Annual Review of Linguistics

The Motivation for Roots in Distributed Morphology

David Embick

Department of Linguistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; email: embick@ling.upenn.edu

Annu. Rev. Linguist. 2021. 7:8.1–8.20

The *Annual Review of Linguistics* is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-040620-061341

Keywords

Distributed Morphology, Roots, lexical category, morphology, locality

Abstract

Within Distributed Morphology, it has been proposed that the lexical vocabulary consists of Roots: category-less primitives. The motivation for Roots is connected with a line of argument reaching back to Chomsky concerning the representation of lexical categories and their role in syntax. At the center of the theory of Roots is the Two Domains Intuition: the idea that there are two different types of domains in which grammatical interactions (form: allomorphy; meaning: allosemy) occur. Roots are posited as part of an argument against lexicalist approaches to the Two Domains Intuition that reduce it to a modular distinction between the lexicon and the syntax. In place of the modular distinction, Root-based approaches hypothesize that domain differences are derivative of syntactic locality effects in a way that connects with the phase theory of Minimalist syntax. This review examines developments leading to current versions of a Roots-and-contexts theory. A particular focus is on the idea that separating lexical Roots from the morphemes that categorize them is essential to defining the distinct locality domains that are posited to explain the effects subsumed under the Two Domains Intuition.

1. INTRODUCTION

The Root is a kind of morpheme employed in Distributed Morphology, a syntactic theory of morphology and of the structure of grammar more generally. Roots are hypothesized to be category-neutral members of the lexical vocabulary. In a standard notation, Roots are primitives like $\sqrt{\text{KICK}}$, $\sqrt{\text{CAT}}$, or $\sqrt{\text{RED}}$, not "basic" verbs, nouns, or adjectives ([$_{\text{V}}$ kick], [$_{\text{N}}$ cat], [$_{\text{A}}$ red]). Why employ Roots instead of categorized primitives?

This review focuses on the question above—the motivation for Roots. Another line of review could address what properties (e.g., phonological, semantic) Roots have, or how they are represented at different stages of a derivation. These are important topics in their own right. The same may be said of the many objections that have been raised against the specific type of Root-based approach that is the focus of this review. I have chosen to examine only the motivation for Roots in this review because motivational issues are clearly more basic. It also seems to me that many discussions of Roots have lost track of the arguments for positing this kind of primitive or proceed as if Roots were a foundational assumption, not (part of) a hypothesis whose predictions continue to be refined and compared with alternatives. This review thus privileges depth over breadth, in two ways. First, the argument for Roots is found in a relatively small number of works that require unpacking. Second, there is a temporal depth to the discussion; it is necessary to consider developments in the treatment of lexical categories in earlier syntactic theories. While this historical angle is essential to understanding the (dis)continuities that define work in this area, the review is not itself a retrospective. Rather, the background discussion provides context that is essential to understanding what is at issue in current work.

Central to the discussion is the idea that there are two domains involved in the formation of complex objects. This idea is called the Two Domains Intuition:

(1) Two Domains Intuition: There are two distinct but systematic sets of interactions and properties shown by objects built in the grammar.

The interactions at issue are contextual effects on both the form and meaning sides. On the form side, they concern the conditions under which morphemes may interact to determine each other's form: allomorphy. On the meaning side, the question concerns allosemy—to a first approximation, the theory of when morphemes in a complex object may interact to produce particular types of meaning effects. Root-based versions of Distributed Morphology propose that the Two Domains Intuition is derivative of syntactic locality and implement this theory in a way that involves the separation of category-neutral Roots from morphemes containing features that define lexical categories. My starting point is thus with the treatment of lexical categories in earlier syntactic theories that provide context for this approach.

2. CATEGORY: "REMARKS ON NOMINALIZATION"

The defining property of Roots is that they are category-neutral; that is, they are part of a theory in which the features defining lexical categories like verb and noun are separated from the lexical vocabulary items that are realized in such categories. Understanding the motivation for positing such objects requires a look at the revisions to the theory of lexical category proposed in Chomsky's (1970) "Remarks on Nominalization" (hereafter, *Remarks*). The discussion in this section shows (*a*) how the analysis of nominalizations relates to this more general project in ways that involve (*b*) an argument that the Two Domains Intuition reflects the operation of distinct syntactic mechanisms, in which (*c*) lexical items unspecified for grammatical category play a crucial role.

Chomsky (1970) starts with the observation that many phenomena allow for competing analyses that involve different components of the grammar. In the framework assumed in *Remarks*, the

choices are between "enriching the lexicon and simplifying the categorial component of the base, or conversely; or. . .simplifying the base at the cost of greater complexity of transformations, or conversely" (Chomsky 1970, p. 185). The main descriptive problem approached in these terms is posed by different types of nominalizations. A starting point (not Chomsky's, but one that is taken up in later work) is shown in examples 2 and 3. The nominalizations shown in examples 2b and 3b are called derived nominals; those of the type shown in examples 2c and 3c are gerunds:

(2a) The Romans destroyed the city.
(2b) The Romans' destruction of the city…
(2c) The Romans' destroying the city…

(3a) Mary refused to leave.
(3b) Mary's refusal to leave…
(3c) Mary's refusing to leave…

To a first approximation (there are refinements below), the semantic ("thematic") relations between local NPs and the forms of *destroy* and *refuse* appear to be the same in the sentences shown in examples 2a and 3a and the derived nominals (examples 2b and 3b) and gerunds (examples 2c and 3c). At a minimum, an analysis of nominalizations must account for why they do (or do not) have the same properties as verbal clauses, and for the differences between these types of nominalizations (see below).

In early treatments of nominalizations (e.g., Lees 1960), the relationship between sentences and both derived nominals and gerunds is handled transformationally—that is, by constructing (something like) the sentences shown in examples 2a and 3a and then employing distinct transformations to produce the different kinds of nominals. The relevant differences had been identified in work prior to *Remarks*. While gerunds are productive, have the internal structure of verb phrases, and always have transparent semantics, derived nominals differ in all of these properties: They are not fully productive, they have internal properties characteristic of noun phrases, and they can display idiosyncrasies in both form and interpretation.

A guiding idea behind *Remarks* is that these systematic differences might fall out of a larger generalization, not simply the properties that certain transformations are stipulated to have. This is the crucial intuition. The mere fact that some derived nominals have formal and semantic idiosyncrasies, while gerunds do not, is not what is at issue. What is important is (*a*) analyzing the systematic properties of both types of nominals and (*b*) the ways in which they differ from each other; then, it can be determined whether there is evidence in favor of analyzing them with one grammatical mechanism versus another.

Regarding systematic properties, it is certainly the case that derived nominals can display idiosyncrasies; for example, unpredictable allomorphy is one well-known difference between derived nominals and gerunds (compare, e.g., *refus-al*, *confus-ion*, and *marri-age* with *refus-ing*, *confus-ing*, and *marry-ing*). Such nominals sometimes have unpredictable meanings as well. Nevertheless, derived nominals are systematically related to corresponding verbal clauses by default. For instance, the word *transmission* has a typical meaning, as in *Mary's transmission of the documents*. It also has a special meaning referring to a part of a car, as in *The car's transmission is damaged*. On the assumption that the two instances of *transmission* are the same, one can conclude that this word is sometimes idiosyncratic in this particular way. However, the observed idiosyncrasies are limited in their scope. *Transmission* does not have argument-taking properties that differ radically from, for example, those of other nominals; quite generally, there are no lexical items X for which *John's X-ion by Mary* denotes an event in which John agentively X's Mary. For that matter, the lexical meanings of verbs and their corresponding derived nominals are systematically related in the

normal case; words like *transmission* are singled out as idiosyncratic precisely because the default expectation is that they should behave predictably. These types of observations indicate that while some unpredictable behaviors are found with derived nominals, they nevertheless have systematic properties that call for a principled explanation; in the words of Chomsky (1970, p. 211), "the internal structure of the [derived] nominal. . .mirrors that of the sentence." The complication is then how to analyze derived nominals while accounting for the systematic ways in which they differ from gerunds.

This second systematic effect—differences between derived nominals and gerunds—can be viewed in terms of correlations. Why should being internally a noun (like a derived nominal) correlate with potential "lexical" idiosyncrasies for both allomorphy and interpretation? Or, why does a nominal with the internal structure of a verb (like a gerund) not show these potential interactions? The correlations implicated in these questions suggest that there are systematic grammatical conditions under which a particular type of idiosyncrasy—one that makes reference to the identity of a lexical item—can and cannot be found, and that the determination of lexical category ("basically a noun" versus "noun derived from verb") plays a key role in defining these conditions.

In the terms introduced above, the idea that these two types of systematic effect deserve a principled explanation is a manifestation of the Two Domains Intuition. Chomsky (1970) compares competing analyses of nominalizations with this idea in mind as a primary thread throughout *Remarks*. To understand these analyses, it is necessary to review certain assumptions about category that form the background for this work. Early work in generative grammar employs a distributional view of category. Categorial rules of the base produce elementary syntactic objects—rules like Sentence → NP + VP, NP → T + N, and VP → Verb + NP. Categorial rules also introduce words—for instance, T → *the*; N → *man*, *ball*, etc.; Verb → *hit*, *took*, etc. (Chomsky 1957, p. 26). Lexical categories are defined in terms of where they appear in expansions—in terms of their distributions with respect to other elements—and not inherently. This type of analysis produces asymmetries in the local environments of different categories. Verbs, for example, can be transitive (and must be, in the rules above) since they appear in an expansion with an NP following them. Nouns, however, do not have this property; they occur after T and with nothing following them. In this type of grammar, the symbols N and V do not have intrinsic content that distinguishes them; that is, they are not distinct in the way that, for instance, nasal and oral segments would be in terms of a feature like [± nasal]. They differ solely in terms of their distributions with respect to other symbols in the base rules.

It is important to note that categorial rules both distribute categories and introduce words into them. Thus, words are restricted in their distributions: They are realized only in a single (distributionally defined) category, and they have no existence independent of the rules that insert them. This manner of effecting lexical insertion has important consequences for cross-categorial relationships. If *refuse* is taken to be a verb, then it has to be introduced by a rule V → *refuse*. But what about *refusal*? If it is introduced by a rule that introduces nouns (e.g., N → *refusal*), it would surface in the output of this grammar fragment. However, there is then no connection between the verb *refuse* and the noun *refusal*; it is only by accident that the grammar introduces these two words that overlap substantially in their form, meaning, and certain syntactic properties. Treating these words—and by extension all related verb/derived noun pairs—as only accidentally connected does not seem desirable. In a framework that treats category as the one under discussion does, there is only one way to connect verbs and nominals systematically: by deriving the nominals from verbal clauses transformationally (see Lees 1960).

Chomsky (1965) introduces syntactic features and treats lexical categories in terms of them. The rules of the base distribute symbols like N and V as shown above. At the same time, the grammar contains a lexicon in which lexical entries are classified in terms of category features like

[±N] and [±V]. The theory thus has two places for lexical categories—they appear both in the base rules and in the lexicon—an idea that is revisited below. For now, it is important to focus on an implication of the introduction of features, which forms the main topic of *Remarks*: the idea that the distributional definition of category can in principle be dispensed with. If, for instance, nouns are defined featurally, they have the property of being nouns independently of the rules of the base that determine their distribution. So, nouns could in principle have a distribution identical to that of verbs (e.g., select complements of different types). Lexical items exist independently of the rules of the base in the lexicon and are therefore not restricted to membership in a single basic lexical category, as they are in a theory that introduces them with categorial rules. In other words, a single lexical item could in principle be inserted into different lexical categories in a syntactic structure as long as the features of the lexical item and those of the syntactic terminal match.

Against this background, Chomsky (1970) assumes on the basis of earlier work that gerunds are analyzed transformationally. The primary tension concerns the analysis of derived nominals, for which two analyses are opposed. One treats them transformationally. The other involves an extension of the (rules of the) base. The extension in question allows the base to admit rules in which nouns, like verbs, also take complements—something like NP → N comp(lement). This change allows a noun like *destruction* to take an NP complement and a noun like *refusal* to take an S complement, just like their verbal counterparts do. When generalized, NPs can also introduce a spec(ifier); when generalized in another way, both nouns and adjectives are treated like verbs in terms of this basic expansion. It is then possible to abstract over all of the lexical categories with the variable X since they have the same distribution (possibilities for immediately local elements). The result is the $\overline{X}$ schema: $\overline{\overline{X}} \longrightarrow$ [Spec, $\overline{X}$] $\overline{X}$; $\overline{X} \longrightarrow$ X. . . .

Along with the $\overline{X}$ schema, Chomsky (1970) also posits (proto-)Roots. In the particular version of extending-the-base that is explored, a lexical item like DESTROY is neither a noun nor a verb inherently (for an earlier version, see Chomsky 1965, pp. 219–20 n. 29). Instead, it becomes a noun or a verb when the phrase-structural position in which it appears contains the relevant category features (e.g., [+N, −V], [+V, −N]). Derived nominals on this analysis are basic nouns. The $\overline{X}$ schema produces an expansion that (with category features added) has a head [+N, −V] taking a complement such as *of the city*. Lexical items like DESTROY, REFUSE, and GROW can be inserted into this basic noun syntax, where they are realized as *destruction*, *refusal*, and *growth*. Crucially, derived nominals are not derived from verbs or verb phrases; they are exclusively nominal.

This analysis requires that DESTROY exist in the lexicon independently of the base rules—a position that Chomsky (1970) sometimes calls "lexicalist," meaning that it requires an independent lexicon, not that it requires derivations involving lexical rules to relate words. Historically speaking, research immediately following *Remarks* takes it only in this latter sense; I return to this point below. Here, what is important is Chomsky's conclusion that the extend-the-base analysis is superior to a transformational analysis of derived nominals. The arguments on this point have the same abstract form: (*a*) If derived nominals were derived transformationally, restrictions of various types (concerning productivity, interpretations, and so forth) would have to be imposed on transformations, but (*b*) these restrictions would have to be stipulated; that is, they do not follow from anything deeper. On the other hand, (*c*) extending the base offers the possibility for a more general, nonstipulative explanation.

Illustrating one of these stipulations makes the line of argument concrete and paves the way for the discussion of Roots below. One observation that plays an important role both in *Remarks* and in later work is that while certain derived nominals allow an agentive interpretation, others do not. For example, while *amuse* appears as a transitive verb in examples like *John amused the children*, the corresponding derived nominal is ungrammatical: **John's amusement of the children*. Or, while the derived nominal *rejection* allows for an agentive interpretation, this is not the case

with *growth* (compare *Mary's rejection of the claims* with **John's growth of tomatoes*). Following earlier proposals, it is suggested in *Remarks* that certain verbs (e.g., *amuse* or *grow* but not *reject* or *destroy*) are transitive (anachronistically, "license agentive subjects") only when they appear in a syntactic structure with an element [+cause]. Illustrating with *grow*, the transitive is [John [+cause] [$_S$ the tomatoes grow]], where [+cause] selects for a clause S. The distribution of [+cause] is restricted in the following way:

(4) Agent-Licensing Assumption (1970 version): Agents of transitivized verbs are introduced in syntactic structures with the feature [+cause]; this feature is found only with base structures that are sentences S, not those that are nouns.

The argument for why, for instance, *growth* cannot be agentive on the extend-the-base analysis is then as follows. With extended base rules, the nominal *growth* is a basic noun. Under the Agent-Licensing Assumption, the element [+cause] does not appear in the base structures for nouns. Thus, it is predicted that derived nominals for the relevant class of verbs will not allow agentive interpretations (i.e., **John's growth of tomatoes*). However, if derived nominals were derived transformationally from sentences, only a stipulation would prevent a transformation from creating *John's growth of tomatoes* from [John [+cause] [$_S$ the tomatoes grow]]. This and related lines of argument conclude that there is a principled explanation for the general set of differences between derived nominals and gerunds if the former are treated by extending the base but not if all nominals are derived transformationally from verbal clauses.

This type of argument plays a significant role in *Remarks*. However, Chomsky's (1970) primary focus is on the consequences of moving toward a theory in which categories are defined exclusively in terms of features; other questions concerning category-neutral lexical items are not addressed directly. Furthermore, while *Remarks* effectively introduces Roots, the period of research immediately following it develops a very different stance on lexical categories and the Two Domains Intuition.

The developments in question are connected to lexicalist theories of grammar. To preview this discussion and the focus of the theoretical oppositions that are outlined in the sections below, it is useful to make a further distinction. Strictly speaking, the argument developed in *Remarks* requires that lexical items like destroy and refuse be capable of appearing both as basic nouns and as basic verbs. However, this requirement does not mean that they must be represented as category-neutral. Rather, category-neutrality plays a role in one particular implementation of the idea that lexical items can surface basically in more than one category. The particular approach outlined in *Remarks* makes categorization contextual: Lexical items are categorized by being combined with the features [±N, ±V]. A development of this contextual approach is at the heart of Root-based Distributed Morphology, where the difference between basic and category-changing derivation also plays a crucial role.

As noted above, however, category-neutrality is involved in one particular way of implementing the idea that a single lexical item may occur as a basic member of more than one category; there are also other ways. An alternative approach holds that, for instance, destroy is represented as a verb [+V, −N] and then posits rules capable of changing this feature set into [−V, +N] to produce the basic noun *destruction*. One could then ask the following questions based on the Two Domains Intuition: What kinds of rules are these? And why do rules that change category features have one set of properties, while rules that add category features to something already categorized have another? Though stated somewhat abstractly, the idea that it is a difference in rule type that is responsible for the Two Domains Intuition is what is pursued in lexicalist approaches, which implement the rule difference architecturally as a lexicon versus syntax division; this division is the topic of the next section.

In summary for this section, the treatment of derived nominals in *Remarks* is syntactic but not transformational; the extension of the base analysis is implemented with category-neutral lexical entries; and in terms of implementing the Two Domains Intuition, the *Remarks* approach posits systematic differences between category-basic derivations (e.g., a lexical item as a basic noun) and category changing (e.g., a lexical item becoming a basic verb and then becoming a noun).

3. LEXICALISM, PROJECTIONISM, AND BEYOND

Despite the syntactic analysis of derived nominals that Chomsky (1970) advocates, developments immediately following that paper go in a different direction, one that posits a nonsyntactic component in which certain words are derived: a (generative) lexicon. This lexicalist movement has direct implications for category-neutral primitives, which more or less disappear from the theory. The links between *Remarks* and this type of lexicalist theory that follow it are complex. My comments here thus concentrate on two points: (*a*) why a theory with lexical rules would not need to make use of category-neutral lexical items and (*b*) how this relates to hypotheses about the Two Domains Intuition.

As noted above, while Chomsky (1970) refers in *Remarks* to the extend-the-base analysis as "lexicalist," he does not posit a generative lexicon as a place where (certain) words are formed, or lexical rules that derive words from certain other words. Chomsky (1970, p. 190) does suggest that "fairly idiosyncratic morphological rules" are needed to produce derived nominals because of the unpredictable allomorphy they display. However, he does not specify where (or when) these rules apply; it could be before or after lexical insertion. If they apply after, as *Remarks* most clearly suggests, then, for instance, DESTROY is inserted directly into a [+N, −V] syntactic head, with later rules producing *destruction* after lexical insertion. Root-based Distributed Morphology moves quite generally in the direction of having "morphology after syntax" in this way, as discussed below. Many theories in the immediate wake of *Remarks* take another path, with morphological rules applying before lexical insertion. That is, these rules apply in the lexicon, so that *destruction* would be a [+N, −V] noun created in that component of the grammar. I focus on this type of lexicalism here: one that has a modular distinction between lexicon and syntax.

Historically speaking, research following *Remarks* appears to have moved almost unanimously toward this kind of modular approach to the Two Domains Intuition, with (at least certain) words being derived in a lexicon. The actual argument of *Remarks* notwithstanding, this version of lexicalism became standard in subsequent syntactic theories.[1] The lexicon, conceived of as a component of the grammar in which words are derived and stored, became an object of study in its own right, particularly as morphology is concerned (see, e.g., Halle 1973, Jackendoff 1975, Aronoff 1976; for a helpful overview, see Carstairs-McCarthy 1992).

In a lexicalist theory of the relevant type, lexical items are inserted into syntactic structures (or project them; see below). So, for example, the categorial rules of the base might generate a VP into whose [+V, −N] head the lexical item [$_V$ destroy] can be inserted. Lexical rules would also create a noun [$_N$ destruction], which is inserted into [−V, +N] positions like the one in *destruction of the city*. The question to focus on is how *destroy* and *destruction* are related to each other. In a theory that uses the lexicon to change the categories of lexical items, it is possible to relate them in a way that does not involve transformations but that is different in turn from the extend-the-base

[1]Viewed in the light of these developments (and not the text itself), the idea that *Remarks* posits Roots and a nonmodular implementation of the Two Domains Intuition is thus a revisionist perspective (S. Anderson & A. Kroch, personal communications).

analysis. This scenario assumes the use of rules applying in the lexicon. With such rules, *destroy* and *destruction* (*a*) can be lexically related and (*b*) can surface syntactically as basic verbs (i.e., inserted directly into [+V, −N] nodes) and as basic nouns (i.e., inserted directly into [−V, +N] nodes).

This kind of lexicalist theory shares a property with the extend-the-base approach advanced in *Remarks*: It does not relate derived nominals transformationally. However, it differs from that analysis in other important respects. As discussed above, Chomsky's (1970) analysis attempts to give a principled explanation for why certain verbs that involve "causative" variants do not have agentive derived nominals. His extend-the-base analysis with category-neutral lexical items provides such an explanation with the help of the Agent-Licensing Assumption (example 4). What about a theory that has categorized lexical items and lexical rules—one that derives both *destruction* from *destroy* and *growth* from *grow*? What principle ensures that the former nominal can be agentive while the latter is not? It is not clear that the lexical theory provides a nonstipulative explanation for the relevant patterns; this point is the gist of Marantz's (1995, 1997) critique, as I show in Section 4.

Looking beyond the specific question of nominalizations, a relevant line of work under the lexicalist umbrella examines the broader question of what kinds of generalizations should fall under the Two Domains Intuition, in a number of domains. Wasow (1977) brings together several of these generalizations, in proposals that draw dividing lines in the domain of passivization and other verbal alternations (see also Anderson 1977). Lexical phonology and morphology, developed by Kiparsky (1982), looks at a different set of (primarily sound-side) phenomena with an emphasis on the internal structure of the lexicon. Other proposals, such as Zwicky & Pullum's (1983) discussion of cliticization (syntax) versus inflection (lexical), are in this vein as well. There is in fact so much work devoted to justifying or making precise a lexicon/syntax modular distinction that it would be impossible to summarize even in a much longer review.

As useful as a survey of core lexicalist positions on the Two Domains Intuition might be, the discussion here proceeds with reference to the specific proposal stated in example 5. Though drawn from a later period (Aronoff 2007), it provides a succinct background statement for the introduction of Roots:

(5) Lexicalism: What happens inside words is qualitatively distinct from what happens outside of them.

It is useful to have the statement from example 5 at hand. The discussion of lexicalism in the work of Marantz (1995, 1997), where Roots are introduced, does not cite any reference as holding the view he criticizes. At the same time, example 5 is expressly adopted by lexeme-based theories of morphology like the one discussed by Aronoff (2007, p. 804) (see also Section 5), which are continuations of the type of lexicalism developed post-*Remarks*.

One issue in identifying specific forms of lexicalism concerns how the modular distinction it posits is linked with syntactic theories. The kind of lexicon identified above begins as part of a hypothesis about the Two Domains Intuition—about words in particular. In the Government-Binding Theory (GB) of the 1980s, it assumes an additional role, one that makes it look indispensable. Early work in GB observes that the information represented in lexical items to ensure their correct distribution in the syntax—strict subcategorization features—is effectively ["largely," for Chomsky & Lasnik (1993)] identical to the information encoded in the categorial rules of the base. Because of this redundancy and other problems identified with the base component, it is argued that the base rules should be eliminated (Chomsky 1981, Stowell 1981). In the approach that moves in this direction, initial syntactic structures (i.e., D-structures) are created by satisfying the specifications of lexical items. As a result, the lexicon becomes a sine qua non for this particular theory since it is impossible to create basic syntactic objects without it. Accordingly, this type of theory is called "projectionist" because of the way that structures are projected from lexical entries [see also Chomsky's (1981) Projection Principle].

The centrality of the lexicon to GB is beyond dispute. However, two subsequent developments make available other architectural possibilities.

First, familiar versions of GB in which D-structure is projected from lexical items take it for granted that the unpredictable allomorphy shown by certain words is handled in the lexicon. However, this assumption is by no means necessary for a projectionist theory. It is possible to sever (*a*) the projection of structures determined by a lexical item's properties from (*b*) the set of processes (e.g., affixation operations, operations determining allomorphy) that produce the form of that object. This is in essence the position of Halle & Marantz (1993) (cf. Halle 1990), where a "morphological structure" is introduced on the PF side of a GB grammar to perform operations that determine the forms of morphemes "in context"—that is, after they have been combined syntactically [cf. Chomsky's (1957) "morphophonemic" rules]. While they take a syntactic view of inflectional affixation, Halle & Marantz (1993) otherwise adopt a standard GB-type architecture, complete with D-structure. This component plays almost no role in their 1993 paper, however, which is primarily concerned with the claim that inflectional morphology is piece-based and syntactic (cf. Anderson 1992).

Second, in a concurrent development, Chomsky (1993, p. 21) returns to the idea that the grammar requires "generalized transformations, or devices like them." With a mechanism for producing (elementary) syntactic objects that is independent of lexical entries (Generalized Transformations, later Merge), the theory no longer requires a lexicon to create basic structures in the way that (baseless) GB does. The return to a generalized structure-building mechanism thus makes possible the development of a nonprojectionist approach to grammar, against the prior lexicalist/projectionist view.

Although these developments show that a projectionist lexicon is not indispensable, they are not necessarily arguments against it or in favor of a particular nonprojectionist alternative. In this vein, Halle & Marantz (1993) provide a partial argument against a subpart of lexicalism in their discussion of Chomsky's (1993) "checking theory" approach to verbal morphology. However, this argument states only that syntactic affixation is required for inflectional morphology; it is not an across-the-board argument against the version of lexicalism that is embodied in example 5. This further step is precisely what is at issue when Roots (re)enter the picture.

In summary for this section, in the wake of *Remarks*, the lexicalist movement posits a lexicon/syntax modular distinction to explain effects related to the Two Domains Intuition (see example 5). The GB framework requires (after certain developments) a projectionist lexicon from which D-structures are projected, making the lexicon a necessity. Subsequent developments (*a*) reintroduce the idea that morphology might be postsyntactic and (*b*) make possible the elimination of projectionism by reintroducing generalized structure-building mechanisms.

4. ROOTS IN DISTRIBUTED MORPHOLOGY

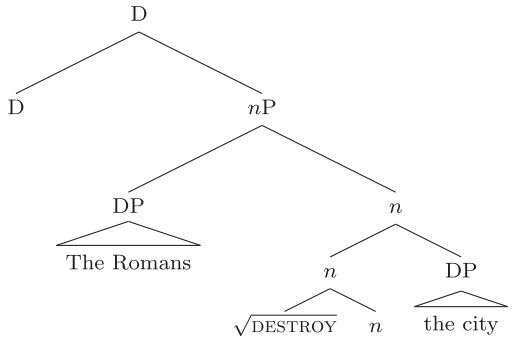
Roots make their appearance in Distributed Morphology in the work of Marantz (1995, 1997) as part of an argument against lexicalism (as defined in example 5) and in favor of a view of the Two Domains Intuition that is based instead on syntactic locality. The patterns of nominalization in examples 6 and 7 play a central role in the argument, which Marantz packages as an update to *Remarks*; while *Remarks* does not explicitly juxtapose these particular patterns, the analysis that Marantz develops for this contrast is implicit in that paper:

(6a) John destroyed the city.
(6b) *The city destroyed.
(6c) John's destruction of the city…
(6d) John's destroying the city…

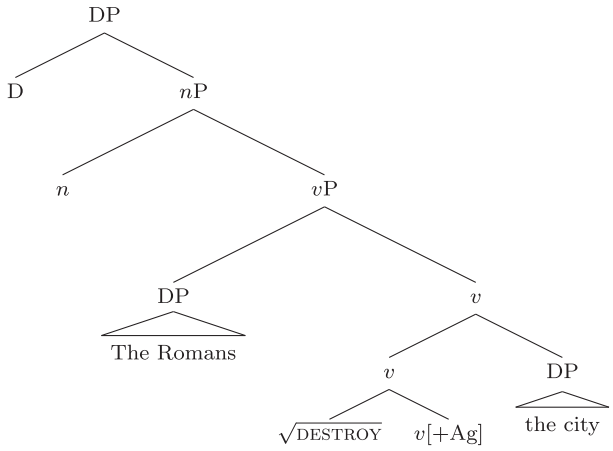
(7a) John grows tomatoes.
(7b) Tomatoes grow.
(7c) *John's growth of tomatoes…
(7d) John's growing tomatoes…

Marantz's (1995, 1997) approach shares with Chomsky's (1970) the idea that both derived nominals and gerunds are derived syntactically. However, while different syntactic tools are used for the two in *Remarks*, Marantz's analysis employs a single mechanism: generalized syntactic structure building (Merge). The differences between nominal types are derivative of locality: informally, how close the Root is to the head that nominalizes it. Derived nominals are formed when a Root is combined directly with nominalizing features—the *n* head in structure 8. Gerunds involve the Root first combining with *v* (and forming a *v*P, with perhaps other heads), with the *n* being added on top, as in structure 9. The general idea, developed in Section 5 below, is that the Root and *n* are local in structure 8 but not in structure 9 because of the presence of intervening *v*:

(8) Structure of *the Romans' destruction of the city*



(9) Structure of *the Romans' destroying the city*



These structures have been modified from those that appear in Marantz's (1995, 1997) papers. The *n* head, for example, is considered in those works but is not depicted; I have shown it here for continuity with later proposals. I have also introduced the possessor with the *n*P in structure 8; this is not crucial for immediate purposes, as long as it is clear that there is no agent-licensing head in this structure.

As in *Remarks*, the classes of Roots represented by $\sqrt{\text{GROW}}$ and $\sqrt{\text{DESTROY}}$ are assumed to be different. The $\sqrt{\text{GROW}}$ type can be interpreted with an agent (i.e., as causatives, in the *Remarks* sense) only when there is syntactic structure that licenses this interpretation. By contrast, the $\sqrt{\text{DESTROY}}$ type can be interpreted as (pseudo)-agentive even without such structure. Roots of this type are called "encyclopedically agentive" because they imply something related to agentivity in their lexical semantics. This fact reflects the flip side of the assumption in *Remarks* that causative verbs like *amuse* and *grow* can be interpreted agentively only in a particular type of syntactic structure, not inherently.

Regarding the syntactic licensing of agents, it can be assumed (in a notationally anachronistic way) that the agent relation is introduced by the feature [+Ag(ent)]. In later work, this is a feature of a head called "Voice" that selects *v* (cf. Kratzer 1996, Pylkkänen 2002). In the same way that the *Remarks* analysis restricts [+cause] to taking only an S complement (never a noun), the [+Ag] feature is restricted in its distribution; it occurs only with *v*, never with *n*. This restriction is stated as the (updated) Agent-Licensing Assumption in example 10 (cf. example 4 above):

(10) Agent-Licensing Assumption (1990s version): Agent thematic relations are introduced by the feature [+Ag], which is either a feature of *v* (and not *n*) or a feature of a head Voice that selects only *v* (and not *n*).

Derived nominals are Root (i.e., basic) nouns (recall structure 8) and lack a verbalizing *v*; they will thus never occur with syntactically licensed agents (they can, however, occur with possessors; see below). Gerunds, by contrast, are first verbalized and later nominalized. Because they contain a *v*, they can have the agent-licensing feature [+Ag].

The contrasts in examples 6 and 7 are then treated as follows. When $\sqrt{\text{DESTROY}}$ appears in the structure outlined in structure 8, it is categorized as a noun. There is no agent-licensing head. However, because this Root is encyclopedically agentive, it allows possessors to be interpreted agentively (through the back door, as it were). Whatever this backdoor mechanism is, it serves to explain why agentivity is a possible (but not necessary) interpretation for the possessor in certain derived nominals. For example, *Mary's destruction of the city* could be the destruction that Mary attended, not necessarily the one that she perpetrated. The flexibility of these alternate nonagentive interpretations is easier to see with some Roots than with others (e.g., *Bill's performance of the play*, on the interpretation "the play that Bill attended"). With $\sqrt{\text{GROW}}$-type Roots, there is no possibility of agentive interpretation through this nonsyntactic mechanism. In the derived nominal *growth*, agent-licensing structure is excluded because it does not occur with nouns, as stated by the Agent-Licensing Assumption in example 10. *Growth* is thus predicted correctly to disallow an agentive interpretation (cf. example 7c). However, the gerund *growing* is predicted to allow agentive interpretations because it is a nominal formed on a *v*P, where agent-licensing structure is possible (cf. example 7d).

Much could said about this argument—from background assumptions to details of implementation to the particular data patterns involved. For the purposes of this review, the issue is how category-neutral Roots fit into the picture. In the *Remarks* analysis, lexical items have to be unspecified for categories like "noun" and "verb" so that they can be inserted directly into X^0 positions bearing different combinations of category features—that is, so that they can potentially be basic nouns, verbs, and so on. This assumption allows for the systematic relations that elements show in different categories to be accounted for, but without transformations, which, Chomsky (1970) argues in *Remarks*, can be restricted to produce the correct results only by stipulation. Marantz's (1995, 1997) Roots have essentially the same properties and motivation. It is crucial for the argument, which parallels Chomsky's, that derived nominals be basic nouns, not deverbal nominalizations of verbs. If derived nominals were formed from verbs, there would

be no principled way of accounting for the asymmetries in argument licensing seen in examples 6 and 7.

The more general conclusion drawn by Marantz (1995, 1997) is that lexicalism as defined in example 5 precludes a principled explanation of the behavior of nominalizations and, by extension, the Two Domains Intuition. The reasoning proceeds along the same lines as Chomsky's argument against treating derived nominals transformationally. In *Remarks*, Chomsky (1970) suggests that if the sentence *John grows tomatoes* were transformationally related to a nominal with *growth*, nothing short of a stipulation could ensure the ungrammaticality of **John's growth of tomatoes*. For Marantz, the argument is the same but framed with reference to the lexicon, not a transformational rule: If the transitive *grow* that appears in *John grows tomatoes* exists in the lexicon, why can't it be nominalized like the transitive *destroy* can be? Some stipulation—Marantz (1997, p. 220) has it as "don't make nominalizations from verbs that are causatives of change of state verbs with internal causers"—is required to prevent this from happening. Marantz argues that the Root-based alternative provides a nonstipulative explanation for this asymmetry. In terms of commitments, this alternative assumes that $\sqrt{\text{GROW}}$- and $\sqrt{\text{DESTROY}}$-type primitives are inherently different, as most approaches do; it also relies crucially on the Agent-Licensing Assumption (see Appendix).

The idea that Roots are differentiated along these lines deserves a brief aside. In particular, the idea that Roots are category-free does not mean that they must be entirely free of inherent properties. In the *Remarks* treatment, for example, items like REFUSE are represented "with certain fixed selectional and strict subcategorization features [but] free with respect to. . .categorial features" (Chomsky 1970, p. 190). The Root differences suggested by the Agent-Licensing Assumption implicate another dimension to Root-meaning space. An active and ongoing line of research has investigated the idea that Root semantics involve an ontology that connects with type-theoretic and event-structural notions such as states, events, and entities (see, e.g., Harley 2005).

Returning to the analysis of nominalizations, a lexicalist account could in principle say that nominalizations like *destruction* have thematic possessors, not true agents, just like the syntactic approach sketched above says. It should therefore be possible to have a more general lexicalist account than the one that is the target of the quote from Marantz (1997, p. 220) above—one that says, for instance, that the affixes that produce derived nominals never attach to verbs that introduce an agent thematic relation (or a [+cause] feature). The affixes that create gerunds, by contrast, would be specified so that they would be able to attach to verbs that have this property. This works mechanically. It can also be implemented syntactically with categorized primitives. A nominalization like *the growth of tomatoes* could be represented as follows: [N_1 [$_{VP}$[+V, grow] [$_{PP}$ of tomatoes]]]. To prevent the derivation of **John's growth of tomatoes*, the N_1 head would be specified so that it does not select for Voice/[+Ag]. Parallel to the lexical account, a second nominal head N_2 that can select for [+Ag] would be posited for gerunds. This analysis, though based on selection, shares essential properties with that of Lees (1960), which employs distinct transformations for different types of nominals.

Both the lexicalist and selection-based syntactic accounts produce the relevant patterns. The question then is, why do we find these patterns and not others? Could we also say the opposite of what these accounts require: that derived nominal affixes select only verbs that have [+Ag]? This point can be generalized. The kind of stipulation involved in these accounts provides no basis for explaining the correlations between sets of properties that involve syntax, form, and meaning that have been identified in the discussion of *Remarks* above. In particular, it does not follow from such an account that the absence of agent-licensing structure/[+cause] correlates with the possible presence of unpredictable Root-determined allomorphy for the nominalization affix or with special meanings for some Root-affix combinations; or, conversely, that the presence of agentivity/[+cause] correlates with the absence of Root-determined allomorphy and with the

fully transparent meanings exhibited by gerunds. In other words, the effects of the Two Domains Intuition are encoded in such an account but not explained.

On the basis of these considerations, Marantz (1995, 1997) draws general conclusions about lexicalism, and about its performance in explaining the Two Domains Intuition, along the following lines. The point of a two-domain approach is that there are systematic differences in the way that certain form–meaning connections are produced, and the grammar must account for this fact. The lexicalist intuition in example 5 is that the word constitutes the relevant domain for one type of form–meaning connection; the syntax creates another type. Marantz's (1995, 1997) papers support the idea that finding two domains is the essential question. However, according to the analysis of nominalizations outlined above, the word cannot do the work it is supposed to do in defining a domain. Why? Because derived nominals show a special effect relative to sentences and gerunds: They preclude agentive interpretations. Crucially, both derived nominals like *growth* and gerunds like *growing* are words. The "special interaction or not" distinction thus does not match up with word/phrase distinction. This is Marantz's point generalized: There are domains that show different types of potential interactions, both for form and for meaning, but the word does not define them. The relevant domains are sometimes smaller than the word (as in the gerund example) and sometimes larger (see Section 5).

This line of reasoning allows for an analysis that goes well beyond the specific case of derived nominals versus gerunds. In particular, it invites a broader investigation of what types of generalizations should be associated with the Two Domains Intuition. Marantz's (1997, p. 223) approach concentrates attention on "the domains for contextual allomorphy and contextual allosemy (special meanings in particular contexts)." The restriction to certain phenomena is worth noting; the Root-based theory does not require that every contrast attributed to the lexicon/syntax modular split must be explained in the same way. Rather, the focus is on particular types of contextual effects, as shown in the following section.

In summary for this section, Marantz's (1995, 1997) approach employs category-neutral Roots and argues for them (and against a lexical analysis) in a way that is parallel to the argument of *Remarks*. However, whereas *Remarks* employs distinct syntactic mechanisms for derived nominals and gerunds, Marantz argues that all derivation is syntactic structure building. The differences between nominal types can be understood in terms of a theory of syntactic locality—one that crucially implicates category-less Roots as part of how locality domains are defined.

5. DEVELOPMENT OF A ROOTS-AND-CONTEXTS THEORY

The goal of a generalized theory of Roots is to explain the effects associated with the Two Domains Intuition in terms of the locality conditions under which morphemes may interact. Because interactions after syntactic composition are essential to this approach, I refer to it as the Roots-and-contexts view. Connecting with Chomsky's (1970) emphasis on the status of lexical categories in the grammar, a central role in this type of theory is played by category-free Roots and categorizing heads such as *v* and *n*. The domains associated with the Two Domains Intuition are sometimes referred to as an inner domain (where close interactions with a Root are possible) and an outer domain (where they are not). Early suggestions concerning this type of approach appear in Marantz's (1995) discussion of idioms and in Embick's (1996) analysis of allomorphic interactions.

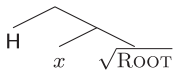
A next step makes connections with other parts of syntactic theory. The two domains that are hypothesized to exist are connected to phase theory (Chomsky 2000, 2001) in Marantz's (2001) lengthy handout, which takes further steps toward specifying what types of generalizations a Root-based theory is meant to account for (see also Marantz 2007). The intuition to be implemented—using the nominal structures 8 and 9 for illustration—is that in derived nominals, the Root and the

nominalizer *n* are in a local relationship and are therefore visible to each other for special form–meaning interactions. In gerunds, however, *n* and the Root are not local to one another because of the presence of *v*, which categorizes the Root first. The connection to phase theory provides a way of implementing these locality effects: Phase-based derivations make spell-out to the form and meaning interfaces PF and LF incremental. A consequence of treating derivations in this way is that when a syntactic object O is spelled out, it contains a number of terminals (i.e., morphemes) that are active in that cycle of form or meaning interpretation. The morphemes (including Roots) that are active in the same PF and LF cycle are those that are capable of interacting with each other in a close way. This type of phase theory also specifies the conditions under which elements are inactive with respect to certain operations (the Phase Impenetrability Condition; see Chomsky 2001). When syntactic object O has been spelled out, it has subparts that cannot be accessed during subsequent syntactic derivation and that are not active at PF or LF in the way described above. In other words, certain morphemes will be phase-inactive for others and thus will not display the close type of interactions exhibited by phase-local morphemes.

Marantz's (2001) version of the phase-based approach introduces the important assumption that category-defining morphemes like *n* and *v* are phase-defining. These heads are called "cyclic" because of the special role they play in determining cyclic spell-out. Other heads are referred to as "noncyclic." This view recognizes two types of interactions between a head H and the Root, schematized in structures 11 and 12, where *x* is a phase-defining head; these scenarios correspond to the inner and outer domains mentioned above:

(11) Inner attachment

(12) Outer attachment



The head H and the Root are in the same phase cycle in structure 11; they may potentially interact. In structure 12, however, the Root is spelled out with *x* and is inactive with respect to H. In this scenario, no interaction is predicted in which H is sensitive to the identity of the Root, or vice versa (though H could be sensitive to properties of *x*).

Marantz's (2001) discussion looks at several phenomena in terms of structures 11 and 12—for example, the difference between eventive and stative (or "verbal" versus "adjectival" passives), a privileged example in the lexicalist literature because of its prominence in the work of Wasow (1977) and related work on passivization (e.g., Levin & Rappaport 1986). While a survey is not possible here, it is worth noting that subsequent proposals have moved toward the idea that not all of these cases involve a simple phase-cyclic difference [see, e.g., Embick (2003, 2004) on passive types and phases].

In one version of the Roots-and-contexts approach, Embick & Marantz (2008) focus on employing the phase theory to explain the patterns of allomorphy and interaction with Root semantics (allosemy) that are summarized in example 13:

Generalizations

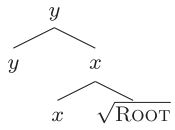
(13a) *Allomorphy:* For Root-attached x, special allomorphy for x may be determined by properties of the Root. A head x in the outer domain is not in a local relationship with the Root and thus cannot have its allomorphy determined by the Root.

(13b) *Interpretation:* The combination of Root-attached x and the Root might yield a special interpretation. When attached in the outer domain, the x heads yield predictable interpretations.
(Embick & Marantz 2008, p. 11)

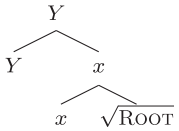
Embick & Marantz (2008) also make use of the Categorization Assumption: the idea that Roots never appear bare (i.e., without being locally categorized). This assumption implies that Roots, though category-neutral, are always interpreted in the context of a category-defining head, both for form and for meaning. There is thus no direct access to Roots; they are interpreted on both the form and meaning sides as a result of passing through a syntactic derivation.

An important question that is left vague in Embick & Marantz's (2008) paper is what types of morphemes x is intended to range over in example 13. Clearly, it cannot be all morphemes. For example, tense and other heads (e.g., aspect, number) show Root-determined allomorphy in any number of languages despite being like H in structure 12. The generalization that appears to be at play is that noncyclic heads in the outer configuration schematized in structure 12 may potentially interact with the Root, while outer cyclic heads may not. That is, a phase head *y* (and any heads above it) cannot interact with the Root in structure 14 (where *x* is also phase-defining), but a nonphase head *Y* can potentially interact with the Root in structure 15 (Embick 2010):[2]

(14) *y* not local to Root



(15) *Y* local to Root



Structure 14 involves category changing: The Root is first categorized by the head *x*, and the category-defining head *y* is added above the *x* head. This kind of structure is found with gerunds. In structure 15, the Root is categorized by *x*, and the non-category-defining head *Y* appears outside of the *x* head. This type of structure is at play when a verb has a tense or aspect morpheme attached to it or when a noun is pluralized. In such situations, the theory predicts that *Y* is phase-local to the Root, so that the two are in principle capable of interacting even though *Y* is outside of the *x* head. Thus, the theory can account for the fact that (as noted above) allomorphic interactions between Roots and several different noncyclic morphemes are quite common cross-linguistically—for instance, with irregular verbs like *ben-t* and *brok-en*, where the *-t* and *-en* exponents realize tense and aspectual heads in a configuration like the one shown in structure 15.

Moving to interpretation, Marantz (2013) explores contextual interactions with Root meanings—that is, (contextual) allosemy—in terms of the theory of locality outlined above. His conclusions parallel those described for allomorphy. An outer cyclic morpheme *y* in a category-changing scenario like that shown in structure 14 can operate only on Root meanings that are

[2]The summary illustrated in structures 14 and 15 concerns the effects that are produced, not the details of how to produce them derivationally; for one proposal, see Embick (2010).

activated by *x*; that is, outer *y* cannot access Root meanings directly. In the scenario schematized in structure 15, examples from Japanese and Greek—both involving participial morphemes that are argued to be noncyclic—are adduced as evidence that noncyclic morphemes like *Y* can directly interact with Root meaning. The details of this argument implicate some additional assumptions. Among these, the most important concerns the scope of what is to be explained. Marantz's approach requires a clear difference between, on the one hand, allosemy as discussed above and, on the other hand, idiomaticity of the type found with, for instance, *kick the bucket*. This distinction is of crucial importance to the Roots-and-contexts approach because words (and phrases) may in principle show both allosemy effects and have idiomatic uses, and the theory of locality outlined above is aimed at the former phenomenon, not the latter. An important task for further research is thus to identify the dividing line(s) between these two types of noncompositional meaning.

These developments on the meaning side of the Roots-and-contexts theory are significant for two reasons: First, they provide a stance on the Two Domains Intuition that makes possible comparison with certain alternatives, and second, they manifest a specific view of how the domains for form and meaning should relate to another. The first of these reasons connects with a general set of questions concerning the derivation/inflection split in morphological theory. While category-defining morphemes are central to the Roots-and-contexts approach, it is not predicted that all derivational morphemes should behave one way with respect to Root interactions while all nonderivational morphemes should behave in another way. As reviewed above, both cyclic and noncyclic heads may interact closely with Roots for the purposes of allomorphy or allosemy, as long as they are local to the Root. Thus, while locality domains are defined in terms of cyclic heads, possible interactions within such domains do not refer to the cyclic/noncyclic morpheme distinction per se; what matters is that conditions of locality are satisfied.

This view contrasts with an alternative one in lexeme-based morphology. According to Aronoff's (2007) representative version of this view, lexemes are essentially words minus inflectional morphology: Only Roots, derivational morphology, and compounding count in defining a lexeme's meaning. The theory thus privileges the word in a way stated in example 5 above, with "lexemes" standing in for "words." Taken at face value, this form of lexicalism holds that what matters for the Two Domains Intuition is only the type of morphology involved: Derivational morphology and Roots can interact directly for meaning in ways that are bounded only by the domain "word," whereas inflectional morphology never interacts in this way. For purposes of comparison, the primary focus has been on the meaning side, where Aronoff (2007, p. 805) argues that because morphemes are not "reliable Saussurean signs," the locus for basic form–meaning connections must be the entire lexeme. As discussed above, the primary claim of the Roots-and-contexts theory is that there are derivational (category-defining) morphemes that may not show close interactions with the Root (as in structure 14) as well as nonderivational morphemes that can (as in structure 15). Thus, it is not clear that the lexeme-based approach could explain the types of generalizations discussed above. It could, of course, deny that there are such generalizations to be explained in the first place; this appears to be the line that Aronoff adopts. A key point of contention is one highlighted above: the dividing line between allosemy and idiomaticity. Although it is not possible to go into the details here, many of the examples that Aronoff adduces in support of the claim that basic meaning relations are properties of the word appear to be idioms, not instances of allosemy.

While the comparison with lexeme-based theories takes a narrow focus within morphological theory, the developments on the meaning side point to much broader issues as well. These concern the proposal that a single type of phase-defined domain defines interactions for both form and meaning; that is, the cyclic operation of syntax creates a domain whose properties regulate possible interactions for allomorphy and allosemy. This idea, which can be called the Single Cycle

Hypothesis, embodies the intuition that form–meaning interactions are restricted in the same way because they are determined by the properties of syntactic derivations—a particular take on the Minimalist impetus in syntactic theory.

An important corollary to this syntactic approach is the idea that the interfaces might enforce locality conditions of their own. Phase domains make certain interactions possible by identifying which morphemes are active at the same time. However, phase-determined locality appears to be only part of what is at issue. A further role is played by the properties of the syntactic structures in which morphemes appear, along with relations particular to the PF and LF interfaces. Proposals have been made concerning further restrictions specific to each of the interface levels, effectively interface-parochial conditions under which morphemes may see each other. Embick's (2010) approach, for example, proposes that morphemes may interact for allomorphy when they are (*a*) active in the same phase cycle and (*b*) concatenated (immediately linearly adjacent). On the meaning side, Marantz (2013) makes a proposal that parallels this approach, holding that morphemes in a phase cycle are local for allosemy when a further condition of "semantic adjacency" is met. The details of these particular proposals, which continue to be discussed, are not at issue for immediate purposes; what is important is that both phase cycles and interface-specific conditions are at play in the grammar of morphemes in context.

In summary, the Root-and-contexts theory employs Roots, category-defining heads, and cyclic spell-out to define the domains for allomorphy and allosemy. This theory does not claim that every difference that has been attributed to a lexicon/syntax modular split is to be accounted for in these terms; it is directed at these types of contextual effects. Category-defining heads, which are (often) typical derivational morphemes, play a crucial role by defining cyclic domains. However, the theory goes beyond typical versions of the derivation/inflection divide. Both cyclic and noncyclic heads can display inner or outer behavior; locality determines possible interactions. The Single Cycle Hypothesis holds that one cycle defines local domains for both form and meaning. Interface-specific locality conditions of different types interact with phase theory to produce observed patterns. Within Root-based approaches, several questions continue to be actively investigated. These questions concern which heads are cyclic (and which are not) and whether there are further conditions on which morphemes may produce allosemy. Another active discussion concerns the theory of locality, both in terms of phase theory and in terms of additional locality conditions in syntax and at the interfaces (for a review, see Embick 2020).

6. DISCUSSION

The introduction of Roots into Distributed Morphology follows closely on a sequence of developments in syntactic theory concerning the notion of lexical category and the relation of lexical categories to the structures in which they appear. At the heart of this line of inquiry is the question of how lexical categories like "verb" and "noun" are defined syntactically and how they relate to lexical primitives like those seen in *cat*, *kick*, and *red*. Many of the empirical problems in this area involve the idea that category-flexible behavior—the ability of primitives to surface in more than one lexical category—is not of one type. Instead, there are systematic differences between types of category flexibility that call for a principled explanation; this is an instantiation of the Two Domains Intuition. The line of thought that runs from the work of Chomsky (1970) through that of Marantz (1995, 1997) is based on the idea that lexical primitives must be separated from the features that define lexical categories, resulting in category-neutral Roots. Roots—and the contexts they appear in—provide the basis in later work for broader explanations of Two Domains Intuition effects, which move beyond derivational morphology in the narrow sense to provide a general theory of the contextual interpretation of morphemes on both the form and meaning sides.

This review has concentrated on the development of a Root-based theory within one specific approach, Distributed Morphology; in concluding, some comments are in order concerning how Roots fit in with other parts of that theory. On the one hand, Roots are not, strictly speaking, a necessary part of this approach. Theories with similar assumptions could be developed without Roots, and category-free primitives could be used in any number of approaches with different starting assumptions. On the other hand, Roots provide a way of making precise the theory's core claims. The defining properties of Distributed Morphology are that it is piece-based and syntactic, with at least some contextual interpretation of morphemes at the interfaces. The Roots-and-contexts approach developed to explain the Two Domains Intuition involves all of these components in formulating claims about form and meaning: The grammar generates hierarchically structured objects that consist of morphemes, whose interpretation for both form and meaning is affected by morphemes in their local context. Morphological relationships, such as the locality conditions that play a central role in this review, are defined by the syntax or by conditions at the interfaces that are themselves defined by syntactic relations. At the heart of this project, the strong claim embodied in the Single Cycle Hypothesis holds that form–meaning relations are not only constrained by syntactic domains but also restricted in fundamentally the same way because of how syntactic derivations operate. Roots—more precisely, the separation of cyclic category-defining heads from Roots—play a central role in defining the locality domains that determine possible interactions. Thus, while Roots are not, strictly speaking, a necessary part of the theory (i.e., its architecture and other foundational assumptions do not require them), they are essentially indispensable to it in its current state.

APPENDIX: THE STATUS OF THE AGENT-LICENSING ASSUMPTION

The Agent-Licensing Assumption (ALA) plays a central role in the arguments that end in positing category-free primitives. It is therefore worth asking why the agent-licensing feature might be restricted in its distribution. That is, does it follow from anything that [+Ag] combines only with *v* (or, that Voice[+Ag] selects only *v*)? One reason to dwell on the ALA involves a tension that it implicates. The majority of *Remarks* (Chomsky 1970) considers the consequences of abandoning the distributional definition of grammatical categories. That is, if categories are defined featurally (as opposed to distributionally), then the expectation is that they should (by default) distribute in the same way in terms of immediate constituent structure, as encoded in the $\overline{\text{X}}$-schema. Thus, restricting [+Ag] to the category *v* merits further attention. It makes verbs distributionally different as the only category that can be in a local relationship with agent-licensing features and hence agents. Thus, while the ALA plays a crucial role in launching the featural approach to category, it appears in itself to assume a distributional difference between verbs and other categories.

Although not directed at this particular question, the approach to external arguments developed by Wood & Marantz (2017) provides a way of resolving this tension. They propose that external arguments are introduced by a head i^* that has the ability to introduce a DP in its specifier. The same i^* is argued to select for different complements (*v*'s, prepositions, nouns). The thematic interpretation of the external argument, which is known to differ in these different structures, is attributed to a contextual effect: a kind of allosemy that applies to functional heads. For example, the external argument of i^* that selects *v* might be interpreted as an agent, whereas the external argument of i^* attaching to other complements would be interpreted differently. In this system, the particular thematic relation that is assigned is determined contextually in a way that is divorced from the (syntactic) function of introducing an external argument. If generalized, then all categories may have the same distribution, in the sense of potentially co-occurring with i^*, even if the particular thematic role assigned to that external argument differs. In this view, all categories are

essentially the same in terms of how they relate to arguments in their local domain. In particular, all categories can occur with external arguments, but because of contextual effects on i^*, categories might differ in terms of the specific thematic relations that are associated with these arguments. In addition to its interesting consequences for how the notion of thematic relation should be understood in syntax versus semantics, this view thus has the potential to explain (away) the apparent category-specificity of the verbs-only approach to agentivity embodied in the ALA.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The author thanks S.R. Anderson, J. Benz, A. Biggs, M. Salzmann, and the anonymous reviewer.

LITERATURE CITED

Anderson S. 1977. Comments on the paper by Wasow. In *Formal Syntax*, ed. PW Culicover, T Wasow, A Akmajian, pp. 361–77. New York: Academic
Anderson S. 1992. *Amorphous Morphology*. Cambridge, UK: Cambridge Univ. Press
Aronoff M. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press
Aronoff M. 2007. In the beginning was the word. *Language* 83:803–30
Carstairs-McCarthy A. 1992. *Current Morphology*. London: Routledge
Chomsky N. 1957. *Syntactic Structures*. The Hague, Neth.: Mouton
Chomsky N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press
Chomsky N. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, ed. R Jacobs, P Rosenbaum, pp. 184–221. Washington, DC: Georgetown Univ. Press
Chomsky N. 1981. *Lectures on Government and Binding*. Dordrecht, Neth.: Foris
Chomsky N. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. K Hale, S Keyser, pp. 1–52. Cambridge, MA: MIT Press
Chomsky N. 2000. Minimalist inquiries: the framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. R Martin, D Michaels, J Uriagereka, pp. 89–156. Cambridge, MA: MIT Press
Chomsky N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. M Kenstowicz, pp. 1–52. Cambridge, MA: MIT Press
Chomsky N, Lasnik H. 1993. Principles and parameters theory. In *Handbook of Syntax*, ed. J Jacobs, A von Stechow, W Sternefeld, T Vennemann, pp. 13–127. Berlin: Walter de Gruyter
Embick D. 1996. Causativization in Hupa. In *Proceedings of the Berkeley Linguistics Society 22*, ed. J Johnson, M Juge, J Moxley, pp. 83–94. Berkeley, CA: Berkeley Linguist. Soc.
Embick D. 2003. Locality, listedness, and morphological identity. *Stud. Linguist.* 57:143–69
Embick D. 2004. On the structure of resultative participles in English. *Linguist. Inq.* 35:355–92
Embick D. 2010. *Localism Versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press
Embick D. 2020. Locality (in Distributed Morphology). In *The Cambridge Handbook of Distributed Morphology*. Cambridge, UK: Cambridge Univ. Press. Forthcoming
Embick D, Marantz A. 2008. Architecture and blocking. *Linguist. Inq.* 39(1):1–53
Halle M. 1973. Prolegomena to a theory of word formation. *Linguist. Inq.* 4(1):3–16
Halle M. 1990. An approach to morphology. In *Proceedings of the 20th Annual Meeting of the North East Linguistic Society* (*NELS* 20), pp. 150–84. Amherst, MA: Grad. Linguist. Stud. Assoc.
Halle M, Marantz A. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. K Hale, S Keyser, pp. 111–76. Cambridge, MA: MIT Press

Harley H. 2005. How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in English. In *The Syntax of Aspect*, ed. N Erteschik-Shir, TR Rappoport, pp. 42–64. Oxford, UK: Oxford Univ. Press

Jackendoff R. 1975. Redundancy rules and the lexicon. *Language* 51:639–71

Kiparsky P. 1982. Lexical morphology and phonology. In *Linguistics in the Morning Calm: Selected Essays from SICOL-1981*, pp. 3–91. Seoul, South Korea: Hanshin

Kratzer A. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, ed. J Rooryck, L Zaring, pp. 109–37. Dordrecht, Neth.: Kluwer

Lees RB. 1960. *The Grammar of English Nominalizations*. Bloomington, IN: Int. J. Am. Linguist.

Levin B, Rappaport M. 1986. The formation of adjectival passives. *Linguist. Inq.* 17:623–61

Marantz A. 1995. *'Cat' as a phrasal idiom: consequences of late insertion in Distributed Morphology*. Work. Pap., MIT, Cambridge, MA

Marantz A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *Univ. Pa. Work. Pap. Linguist.* 4(2):14

Marantz A. 2001. *Words*. Course Handout, MIT, Cambridge, MA. **https://babel.ucsc.edu/~hank/mrg.readings/Marantz_words.pdf**

Marantz A. 2007. Phases and words. In *Phases in the Theory of Grammar*, ed. SH Choe, pp. 199–222. Seoul, South Korea: Dong-In

Marantz A. 2013. Locality domains for contextual allomorphy across the interfaces. In *Distributed Morphology Today: Morphemes for Morris Halle*, ed. O Matushansky, A Marantz, pp. 95–116. Cambridge, MA: MIT Press

Pylkkänen M. 2002. *Introducing arguments*. PhD Thesis, MIT, Cambridge, MA

Stowell T. 1981. *Origins of phrase structure*. PhD Thesis, MIT, Cambridge, MA

Wasow T. 1977. Transformations and the lexicon. In *Formal Syntax*, ed. P Culicover, T Wasow, A Akmajian, pp. 327–60. New York: Academic

Wood J, Marantz A. 2017. The interpretation of external arguments. In *The Verbal Domain*, ed. R D'Alessandro, I Franco, Á Gallego, pp. 255–78. Oxford, UK: Oxford Univ. Press

Zwicky A, Pullum G. 1983. Cliticization vs. inflection: English *n't*. *Language* 59:502–13