Localism versus Globalism in Morphology and Phonology

David Embick
University of Pennsylvania

Version of December, 2008

Contents

1

Introduction: (Morpho)syntax versus (Morpho)phonology

Theories of grammar (and of language more generally) make specific claims about how the different facets of language are analyzed, often in ways that create partitions that are at odds with descriptive works, and, notably, at odds with each other. It is not difficult to find linguistic theories that take prima facie radically different views on large-scale issues of organization. However, in spite of the fact that different theories propose very different models of the grammar at an architectural level, the questions involved in distinguishing among competing theories are often quite subtle. Although a fair amount of the discussion of these issues derives from conceptual concerns, often where the contrasting positions are compared in terms of how they line up with opposing *research intuitions*, the ultimate assessment of questions of this type is empirical. So, to move closer to the topic of this work, there is no way of knowing based on conceptual or *a priori* considerations whether or not, e.g., "phonology" and "morphology" constitute one component of the grammar, or more than one. This is a question that has to be determined by taking specific models that make competing claims about these facets of linguistic knowledge, and comparing the predictions that these models make. While conceptual considerations about a particular type of explanation are discussed to some extent below– mostly to highlight why the empirical questions are the most important– part of what must be emphasized from the outset is that the crucial comparisons are always to be found in the empirical predictions made by different theories.

1.1   Localism/Globalism; Serialism/Parallelism

The question that is central to this monograph concerns how the system (or systems) responsible for deriving and representing the morphological or morphosyntactic properties of complex expressions is related to the system that computes the phonological form of these expressions. In terms that look ahead to the details that are examined below, this is the question of whether morphology is computed in the same system as phonology, or whether these parts of language are computed by distinct linguistic systems.

There are two specific parameters that are implicated by this question, that is, two diametrically opposed pairs of positions in terms of which current research programs can be defined. These are, first, whether the grammar functions in *Localist* terms, or whether *Global* considerations play a role in the grammar; and, second, whether the grammar employs *Serial* computations, or whether computation proceeds in *Parallel*. In the contemporary theoretical context, it is striking that the prevailing views in syntactic theory and in phonological theory offer opposing views on these basic questions.

In syntactic theory, the Minimalist Program of Chomsky (1993) and subsequent work continues a great deal of earlier research in advancing a theory in which syntactic relations are inherently local. Particular emphasis in this approach is placed on the idea that derivations are serial. Each computational operations is given a step in a derivation, and these computational steps are ordered

so that the output of one step is the input to the next. Serial derivation enforces a kind of Localism, by restricting the information that is available at any particular stage of computation. A further programmatic position is that, in addition to serial derivation, the sound and meaning interface systems play a defining role in syntactic explanation. This program and the theories that derive from it are *Localist* and *Serialist* in nature.

Phonological theory at present is dominated by Optimality Theory (McCarthy and Prince 1993, Prince and Smolensky 1993), which takes a *Globalist* and *Parallelist* view of the grammar. Optimality Theory dispenses with many of the assumptions of earlier generative phonology, in which an underlying representation is subjected to a serially ordered set of rules that effect local changes to the representation, and ultimately derive a surface form. This view is replaced by an architecture in which an input form is paired with a set of potential surface expressions, with a system of ranked constraints selecting the winner of this competition. A defining property of this type of view is that the factors that force a change in the output relative to the input need not be structurally close to the locus of the alternation.

Another defining property of Globalist theories like Optimality Theory is that morphology and phonology are not serially related to one another, but are instead computed in the same system. This architectural premise constitutes another departure from earlier models of phonological computation. In Chomsky and Halle's (1968) *The Sound Pattern of English* (SPE) and later versions of generative phonology, morphological processes are followed by phonological rule application. Although these distinct systems are interleaved in some theories (e.g. Lexical Phonology and Morphology, as in Kiparsky 1982), the ways in which they can interact are restricted by their serial organization.

The opposing positions defined by Serialism versus Parallelism and Localism versus Globalism are particularly acute in the domain of morphology, where current theories of (morpho)syntax and current theories of (morpho)phonology take positions that are simply incompatible with each other.

The morphosyntactic theory developed here, Distributed Morphology, takes a Localist and Serialist view of syntax and sound (and meaning as well), holding that phonology interprets the output of the syntactic derivation. In frameworks like Optimality Theory, as just mentioned, morphology and phonology are computed in the same system. In such a theory, it is predicted that phonological constraints may in some cases outrank syntactic or morphological constraints. This model therefore allows the morphological properties of an expression to be determined by output phonology, or by the global properties of surface forms, in ways that cannot be formulated in Localist and Serialist theories. This type of prediction is especially important in the domain of allomorphy, as will be made clear below.

While the theories discussed above differ in practical terms, in the sense that research in Distributed Morphology is more oriented towards syntax, and research in OT more oriented towards phonology, they overlap considerably in terms of what they seek to explain, and it must be asked directly why they differ so fundamentally. The opposing views of grammar hypothesized by these frameworks makes for a sort of schism between (morpho)syntax and (morpho)phonology. To a first approximation, this schism suggests two possible discoveries. The first is that one of the two theories is simply incorrect. The second is that they both are correct, and that morphosyntax and phonology are distinct and disconnected systems, in some profound sense. These are large points, and progress can be made by comparing the different predictions made by Localist/Serialist and Globalist/Parallelist theories in key domains, where each has something to say.

Since the primary point of contention centers on whether grammar functions in local terms, or whether at least some global considerations play a role in computation, the terms *Localist* and *Glob-*

alist are used throughout the book for the two types of architectures just outlined. These headings are used to refer both to different types of frameworks, and to specific theories that can be framed within these architectures. In the course of the argument in this monograph, the specifics of different proposals are articulated; as an initial step, it is useful to specify some reference points concerning Global interaction in particular, so that the scope of the discussion is clear.

For the part of this book that compares architectures, the primary question is whether there is a single computation in which the morphological form and phonological form of morphemes is determined simultaneously, with the potential for global interaction. Different types of Globalist answers can be framed to this general question. A *fully* Globalist theory of morphology and phonology would hold that the morphology and phonology of entire words is computed in a way that allows for interaction between structure, allomorphy, and sound; perhaps with syntax included in this computation as well (cf. McCarthy 2002:142). *Limited* Global interaction can be implemented as well. For example, in stratal or cyclic versions of OT, only subparts of a given word are subject to simultaneous morphological and phonological computation. While theories of this type rule out fully global interactions across entire words, they nevertheless predict that in a given cyclic domain, there should be global interaction among morphology and phonology.

In the course of examining specific theories below, the finer distinctions between full and limited Globalism are made when required. The overall point, though, is that theories with even limited global interaction between morphology and phonology make very different predictions from Localist theories about how phonology and morphology can interact, and this allows for a direct comparison of the different frameworks.

1.2   (Phonologically Conditioned) Allomorphy

This monograph examines the predictions that Localist and Globalist theories make for *allomorphic* interactions. Allomorphy in the broad sense is a term that covers any variations in the surface form of a morpheme. Whether all such variations are the result of one type of operation in the grammar, or different operations, is something that different theories make different claims about.

As an initial example of allomorphy, consider the behavior of the past tense morpheme T[past] in English. According to a standard analysis, the default shape of this morpheme is *-ed* (i.e. /d/), as in *play, play-ed*. As is well-known, the past tense morpheme has allomorphs besides *-ed* which occur when T[past] occurs next to other verbs; putting aside processes that change the phonology of the verb stem itself (such as *broke* from *break*), a simple description is given in (1):

(1)   Allomorphs of T[past] in English

    a.   *-Ø*: *hit/hit-Ø*, *sing/sang-Ø*, *break/broke-Ø*, etc.

    b.   *-t*: *bend/ben-t*, *leave/lef-t*, *buy/bough-t*, etc.

    c.   *-ed*: Elsewhere

Allomorphic interactions of this type appear to be highly constrained. Informally, for allomorphic purposes one node sees another only when the two nodes are "close" to each other in a way that must be made precise.

The kind of allomorphy exhibited by English T[past] is *grammatically conditioned*. Knowing whether a particular verb takes a particular allomorph from (1) is something that does not follow from other factors. In particular, it is not predictable based on the phonology of the verb. Rather, the

conditioning element is a locally visible, grammatical object; in the case of (1), the identity of the particular verb that the node T[past] is attached to.

This kind of allomorphy is a type of what is called *contextual allomorphy*. Something in the grammar specifies that the pronunciation of T[past] has one of the non-default forms in (1) (i.e. either (1a) or 1b)) when it occurs in the *context* of another element. Part of any theory of morphology is the theory of the conditions under which elements can show contextual allomorphy in this way. The first part of this monograph develops a Localist and Serialist theory of allomorphy, in which linear adjacency and cyclic locality interact to produce a constrained theory of allomorphic interaction.

A second type of allomorphy, which allows for a direct comparison of Localist and Globalist frameworks, is *phonologically conditioned allomorphy* (PCA; see Carstairs (1988) and subsequent work). This is a type of contextual allomorphy in which the choice of a particular allomorph of some morpheme is determined by phonological factors. Some examples are given in (2):

(2)   a.   Korean nominative suffix

Allomorph	Env.	Example	Gloss
-i	/C__	pap-i	'cooked rice'
-ka	/V__	ai-ka	'child'

b.   Seri passive suffix (Marlett and Stemberger 1983)

Allomorph	Env.	Example	Gloss
p-	/__V	-p-eši	'be defeated'
a:ʔ-	elsewhere	-a:ʔ-kašni	'be bitten'

c.   Haitian Creole definite suffix

Allomorph	Env.	Example	Gloss
-la	/C__	liv-la	'book'
-a	/V__	tu-a	'hole'

These examples are chosen to illustrate two types of effect that are found in PCA, as viewed from the perspective of the output phonology of the affixed word.

The first case, Korean *-i* and *-ka*, is a case where the distribution of allomorphs could be seen as having a phonological motivation. The vowel *-i* after consonants creates syllables that are "better" than those that would be created by affixing *-ka* to such forms. Similarly, affixing *-ka* to vowel-final hosts avoids the hiatus that would be created by the affixation of *-i*. In this sense, it might appear that the "morphological" choice of allomorphs is driven by the output phonology, in a way that fits nicely with a Globalist phonological theory in which syllable-structure markedness constraints that favor CV- syllables without codas can effect allomorph selection.

The Seri passive morpheme shows this kind of effect in part. While affixing *p-* to vowel-initial hosts seems phonologically natural, affixing *a:ʔ-* to consonant-initial hosts creates a coda. A "better" type of affixation in the latter case would, all other things being equal, involve a prefix with the shape (C)V-. In terms of the phonological forms that the language happens to provide for the realization of the passive morpheme, however, the distribution of allomorphs could be seen as phonologically optimal.

Finally, the Haitian Creole allomorphy is "perverse" from the perspective of syllable-structure markedness. Affixing *-la* to consonant-final hosts creates codas, and affixing *-a* to vowel final hosts creates hiatus. Each of these problems would disappear if the reverse distribution of allomorphs obtained.

Intuitively, the importance of PCA as a case-study derives from the fact that it involves the interaction of morphological and phonological factors in the determination of a form.

In the Localist theory developed in Part I of this monograph, all of the cases of contextual allomorphy seen above receive the same analysis. The theory says that the phonological "spell out" of a morpheme, which occurs in a process called *Vocabulary Insertion*, can be sensitive to items that are in the local environment of the morpheme being spelled out. Crucially, while this theory can account for the distributions in (2), it cannot say *within the grammar itself* that these distributions happen for a reason; i.e., that they are (in part) driven by surface phonological optimization. This theory can generate the forms that it derives mechanically, but it does so without reference to ultimate output forms; in this sense, it is a theory of morphology without teleology.

In Globalist theories like Optimality Theory, the architecture allows phonological constraints to determine allomorph selection. The reason for this is that morphology and phonology are one system, in which phonological constraints can outrank morphological ones. It is therefore possible in such theories to say that allomorph selection– part of the morphology– happens the way it does *because* of the way that affixation creates particular phonological patterns. In the Korean case (2a), for example, it is possible to give an analysis in which the candidates that are in competition consist of the host plus each of the different allomorphs, so that both e.g. *pap-i* and *pap-ka* are generated for the input "nominative of *pap-*". The constraint ranking– and phonological constraints governing syllable structure in particular– then work together to derive the pattern of allomorph selection. In such a theory, it is possible to say *in the grammar* that the distribution of allomorphs is the way it is for a reason.

A theory with the capacity to say that surface forms drive allomorph selection must have certain formal properties; in particular, it must have at least a limited amount of Global interaction in morphology and phonology. Because of this, such theories make predictions that are distinct from those that derive from a Localist view. These differences can be seen clearly by looking at the general question of which factors could conceivably play a role in determining the phonological form of a complex expression.

1.3 Surface Forms, Competition, and The Schism

Taking grammars to be theories of how sound/meaning connections are derived, it can be asked at a very general level what different theories have to say about the factors that may play a role in determining the surface form of an expression.

Optimality Theory implements Global and Parallel computation by generating an infinite set of output candidates for any given input, with constraints selecting a winner from these competitors. The output candidates differ from the input in ways that potentially involve more than one phonological "change". This computation of forms is Global in at least two ways: first, because it is anti-modular, phonological and morphological constraints can interact in a manner that is not possible in alternative theories; and, second, because the constraints could be ranked in such a way that there are non-local interactions within a word.

The central principle that allows output forms to be compared for well-formedness is *competition*. Competition is a central concept in grammatical theory. It is implicated in morphological discussions in the study of *blocking* effects, initiated in the modern era in work by Aronoff (1976). According to Aronoff, for example, the word **gloriosity* is derived by the rules of the grammar, but cannot be the "abstract noun for GLORY" because *glory* exists and blocks it. In order for this analysis to work, the grammar must supply more than one object for the potential expression of

a particular meaning (in this example, both *gloriosity* and *glory*), and it must supply a means of determining the winner of this competition.

Part of the OT program is the idea that surface forms are the way they are for a reason, and that the grammar must state these reasons directly. In order to implement this idea, competition is required. From the infinite set of possible output forms, the winner is the one that is optimal with respect to the constraint ranking. If there were not multiple competitors– i.e., if the grammar only made available one representation in any given computation– then there could be no "optimization".

The two types of potentially Global interactions mentioned are a consequence of having infinite competition in this way. The fact that phonological and morphological constraints interact to select a winner means that in principle, phonological properties of surface forms could determine what happens morphologically, by forcing a particular affix to be selected because of its effects on the phonology of the whole word.

The Globalist perspective on phonological form is incompatible with the view of the grammar that is advanced in Localist morphosyntactic theories like Distributed Morphology. The source for many of these incompatibilities is highlighted above: competition.

The prevailing view of "blocking effects" in the broad sense is that they require competition of the type outlined by Aronoff; this can be implemented in various ways. More recent work argues that there is no blocking of the type discussed above; this is the conclusion presented in Embick and Marantz 2008 and and Embick 2007a. These papers examine arguments for blocking among words and larger expressions, and conclude that there is no motivation for a competition-based analysis of such phenomena. Rather, put somewhat simply, what surfaces in the grammar is what is derived by the grammar; other putative competitors for a particular meaning are simply never derived, and therefore do not need to be blocked.

According to the theory of Embick and Marantz (2008), competition is strictly local: it is restricted to the procedure that determines the phonology of a single node, the Vocabulary Insertion operation mentioned above. A consequence of this view is that there is no competition among complex objects; i.e. no word/word, word/phrase, phrase/phrase competition. In short, complex objects are assembled in syntactic structures, and this simultaneously accounts for how they are represented, and how they are distributed.

Importantly, this Localist theory also has consequences for phonological relatedness, especially the shared properties of *lexically related* forms like *plays*, *played*, etc., where it places a number of restrictions. Specifically, the theory says that the phonological form and phonological relatedness are derivative of the following factors:

- Complex, lexically related forms are built in syntactic structures and contain the same Root.

- In a given structure (Roots, functional heads), a single output is derived; this output is what exists, and therefore what must be used in that grammatical context.

- Complex, lexically related forms share phonological material in a consistent way because they

    – are based on the same Root, which has an underlying representation (UR)

    – appear in syntactic structures whose heads have consistent phonological expression (up to allomorphy)

    – the phonology involves the same rules/constraints (up to exceptionality that must be listed).

These answers are directly related to the fact that this theory has limited competition. In the course of any derivation, only one object is produced. It is thus not possible to generate multiple competitors and select a winner based on properties of the output. This precludes, among other things, generating a word with all of the different allomorphic possibilities the language allows, and then choosing the winner on the basis of e.g. phonological well-formedness.

In sharp contrast to the Localist view just outlined, the essence of Globalism as manifested in Optimality Theory is unlimited competition, and the essence of competition is that there be multiple possible outputs for any given input. This is exactly what the Localist morphosyntactic theory says is impossible. Putting these different incompatibilities into focus, it is clear that these views of morphosyntax and morphophonology define a schism:[1]

(3) THE SCHISM: Globalist theories of morphophonology require competition between multiple potential expressions of a given input; according to the Localist morphosyntactic theory, this is impossible because such competitors are not derived by the grammar.

By concentrating on the schism, the goal of the second part of this monograph is to bring empirical arguments to bear on the large-scale architectural matters implicated by (3). As mentioned in 1.1, there are two possible discoveries that could stem from focus on the Schism, and each of them is significant. The first possible discovery is that (morpho)phonology is simply profoundly different from (morpho)syntax. It is in principle possible to construct a theory in which each of the two views above is correct: i.e., "No Competition" is correct for morphosyntax, and then "Competition" is correct for morphophonology. In such a theory, the syntax and morphology operate in terms of local, serial derivations, but the output of this system in some part of the phonological computation involves multiple or infinite competitors, so that global considerations can play a role in the determination of surface forms. One question to ask is whether this would be a sort of "worst case" scenario, architecturally speaking, since it would divorce the system of combinatorics from the system for computing sound forms in an extreme way. The other possible implication of the schism is that either the Localist or Globalist theory is untenable; i.e. (i) that the "generative" Localist view of (morpho)syntactic theory is incorrect, or (ii) that the Globalist, competition-based theory of (morpho)phonology is incorrect.

These are large points, and they resonate with other aspects of grammatical theory in numerous ways.

This monograph is divided into two major components. Part I develops a Localist theory of allomorphy. Part II makes explicit comparisons between Globalist theories of various types with the core predictions of a Localist framework. The core proposals are that the Localist theory of Part I makes correct predictions about allomorphy in natural language, and that the predictions of Globalist theories are not found.

1.4 Prospectus: A Localist Theory

Part I of this monograph articulates a Localist theory of contextual allomorphy. The defining properties of this theory are that patterns of contextual allomorphy are defined by both phase-cyclic and linear notions of locality.

Contextual allomorphy in Distributed Morphology results from the operation of *Vocabulary Insertion*. Morphemes are terminals in a syntactic structure. Some of these morphemes, the func-

---

[1]In the broader background it is worth noting that this sort of tension has been discussed from the reverse perspective; see Bromberger and Halle (1989).

tional heads, have no phonological form as part of their underlying representation. Rather, these morphemes receive phonological content in the PF component of the grammar. This is the role of Vocabulary Insertion; individual *Vocabulary Items* (VIs) in memory compete for insertion at a given node, and the most specific that can apply gives that node its phonological matrix. In the example of the English past tense, the syntax generates a structure that contains the past tense node T[past]. In the PF computation, the Vocabulary Items in (4) compete for insertion:

(4) Vocabulary Items for Tense

$$
\begin{array}{lll}
\text{T[past]} & \leftrightarrow & \text{-t/}\underline{\quad} \ \{\sqrt{\text{LEAVE}}, \sqrt{\text{BEND}}, ...\} \\
\text{T[past]} & \leftrightarrow & \text{-}\emptyset\text{/}\underline{\quad} \ \{\sqrt{\text{HIT}}, \sqrt{\text{SING}},...\} \\
\text{T[past]} & \leftrightarrow & \text{-ed}
\end{array}
$$

When Roots like $\sqrt{\text{BEND}}$ and $\sqrt{\text{HIT}}$ are present, the Vocabulary Insertion process inserts *-t* and *-Ø* into the T[past] node respectively; in other cases, the default *-ed* is inserted.

The general research question that motivates this work centers on the factors that could play a role in contextual allomorphy. According to the specific view that is developed below, possible patterns of allomorphy are determined by the interaction of distinct sets of locality constraints. The core intuition is as follows: contextual allomorphy, where one node X can see another node Y for the purposes of Vocabulary Insertion, is possible only when X and Y are concatenated; i.e., in the most local linear relationship possible. This linear condition interacts with a further set of restrictions that are imposed by the assumption that syntactic derivation proceeds in terms of *phases* (in the sense of Chomsky 2000, 2001) that are spelled out cyclically. Phase-based derivation places sharp constraints on the amount of information that is available in a particular cycle of PF computation, and restricts potential allomorphic interactions accordingly.

For cyclic derivation, the theory presented below assumes with Marantz (2007) and Embick and Marantz (2008) that *category-defining* heads like v, n, and a define phases. According to this view, heads of this type *categorize* the elements that they attach to. So, for example, a head v merged syntactically to a $\sqrt{\text{P}}$ headed by a category-neutral $\sqrt{\text{ROOT}}$ creates a vP (5); when the Root and the v head are combined into a single complex head as shown in (6), the result is a "verb":

(5) v merged with $\sqrt{\text{P}}$
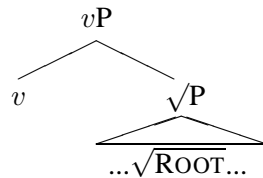
(6) Complex head
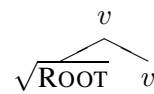
The category-defining heads are *cyclic* in the sense of phase theory. What this means is that when they are merged to a structure, they trigger spell out. Other heads that appear in complex words, such as Tense morphemes, plural morphemes, etc., are not cyclic in this way. This difference between cyclic and non-cyclic heads is manifested in many domains, including possible allomorphic interactions.

The example in (5-6) shows a single cyclic head v attached to a Root. Category-defining heads may also be merged to structures that are already categorized. So, for example, a verb like *break*, which is a Root combined with v, may be combined with a "potential" adjective head a to yield *breakable*, an adjective derived from a verb, as shown in (7):

(7) $[[\sqrt{\text{BREAK}} \, v] \, a]$

When a category-defining head is the first that is merged with a Root, as is the case with v in (5) and (7), the this head is said to be Root-attached, or in the *Inner* domain. When a category-defining head is attached to a structure that has already been categorized, like the a in (7), the additional cyclic head is said to be in the *Outer* domain.

A central idea in linguistic theory is that cyclic domains define possible interactions for syntax, phonology, and semantics. One proposal that has been discussed in the literature is that syntactic configurations in which a Root is merged with a category-defining head– the Inner domain– appears to be special for the purposes of both sound and meaning. In the formulation of Embick and Marantz (2008), the generalizations about what is special about this inner domain are as follows:

(8) Cyclic Generalizations

    a. **Allomorphy:** For Root-attached x, there may be special allomorphy, determined by properties of the Root. A head x in the Outer domain is not in a local relationship with the Root, and thus cannot have its allomorphy determined by the Root.

    b. **Interpretation:** The combination of Root-attached x and the Root might yield a special interpretation. The heads x attached in the Outer domain yield predictable interpretations.

For the purposes of a Localist account of allomorphy, what (8a) highlights is the possibility that contextual allomorphy could be found only with Root-attached cyclic nodes.

An important discovery in this context is that a "Root-attached" theory of contextual allomorphy is too restrictive. This point is discussed with reference to allomorphy in participles in Embick 2003, and arises in cases like the English past tense as well. English past tense verbs have a structure consisting of a Root, a v head, and the node T[past]:

(9) English Past Tense
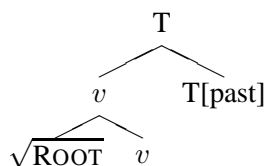
The T[past] node shows contextual allomorphy, yielding the familiar allomorphs in e.g. *ben-t* and *hit-Ø* versus the default *-ed* in e.g. *play-ed*. Crucially, the T[past] node is not Root-attached, but nevertheless shows irregular allomorphy conditioned by the Root, contrary to what is predicted by (8a).

The conclusion that emerges from examples of this type is that the most restrictive phase-cyclic account of allomorphy (8a) is incorrect. The challenge is therefore to present a theory that is capable of accounting for the attested patterns of contextual allomorphy, while nevertheless being restrictive enough to make strong empirical predictions.

While the past tense type of case outlined above suggests that heads outside of the Inner cyclic head may show Root-determined allomorphy, the possibilities for allomorphic interaction are still sharply constrained. The restrictions are of two types.

First, it appears that a morpheme can show contextual allomorphy determined by another morpheme only when these two pieces are linearly adjacent to one another; i.e., when no overt morpheme appears between the two. This generalization suggests a strict linear constraint on allomorphic interactions.

Second, while the cyclic theory based on (8a) is too restrictive, cyclic structure is still relevant for allomorphic interactions. This is clear from another generalization: it appears that Outer cyclic heads cannot show contextual allomorphy that is determined by elements in the domain of an Inner cyclic head. So, for example, in a "category-changing" structure with two cyclic heads x and y like (10), the Outer cyclic head y never shows Root-determined allomorphy:

(10)   Structure with two cyclic heads
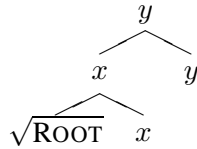
An example of this is provided by English gerunds, like *John's destroying the files*. Unlike special nominals, like *laugh-te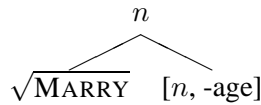r*, *marr-iage*, *destruct-ion* and so on, which show a great deal of Root-determined allomorphy, gerunds always show *-ing*: *laugh-ing*, *marry-ing*, *destroy-ing* etc. In special nominals, the n head is Root-attached. In Gerunds, on the other hand, the nominalizing n morpheme attaches to structure that is verbalized by v. The structures at play here are those in (11) and (12):

(11)   *marriage*

(12)   *marrying*
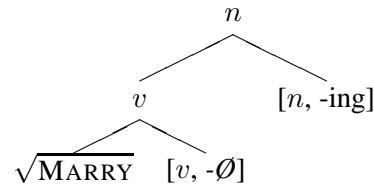
This Outer n seen in (12) shows no Root-determined allomorphy: it is always has the phonological form *-ing*, even though it is superficially adjacent to the Root.

There is thus an asymmetry between non-cyclic and cyclic heads in allomorphy: the former type of head can see across an Inner cyclic node, but the latter type of head cannot. With α as the element potentially conditioning allomorphy, these important generalizations are schematized in (13), where lower-case x, y are cyclic heads, and upper case Z is a non-cyclic head:

(13)   a.   ...α] x] Z]

   Generalization: Non-cyclic Z may show contextual allomorphy determined by α, as long as x is not overt

b.   ...α] x] y]

   Generalization: Cyclic y may *not* show contextual allomorphy determined by α, even if x is not overt

The asymmetry in (13) presents a further empirical challenge for a restrictive theory of allomorphy: not only must the cyclic theory be extended to allow the (13a) cases; the extension must be executed in such a way that Outer cyclic heads in (13b) cannot be sensitive to elements in x's complement.

The theory of Part I faces this challenge by developing a theory in which linear and cyclic locality interact to constrain allomorphic interaction, as stated in (H1) and (H2):

(H1) Contextual allomorphy is possible only with elements that are concatenated.

(H2) Cyclic spell out domains define which nodes are present in a given cycle of PF computation, and thus potentially "active" (capable of being referred to) for the purposes of contextual allomorphy. In some cases, superficially adjacent nodes cannot influence each other allomorphically because in terms of cyclic spell out, they are not active in the same PF cycle.

The linear condition in (H1) is straightforward: it holds that one node can only show contextual allomorphy determined by another node when the two are immediately next to one another; i.e., when there is no intervening morpheme.

The essential properties of the cyclic part of the theory (H2) can be illustrated with reference to (14), where lower case x and y are cyclic heads, and upper case W, Z are non cyclic heads:

(14)   a. $[\sqrt{\text{ROOT}}\ x]$
        b. $[[[[\sqrt{\text{ROOT}}\ x]\ W\ ]\ Z\ ]\ y]$

The basic premise of the theory is that cyclic heads force the spell out of cyclic domains in their complements. With reference to (14), this means that when x is merged syntactically in (14a), there are no cyclic domains in the complement of x, so that there is no spell out in this particular case.

Subsequent syntactic derivation merges non-cyclic W and Z. When the head y is merged (14b), a PF cycle is run on the cyclic domain centered on x. This includes the Root, x, and the non-cyclic heads W and Z. In this cycle, Vocabulary Insertion occurs at x, W, and Z, and gives phonological form to these morphemes. These heads could potentially show Root-determined allomorphy, as long as no overt morphemes intervene.

Later in the derivation, another cyclic head (not shown in (14)) triggers spell out of material in its complement, which includes the phase centered on y. The elements that are present in this PF cycle are x (the *edge* of the xP phase), W and Z, and y. Crucially, while y could show contextual allomorphy determined by x, W, or Z, it could not show Root-conditioned allomorphy. The reason for this is that the PF cycle in which y is given phonological form does not involve the Root; it (and other elements that could be in the complement of x) are derivationally closed off.

The cyclic aspect of the theory restricts the amount of information that is available to condition the allomorphy in two ways. First, in a complex word that contains multiple cyclic domains, the computation of the phonological form of the inner domains takes place at a stage when the "Outer" material is not present. This outer material can therefore play no role in determining the phonological form of inner nodes. Second, for computation in Outer cyclic domains, inner material is inaccessible, in the way outlined above. As a result, Outer cyclic nodes cannot be influenced by certain nodes in the Inner domain, restricting potential allomorphic interactions further.

In sum, the guiding insight of the theory presented here is that the interaction of (phase) cyclic domains and a strict linear notion of locality are responsible for possible patterns of contextual allomorphy. Reflecting the interaction of cyclic and linear factors, the approach that is advanced in

Chapter 2 is called the "$\mathbb{C}_1$-LIN" theory, where the $\mathbb{C}_1$ stands for the cyclicity condition, and LIN stands for the linear condition.

After developing the details of this theory in Chapter 2, a number of illustrations and consequences of the approach are presented in Chapter 3. This includes a discussion of (linear) Intervention Effects, and cyclic Edge Effects, along with a series of more complex case studies examining how "the same" pieces of morphology can appear in different cyclic domains. Some comments concerning how morphosyntax and morphophonology interact in the theory are also presented, paving the way for some aspects of the comparison of frameworks in Part II.

1.5 Prospectus: Localism versus Globalism

Part II returns to the fundamental tension between Localist morphosyntax and Globalist phonology outlined earlier in this chapter, by looking at the empirical predictions that these theories make for allomorphy.

The morphosyntactic theory developed in Part I of the book restricts competition in the grammar to allomorphy of a single node: this is the process of Vocabulary Insertion. The theory thus disallows competitions in which multiple competitors like "words" are derived and compared for well-formedness. This effectively restricts the factors conditioning a case of contextual allomorphy to elements in the immediate context of the node being spelled out.

This view differs fundamentally from that offered by a Globalist theory of the type that is assumed in much current work on phonology. In a theory like Optimality Theory, in which the grammar generates an infinite number of candidate expressions that are potential surface realizations of a given input, the inputs involved are complex; i.e., they involve more than one morpheme. Thus, this theory is responsible for morphology as well as phonology. Since morphological and phonological properties are determined in the same computational domain, this type of framework allows for global interactions in which, for example, non-local properties of surface forms play the defining role in allomorphic selection.

The full range of predictions that separate the Localist and Globalist views on allomorphy emerge from an examination of the following closely interconnected questions:

- GLOBAL MORPHOLOGY/PHONOLOGY INTERACTIONS: Is there evidence that morphology and phonology are computed in a single, Global/Parallel system (*Global-MP*)? Or do the facts on interaction suggest an organization in which phonology acts on the output of allomorph selection, as in the Localist theory?

- PHONOLOGICAL SELECTION: Is there *Phonological Selection*, in which surface phonological well-formedness forces a choice among allomorphs, such that phonology drives allomorphy in ways that are impossible in a Localist theory?

- GLOBAL CONSIDERATIONS OVER LOCAL: Is there evidence that the factors determining allomorphy are global in any sense? I.e., are there cases in which Localist and Globalist approaches make different predictions about which allomorph should be chosen for a particular position, and the Global considerations win out, in a way that cannot be stated in a Localist theory?

Part II of the monograph begins in Chapter 4 with a look at the answers to these questions that derive from the architectural assumptions of Globalist theories like OT. The discussion centers

on the types of arguments that could, conceivably, provide evidence for the Globalist architecture. In paving the way for the empirical comparison, the conceptual ground is cleared with a crucial clarification. As discussed above in 1.2 with reference to the initial examples of allomorphy from Korean, Seri, and Haitian Creole, a Localist theory cannot say that a pattern of allomorph selection happens *because* of some output property, phonological or otherwise. To the extent that there are generalizations about surface forms to be made, the Localist theory can make them, but they must be derivative of another part of language in the broad sense. That is to say, not part of the grammar, but the result of diachrony, acquisition, etc.

As noted in 1.2 above, these considerations lead to a kind of conceptual argument that is often advanced in favor of Globalist theories. In theories with Global-MP it is possible to say that patterns of allomorphy happen *for a reason*, within the grammar. So, for example, the case of Korean *-i/-ka* allomorphy can be treated in terms of syllable structure constraints. An OT analysis can then say that the (phonological) grammar forces the attested distribution of allomorphs, and, moreover, the grammar *explains* the distribution by having morphological selection driven by optimization of the phonology of the output. The charge that is levelled against Localist theories is that, while they might account for the distribution of allomorphs, they do not account (within the grammar) for why the distribution is as it is. This type of argument is based on the Localist theory's *Putative Loss of Generalization*, or PLG. In the domain of phonological rules, the question of whether Localist theories are missing generalizations about outputs has been actively discussed since at least Kisseberth 1970. The same kind of considerations about patterns in surface forms motivate Globalist views of morphology/phonology interactions, and allomorph selection in particular.

In many cases that have been studied in the literature, Localist theories and Globalist theories are both able to account for the facts. In such cases, only conceptual arguments, such as appeal to PLG, can be deployed against a Localist theory; there is no empirical basis for determining which of the two frameworks is to be preferred. Rather, the choice reduces to whatever combination of conceptual, aesthetic, or other factors regulate the intuitions that individual researchers have about what explains what. A key point that moves the argument from the conceptual to the empirical is that Globalist theories predict a number of types of global interaction that simply cannot be expressed in the Localist theory. The direct comparison of frameworks must be directed at such cases.

The examination of these predictions goes in two steps. Chapter 5 begins by outlining the best case scenario for Globalist theories: the hypothesis that the phonological grammar determines all cases of Phonologically Conditioned Allomorphy (PCA), and nothing more needs to be said about allomorph distribution. This position was shown to be incorrect in early works exploring the Globalist research program like Kager 1996. However, the possibility remains that there are nevertheless *some* instances in which surface phonology drives allomorph selection, in ways that cannot be analyzed in a Localist framework.

In order to highlight the empirical issues, and the intuition behind the Globalist program, Chapter 5 then moves to an examination of *systems* of PCA. This part of the discussion is not a formal argument against Globalism or for Localism. Rather, it examines the intuition that Globalist theories are based on: the idea that patterns of PCA are the way they are for reasons that should be expressed in the grammar, and that these reasons should be phonological in nature. The empirical basis for this chapter is provided by systems of case endings found in two Australian languages, Djabugay and Yidiɲ, where there is a large amount of PCA. Although looking at isolated subparts of such systems might make it look like there is motivation for a Globalist theory in which output phonology determines allomorph selection, this impression disappears when the systems are analyzed in detail. The particulars of the analysis show that the case systems in these languages derive from the interaction

of stored information about the shape of morphemes with sometimes exceptional phonological and morphological rules, in a way that implicates serial organization between morphology and phonology. A further argument, extending this, is that Yidiɲ case allomorphy looks at a first glance like it might be driven by simple phonological constraints, but analyzing the system in surface-based terms obscures key generalizations about other aspects of the language's morphophonology.

Moving past intuitions, the ultimate comparison of different frameworks must be made on the basis of the different empirical predictions that they generate. The second step in the argument in Chapter 6 is centered on the fact that theories with even a restricted form of Global interaction between morphosyntax and phonology predict effects that cannot be stated in a Localist theory. These effects can be seen in cases in which a morpheme X has more than one phonologically conditioned allomorph, say x_1 and x_2, and X appears in words with other morphemes like Y and Z:

(15)   Root-X-Y-Z

There are cases of this type in which the *local* environment predicts insertion at X of the x_1 allomorph, while the *global* environment– i.e., phonological properties of the entire word– predicts insertion of the x_2 allomorph.

In a Localist theory of the type developed in Part I, choice of allomorph at X must be determined by grammatical or phonological information that is visible at the point when insertion occurs. In cases like (15), the Localist theory predicts that the locally-selected x_1 allomorph will be found.

On the other hand, in a Globalist theory in which morphology and phonology are computed in the same system, it is possible for the x_2 allomorph to be inserted, in a way that is driven by the output phonology. This prediction is not the exclusive property of "full" Globalist theories. Even restrained, cyclic Globalist theories make the same prediction, as long as the affixes in question are not in different strata. That is, the prediction that Z's form (or the form of the entire word) could affect allomorphy at X is made by any theory in which the computation of the morphophonology of X, Y, and Z occurs in the same domain.

The allomorphy of perfect heads in certain Latin verbs, discussed in Mester 1994, provides an example of the type schematized in (15). The perfect head in question has two allomorphs: *-u*, generally taken to be the default, and *-s*. Mester's argument is that choice among these allomorphs is determined by the prosodic structure of affixed words. Specifically, the non-default *-s* allomorph is inserted only when the *-u* allomorph creates a form with an unfooted medial syllable, what are called a (medial) *trapping* configuration. The idea is that the prosodic undesirability of trapping is what drives the insertion of the non-default *-s* allomorph with certain verbs.

The effects of this analysis are shown for the verbs *monēre* 'warn' and *augēre* 'grow' in (16). These verbs differ in the metrical weight of the stem (light versus heavy), which results in different metrical parses with the *-u* affix. These verbs show different allomorphs of the perfect head:

(16)   a.   Perfect Allomorph: *-u* with light Root
           [mon**u**]⟨ī⟩
       b.   Perfect Allomorph: *-s* with heavy Root
           *[au]g**u**⟨ī⟩ (trapping)
           [aug]⟨**sī**⟩

According to the Globalist theory advanced by Mester, the perfect morpheme has its allomorphy determined by the output prosody of the word. The grammar generates both *auguī* with the default *-u* allomorph, and *augsī* with the *-s* allomorph, and prefers the latter because of the output prosody.

14

In this case, the Globalist theory predicts– unlike the Localist theory– that the allomorph choice for the perfect may vacillate, depending on the phonological properties of outer affixes. In this particular case, the Globalist theory predicts that in pluperfects like those in (17), the allomorph selected for *augēre* should switch from *-s* to *-u*, because this yields a better prosodic structure (17b). But this does not happen; the grammatical form has the *-s* allomorph as in (17a), in spite of the fact that this creates trapping:

(17)   1s Pluperfect of *augēre*

      a. With *-s* allomorph:
           aug**s**eram = $[\bar{\sigma}]\breve{\sigma}\langle\bar{\sigma}\rangle$

      b. With *-u* allomorph:
           *aug**u**eram = $[\bar{\sigma}][\breve{\sigma}\breve{\sigma}]\langle\bar{\sigma}\rangle$

In this and other cases, the locally determined allomorph is selected, and there is no evidence for the type of Global interaction– allomorph vacillation based on output phonology– that would provide evidence for Globalism.

The general line of argument in Chapter 6 is that any sort of interaction of the type outlined above would be an argument for a Globalist view, but that no such interactions are found. In cases where this type of prediction can be seen, languages show local determination of allomorphs of the type that is predicted by the Localist theory.

As stressed above, the differences in predictions between Globalism and Localism are clearest when a "fully" Globalist position– i.e., one with interacting syntax, semantics, phonology, etc.– is considered, but cyclic theories with limited global interaction also make predictions that go beyond what the Localist theory allows. The arguments advanced in this monograph extend to theories with even highly restricted forms of global interaction: there is no evidence for global interaction in even the restricted form that could be stated on a cyclic OT theory. It seems that even a little bit of global interaction leads to incorrect predictions about allomorphy.

1.6   Implications

Chapter 7 synthesizes the consequences of Parts I and II of the monograph. If the theory of Part I is correct, then allomorphy is subject to strict locality conditions of a type that derive from a Localist syntactic theory. If the conclusions of Chapters 5 and 6 are correct, there are no empirical arguments for the strong predictions of Globalism. This point has implications for the status of generalizations about surface forms, along the lines of what was discussed under the heading of *Putative Loss of Generalization* above. In order to account for why certain patterns of allomorphy occur, a theory must have Global-MP. It is only in such a theory that the grammar can refer to properties of output forms in the allomorph selection process. However, theories with Global-MP make formal predictions about global morphology/phonology interactions that are not borne out. This is an argument against the Globalist architecture *and* an argument against the idea that the grammar itself must say why certain patterns of allomorph selection are found.

A second implication of this argument is that OT is a theory of phonology without a theory of morphology. There are many different potential responses to this line of argumentation, and almost all of them have deep consequences for theories of grammar. One obvious response would be to hold that there are fundamental differences between morphosyntax and (certain aspects of) phonology, and OT is a theory of the latter. Another possibility is that the type of Globalist system espoused by

OT must be abandoned, or modified in some extreme way. Importantly, since incorrect predictions about allomorphy appear to arise even in systems with a limited amount of Global interaction, appealing to stratal or serial versions of OT does not appear to be an adequate response.

My view– a research intuition– is that the success of Localist theories of morphology and morphophonology in this domain motivates a return to theories in which the sound form of complex expressions is linked as closely as possible to the generative procedure that builds them. This means programatically that there must be an assessment of the empirical arguments in favor of Localism and Globalism. At the very least, such investigations will make precise the dividing line between structure/sound relationships that function in Localist terms, and those that (putatively) require Globalism in some form.