the important locality relation is hypothesized to be the concatenation of morphemes, by (L1) of section 11.3. In the case of morphophonology, the critical representations are phonological in nature, by (L2). As stressed in 11.4.3, though, morphophonological alternations appear to have a morphological locus, in the sense that they behave as if the rule applies from a particular position in the word (i.e. a morpheme). If it is true that all such alternations operate in terms of the locus that is predicted by a morpheme-based theory, then a theory that operates on unstructured feature matrices is clearly missing an essential generalization about form/meaning connections.

It is important to stress that this comparison of DD and PSD does not make the claim that the facts considered above are impossible to state in a PSD theory. That is certainly not the case; unless there is some substantive paradigmatic theory that restricts possible stem distributions (see below), a PSD theory is capable of stating any pattern of stem distributions, including the ones found with Latin *esse*, Italian *andare* and *-i/-isc*, and so on. That is, PSD could certainly encode the patterns that DD predicts in terms of locality. But this would be missing the point: if there is no general basis for appealing to one of the methods for associating stems with cells versus another, the theory is merely restricting itself artificially to express what is predicted by a DD theory. An argument for PSD would have to show that a paradigmatic theory is correct about stem distributions in a way that is *different* from what DD predicts or could predict; and advocates of PSD have offered nothing along these lines. For example, Aronoff (2012: 32) waves at implementing stem distributions with default inheritance hierarchies, but without reference to possible restrictions on distributions. More explicit attempts to formalize stem relations cited by Aronoff seem to do little more than stipulate dependencies found in the data. For example, Bonami and Boyé (2002) appear to recognize that there should be at least some limits on possible patterns of stem distribution. Their analysis, though, states dependency relations on stems in way that is determined post hoc from what occurs

in the language being analysed. Thus, as far as the structure of the approach goes, no dependencies between cells are ruled out in principle; only in practice, by stipulation.[18] It should be clear that specifying stem dependencies on an ad hoc basis is very different from positing universal conditions on trigger/target locality that determine when alternations could in principle occur.

Moving past stipulating dependencies, another conceivable line of response—true to the 'mophology by itself' dictum at the centre of many PSD theories—would be for advocates of PSD to try to demonstrate that stem distributions do not respect any contextual locality conditions. This would mean that the kinds of examples considered earlier in this section show restricted trigger/target relations for reasons that are completely accidental, as far as the sychronic grammar goes. This a possible finding; one that would be an unreservedly negative result, as far as this part of grammar is concerned. It would amount to the discovery that—in spite of the fact that questions about the representation of primitive units and their composition into complex objects are central in phonology, syntax, semantics, derivational morphology, etc.—stem distributions are just different. If this were true, it is not clear what theoretical interest stems would hold, since the data structures involved are finite (given the finite number of stem alternations in any language), and therefore formally trivial. It is perhaps for this reason that, to my mind, current synchronic theories of stems like those of Stump (2001), Bonami and Boyé (2002), and Aronoff (2012), do not hold more theoretical interest than an enumeration of the facts. That is, if they are correct, then there is really very little to say about this part of the theory of grammar.[19]

In any event, there is little reason to dwell further on possible PSD responses, or on what general theoretical interest PSD might have. For the reasons that I have articulated above, I see little evidence at this point that the DD view should be abandoned in favour of the PSD alternative; and the large amount of productive work devoted to investigating questions (Q1) and (Q2) in ways that connect with the general theory of morphological and phonological alternations (recall the references at the end of section 11.3) suggest that this is not an isolated conclusion.

11.5 *Separation* and the issue with morphomic distributions

As mentioned in the sections 11.1 and 11.2, Aronoff's (1994, 2012) approach to stems involves two components: it says (i) that languages have morphomic stems, and (ii) that these call for a PSD theory of stem distributions. The main focus of section 11.4 is part (ii). This section offers some comments on (i), the putative

[18] The fact that any relation between cells is in principle possible is perhaps unsurprising, given the connections that they see between their approach and the Rules of Referral employed in Zwicky (1985) and Stump (1993), which allow connections between paradigmatic cells to be stated in an unrestricted way.

[19] On the other hand there would still be interesting diachronic questions at play, as emphasized by Maiden (1992, 2005) and others.

morphomic distributions themselves. More precisely, I will look schematically at a kind of distribution that appears to present challenges even for a theory that admits both underspecification of Vocabulary Items and Disjoint Conditioning of morpho-phonological rules (recall section 11.3). With this specific point in mind, the argument of 11.5.1 and 11.5.2 is that *if* there are real morphomic distributions, it need not be concluded that stem derivation is separate from stem distribution.

Throughout the discussion, I will put to the side the question of whether the (by definition) morphomic distributions are actually found, and illustrate the main points of the argument schematically. In my view, many of the arguments for putative morphomes advanced in the literature are less than conclusive, but I will abstract away from this point here.

11.5.1 The problem, schematically

Terminologically, I will use *morphomic*, etc. in a restricted sense, for the specific kind of distribution that is illustrated in this section; the main elements involved in this definition are adapted from Aronoff's 2012 discussion.[20] Consider a language that has different Root classes, in which a particular stem form—concretely, Stem2—is related to the Root (i.e. the underlying phonological form) by a number of distinct phonological changes R_1, R_2, etc.:

(35) Relation between alternants

a. Root type 1:

$$\text{Root} \xrightarrow[R_1]{} \text{Stem} \xrightarrow{} \text{Stem2}$$

b. Root type 2:

$$\text{Root} \xrightarrow[R_2]{} \text{Stem} \xrightarrow{} \text{Stem2}$$

c. ⋮

That is, Stem2 is formed in different ways for different Root classes; it could be raising of the Root vowel for class 1, palatalization of a final consonant for class 2, and so on, just as long as there are distinct phonological rules *R* involved.

Moving past the relation between stem alternants to the distribution of stems, suppose further that the Stem2 forms are found in a non-natural class of synsem features:

[20] I emphasize the terminology because there seem to be a number of distinct but related senses of *morphome* at play in the literature (to judge from e.g. Maiden et al. (2011)).

(36) Distribution: Stem2 forms appear in feature combinations $\{F_1 \ldots F_n\}$, where there is no unifying feature common to these (i.e. the distribution is not a natural class).

In this type of scenario, a theory with morphophonological rules and Disjoint Conditioning does not have a problem with listing the environments in (36) per se. However, on the face of it there is a problem: it would have to repeat the conditioning environment in a number of distinct rules. Specifically, each of the rules *R* would have to be specified to apply in $\{F_1 \ldots F_n\}$, as shown in (37):

(37) Rules
a. R_1 applies to Roots in class 1 in environments $\{F_1 \ldots F_n\}$
b. R_2 applies to Roots in class 2 in environments $\{F_1 \ldots F_n\}$ etc.

This analysis accounts for the facts. But the fact that each of the rules *R* applies in the same (non-natural) set of environments is listed with each rule, a loss of generalization.

11.5.2 Morphophonological abstraction

The loss of generalization identified immediately above in 11.5.1 can be fixed by looking carefully at the ways in which morphology and phonology interact. I will assume here that morphologically conditioned phonological rules reflect the operation of diacritics. In the simple case when a morpheme with a diacritic $[\alpha]$ triggers a phonological change, the phonology interprets such a diacritic by activating a rule: $[\alpha] \rightsquigarrow R$, using $\rightsquigarrow$ for the relation between a diacritic $[\alpha]$ and a rule *R*. For Disjoint Conditioning, each member *F* of a set of morphemes $\{F_1 \ldots F_n\}$ possesses a single diacritic $[\alpha]$, with the result that each of these is associated with the same morphophonological rule.

In the more complex case of a morphomic distribution, the diacritic view can be adapted by letting $[\alpha]$ activate different rules, depending on the context in which it appears (this idea is similar to proposals advanced in Trommer 2010). There are some different ways of implementing this intuition; I will outline one here, with the understanding that there are a number of related alternatives that could be explored to good effect.

One of the important properties of morphomic distributions is that the particular phonological rule that is triggered by $[\alpha]$ depends on the identity of the Root that is local to $[\alpha]$. In other work (Embick 2010b, 2012), I have hypothesized that morphophonological rules that require reference to the morphological identity of both the trigger and the target are restricted by (L1), i.e. they operate under concatenation$^\frown$. With this in mind, it is possible to specify the Roots with the

morphomic stem distribution with a diacritic [α] that is activated when local to the triggers $\{F_1 \ldots F_n\}$; that is:[21]

(38) Activate [α] in context [Roots ___]$\frown\{F_1 \ldots F_n\}$

What it means to be *active* in this sense is that [α] is visible for the phonology (and not for Vocabulary Insertion in a way to be explained below). In the case where there is a single phonological rule that is triggered in those environments (as is the case with German umlaut; (Wiese (1996)), this means that activating [α] and calling the phonological rule (symbolized with $\rightsquigarrow$) look like the same thing, even though there are two steps involved:

(39) Two steps (simple case revisited)
a. Activate [α]
b. Call phonological rule R ([α] $\rightsquigarrow R$)

My suggestion for morphomic distributions is built directly on the idea that diacritics call phonological rules in the way shown in (39). The idea is that in more complex cases, a single diacritic calls a number of rules, in a way that is determined contextually by the (list of) Roots with [α]; that is:

(40) a. [α] $\rightsquigarrow R_1$/List1$\frown$__
b. [α] $\rightsquigarrow R_2$/List2$\frown$__
etc.

In this analysis, [α] functions as an instruction to the phonology that indicates that a rule is to be called. Rather than calling a single rule, though, it functions as a variable whose value is determined by the context in which it appears.

There are a few observations to be made about this analysis. First, in this way of implementing morphologically conditioned changes, it makes sense to have the $\rightsquigarrow$ statements not block each other. This would allow some Roots to undergo both R_1 and R_2 (by virtue of being on both List1 and List2), while other Roots could undergo either one or the other, or some other combination when more Roots are involved. Although I cannot illustrate this point here, it seems typical for systems of

[21] For some purposes this might be equivalent to inserting [α] in the context shown in (38):

(i) Root $\rightarrow$ Root[α]/ ___$\{F_1 \ldots F_n\}$

For other purposes, activation and inserting features might be distinct. For example, Maiden (2005) argues that diachrony shows morphomically patterned Roots to 'converge' in different ways. In terms of the analysis in the text, the relevant changes could treat the Roots that possess [α] inherently as a class; inserting [α] as in (i) might not (depending on other assumptions) achieve the same effect. Since these questions go beyond the scope of the schematic argument that is offered in the main text, I will not explore alternatives here.

morphophonology to involve multiple rules that apply to different Roots sometimes individually, and sometimes in combination.

Second, as defined here, the [α] diacritics are active only in the phonology. As such, they are invisible to the Vocabulary Insertion process. This ensures that while such diacritics may be employed to produce Disjoint Conditioning of morphophonological rules, the same is not true for Vocabulary Insertion. This is important, because if a strong theory of syncretism is to be maintained, Vocabulary Items must not apply to non-natural classes of morphemes. One possibility is that diacritics like [α] are simply by definition only legible to the phonology proper, where they call specific rules. While it might be possible to derive this invisibility from other factors (or, e.g., from alternative implementations of the [α] analysis), I will not dwell further on this issue here, since it is one that can be investigated only in the context of a detailed case study.

Finally—and most importantly—introducing the [α] diacritics is able to account for morphomic distributions *while maintaining locality conditions on targets and triggers*. In order to be active, the [α] diacritic has to be in a local relationship with the triggers *F*. So, for example, if a morpheme intervened between *F* and Root from the list List1, [α] could not be activated, and the rule R_1 would not apply. Thus, the main argument levelled against PSD theories in sections 11.3 and 11.4 cannot be applied against theory of [α] diacritics.

In summary, this section shows that the main argument advanced for separating derivation from distribution, morphomic stem distributions, can be accounted for in a way that does not require a PSD theory of stems. I will leave for future work the question of whether there truly are morphomic distributions (and the [α] diacritics that might be used to analyse them).

11.6 Discussion

The central theme of this paper is how much Separation there is in the grammar. With respect to stem alternations in particular, the theoretical question is whether stem derivation and distribution go hand in hand (DD); or whether stems are represented independently of any particular distribution, and then associated with paradigmatic cells (PSD). The primary line of argument advanced in sections 11.3 and 11.4 above is that DD theories make strong and evidently correct predictions about the contextual locality between the triggers and targets of stem alternations, whereas PSD theories do not make these (or any other) predictions about the role of context in defining stem distributions. For these reasons it is concluded that PSD should be rejected in favor of DD. This argument validates the intuition that is at the core of many different theories; while I advanced one particular DD approach above, it bears repeating that the precise nature of contextual locality effects is

being investigated actively in a number of different theoretical frameworks (recall section 11.3).

The clearest argument for separating stem derivation from stem distribution is based on morphomic patterns of stems. Putting aside the question of whether such distributions are actually found, section 11.5 shows schematically how it would be possible to treat morphomic distributions without the further separation that is inherent in PSD approaches. The proposal outlined in section 11.5 is centred on the idea that phonologically active diacritics possessed by certain Roots (and perhaps other types of morphemes) may call phonological rules in a way that is contextually determined. This proposal connects with a number of important and difficult questions in morphophonological theory, concerning the nature of morpheme-specific phonological processes in general, and how particular ways of representing such information make predictions in both the synchronic and diachronic domains.

The arguments developed above concentrate exclusively on an empirical matter, concerning DD and PSD theories' predictions about contextual locality effects. But there is more to the opposition between these two types of theories than that. There is also a fundamental difference in research intuitions instantiated in the approach on stems advanced here, versus the one assumed by advocates of the morphome. As far as I can tell, Aronoff (1994, 2012) seems content to make the argument that morphomic distributions exist, and to limit himself to the conclusion that stem derivation is separate from stem distribution. Effectively, this means that there is some part of 'morphology' that is sufficiently by itself to deflect any attempts at further analysis that connects with other parts of the grammar (syntax, phonology, probably semantics as well). In the synchronic realm, anyway, it appears that there is little more to do with the morphomic theory, if it turns out to be correct (I am exempting here, for example, Maiden's work, which raises interesting diachronic questions). Specifying how (a finite number of) stem alternants are associated with paradigmatic cells does not raise any questions of theoretical significance. If there is more to the (synchronic) morphome programme than this, it is not clear what it is; and recent work in the morphome tradition, such as Aronoff (2012), is emphatically short of guidance on this point.

Fortunately, there are ways of cutting against the pessimistic grain. Most theories that address (Q1) and (Q2) are actively exploring form/meaning connections in ways that make questions about morphology resonate with theories of syntax, semantics, and phonology. For example, the (conjectural) discussion of the [α]-diacritic analysis in section 11.5 surrounds a type of feature that appears to serve an exclusively morphological (or, more properly, morphophonological) function. But even with an element of this 'morphological' nature, the questions that must be investigated involve connections with other grammatical systems (e.g. the relation of such features to individual Roots, or synsem features; the relation of [α] diacritics to other aspects of phonological organization; etc.). Positing a purely morphological

or morphophonological component to the theory is thus not an end in itself; it is a move that raises questions that have implications for several grammatical domains. Even more importantly for the future development of linguistic theory, specific answers to these questions are being used to investigate the pscyho- and neurological computation of language, in ways that are laying the foundations for a truly integrated theoretical framework (for different aspects of this, see e.g. Embick and Marantz (2005), Stockall and Marantz (2006)], Poeppel and Embick (2005), Embick (2010b), and Marantz (2013b)).

Research intuitions are important because they connect with frameworks for guiding investigation and interpreting empirical results. The research programme advanced in this paper and in related work is about as far as it is possible to get from 'morphology by itself'. It is founded on the idea that morphological phenomena must be analysed with reference to articulated theories of syntax and phonology. In my view, it is only within such a framework that the difficult questions about form/meaning relations in language can be addressed meaningfully, and I hope that this paper provides a concrete instantiation of this intuition put into practice.

Acknowledgements

Thanks to the editors for including me in this project and for feedback on an initial draft, and to Akiva Bacovcin, Rolf Noyer, and Kobey Shwayder for very helpful discussion of the issues addressed in the paper. The author's research is supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD073258.