Heidi Harley

On the identity of roots

Abstract: This paper attempts to articulate the essential nature of the notion 'root' in the morphosyntax. Adopting a realizational (Late Insertion) view of the morphosyntactic model, the question of whether roots are phonologically individuated, semantically individuated, or not individuated at all in the syntactic component are addressed in turn. It is argued that roots cannot be phonologically identified, since there are suppletive roots, and they cannot be semantically identified, since there are roots with highly variable semantic content, analogous to 'semantic suppletion'. And yet, they must be individuated in the syntax, since without such individuation, suppletive competition would be impossible. Roots must therefore be individuated purely abstractly, as independent indices on the √ node in the syntactic computation that serves as the linkage between a particular set of spell-out instructions and a particular set of interpretive instructions. It is further argued that the syntactic √ node behaves in a syntactically unexceptional way, merging with complement phrases and projecting a √P. The correct formulation of locality restrictions on idiosyncratic phonological and semantic interpretations are also discussed.

Keywords: Distributed Morphology, Hiaki (Yaqui), idioms, Elsewhere Condition, competition, allosemy, allomorphy, one-replacement, unaccusatives

DOI 10.1515/tl-2014-0010

1 Introduction

Lexical items are typically built around a core element, identifiable by linguists, though not always by speakers, as a *root*. Factors that a linguist might take into account in identifying occurrences of a root across different contexts include identity or similarity of form, identity or similarity or meaning, and purely morphological behaviors, such as idiosyncratic selectional restrictions with respect to affixation or other morphological processes. For example, a Semiticist faced with the semantically highly variable but phonogically consistent consonantal root *b.x.n*, which might be glossed 'related to examining', might conclude that it

Heidi Harley: Department of Linguistics, University of Arizona, USA.
E-mail: hharley@email.arizona.edu

is the phonological form – the particlar consonants in a particular sequence – which crucially individuates the formative: the root is √*bxn*, with different interpretations in different morphosyntactic contexts. In contrast, a Uto-Aztecanist, faced with a semantically invariant but formally suppletive verb such as *mea* ~ *sua*, 'kill (singular object) ~ kill (plural objects)', might conclude that it is the meaning – the abstract concept of 'killing' – which identifies the formative: the root is √KILL, with different phonological realizations in different morphosyntactic contexts. This paper investigates whether a unified theory of roots can be constructed which allows a motivated approach to root identity at both extremes.

Although the term 'root' traditionally designated a descriptive morphological category, in Distributed Morphology (as in many morphological theories), the term names a particular theoretical construct which plays an important role in the framework. Here, some empirical evidence is brought to bear which illuminates the nature of roots in this model, and which has implications for other models that make use of a similar construct. It is argued that neither phonological properties nor semantic properties are sufficient to individuate root nodes in the syntax. In consequence, a purely formal notion of root identity is needed for use in syntactic computation, to which phonological and/or semantic properties can be attached at the relevant point, both potentially contingent upon particular morphosyntactic contexts.[1]

The conclusion, then, is that syntactic roots are individuated as pure units of structural computation, lacking (in the syntax) both semantic content and phonological features. Following Pfau (2000, 2009) and Acquaviva (2008), an index notation is adopted, according to which individual syntactic roots are referred to simply by a numerical address. The idea is that the address serves as the linkage between a set of instructions for phonological realization in context and a set of instructions for semantic interpretation in context.

Having established this framework, a further pair of questions can then be asked: First, how do root nodes behave in the syntactic component, and second, what kinds of conditions are imposed on their semantic and/or phonological interpretation at the interfaces? In the second half of this paper, arguments are given that roots can and do take complements and project, and again, the empir-

1 Although the particular conclusions argued for here, taken individually, are for the most part uncontroversial outside the Distributed Morphology framework, the empirical results presented in support of them are relatively novel and should be of interest to investigators working from a broad range of perspectives. Furthermore, the overarching moral drawn from the conjunction of the empirical results – that root individuation is neither phonological nor semantic – is a purely general one, relevant to any model of morphosyntax, even though implemented here using Distributed Morphology technology.

ical basis for the argument draws on both semantic and morphophonological data, as well as syntactic evidence. This discussion is tightly connected to the second question, concerning constraints on the semantic and phonological interpretation of root nodes. It is clear that different morphosyntactic environments can trigger both special meanings and special pronunciations of roots. Some proposals (Marantz 2001, 2008; Arad 2003, 2005) argue for a very stringent locality condition on root interpretations. With (Borer 2009), I argue that the constraints cannot be quite so restrictive, and argue for a return to the view of the relevant locality domain originally advanced in Marantz (1995b, 1997), according to which the projection which hosts the external argument marks the domain edge.

The paper is laid out as follows. In section 2, the relevant aspects of the Distributed Morphology model are reviewed, and its original concept of an unindividuated acategorial root node is introduced. In section 2.1 arguments are presented which point to the conclusion that roots are in fact individuated in the narrow syntax. Further consideration shows that the basis for this individuation is neither phonological (section 2.2) nor semantic (section 2.3). The consequences of this discussion are spelled out in section 2.4, where an overview of root individuation, phonological realization, and interpretation is provided. In section 3, arguments are provided in favor of treating root nodes as conventional syntactic entities, capable of taking complements and heading phrasal constituents. The first such argument, in section 3.1, is syntactic, based on the analysis of *one*-replacement in English from Harley (2005b). The second, in section 3.2, is based on the conclusions of Kratzer (1994, 1996) concerning the differential constraints on idiomatic interpretations of verbs with respect to external and internal arguments. The last, in section 3.3, relies again on the suppletive root phenomena discussed in section 2.1, showing that the conditioning environment for suppletive root insertion in Hiaki is maximally local (Haugen et al. 2009, Harley et al. to appear; Bobaljik and Harley to appear). Finally, in section 4, the correct characterization of the locality conditions on idiosyncratic root interpretations is discussed. Section 5 concludes.

2 Root individuation in Distributed Morphology

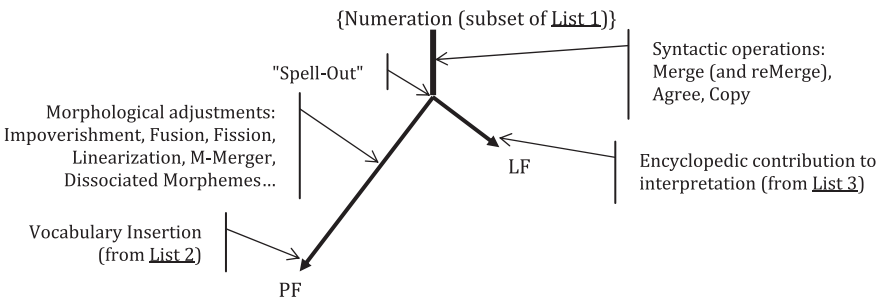
Distributed Morphology (Halle & Marantz 1993) provides a unifed framework within which both morphosyntactic and morphophonological phenomena can be modelled, and which integrates with the core Y-model of Chomskyan generative linguistics in a straightforward way. Analyses couched within the model have ramifications and make predictions concerning phenomena far from the

traditional bailiwick of morphologists, particularly with respect to the LF branch of the Y-model derivation.

The model's name reflects Halle and Marantz's insight that the properties of traditional lexical items actually are distributed across separate components of the grammar, rather than being collected in a single list of sound/meaning correspondences with structural annotations, as in a more traditional lexicon. Instead, there are three such lists, each of which is relevant to only a subset of the fuctions of the lexicon in a lexicalist theory. One list contains the formatives which enter the syntactic computation. These are bundles of morphosyntactic features specifying structural relations, satisfied in the syntax by the usual syntactic operations – Merge, Move and Agree, in current Minimalist terminology. A second list specifies the phonological forms which compete to realize the terminal nodes of a completed syntactic derivation, after Spell-Out to the PF branch. The third list specifies interpretive operations which similarly 'realize', in a semantic sense, the terminal nodes of a completed syntactic derivation. These interpretations will compose with each other, if all proceeds convergently, to produce the meaning of the final structure.

The model is illustrated in (1) below. The points in the derivation at which the elements from List 1, List 2, and List 3 are accessed are indicated.

(1) The model: Distributed Morphology (Halle & A. Marantz, 1993)



List 1: *Feature bundles*: Syntactic primitives, both interpretable and uninterpretable, functional and contentful.[2]

List 2: *Vocabulary Items*: Instructions for pronouncing terminal nodes in context

List 3: *Encyclopedia*: Instructions for interpreting terminal nodes in context

2 In the DM literature, elements of List 1 are termed 'abstract morphemes' which have 'positions of exponence', while elements of List 2 are 'vocabulary items'. In the present paper, the term 'abstract morpheme' is avoided in favor of 'terminal node', or 'feature bundle'; 'position of exponence' may also occur. List 2 items are 'vocabulary items', 'phonological realizations' or 'exponents'.

A derivation begins with a selection of several feature bundles from List 1, including some roots, whose category is notated √, following Pesetsky 1995. This selection produces a set called the Numeration, in the sense of Chomsky 1995. The syntax constructs a well-formed structure from these elements, which is at some point (perhaps in several phasal iterations) handed off to PF and LF, in an operation called 'Spell-Out'. On the PF branch, some morphological operations idiosyncratic to the language may apply, altering the syntactic structure in certain constrained ways to conform to morphological requirements. Following the morphological step, elements from List 2 are accessed. Each terminal node in the structure emerging from the syntax represents a "position of exponence", which must receive some phonological interpretation. List 2 elements compete to provide phonological realizations for these positions of exponence according to the Subset Principle (Halle 1997), a version of Kiparsky's (1973) Elsewhere Condition.[3] The Subset Principle requires that the element of List 2 which realizes a given position of exponence is the most highly specified appropriate realization node. This ensures that more highly specified forms will block the insertion of equally compatible but less-specified forms, in the familiar pattern – the irregular, more specified participle suffix *-en* in *beaten* blocks the regular, less specified participle suffix *-ed*, predicting the ill-formedness of **beated*, for example. On the other interpretive branch, the conceptual/intensional interface looks up model-theoretic interpretations for each terminal node – the elements of List 3 – providing semantic realizations for every feature bundle (and root). These interact with each other in standard model-theoretic fashion to derive a compositional interpretation for the entire structure.

In the original vision of the framework, different roots were not individuated in List 1, nor, therefore, were they individuated in the syntactic derivation. Only features relevant to the syntax were represented in List 1, and extraneous information which the syntactic computation did not attend to was only considered to be accessed when it became necessary, at PF and LF. (Marantz 1995b: 16) wrote:

> There are two basic reasons to treat "cat" and all so-called lexical roots as we treat inflectional affixes, and insert them late. . . . First, it's extremely difficult to argue that roots behave any differently from affixes with respect to the computational system. No phonological properties of roots interact with the principles or computations of syntax, nor do idiosyncratic Encyclopedic facts about roots show any such interactions.

In other words, the phonological and encyclopedic information which differentiate 'cat' from 'dog' are not present in the root nodes drawn from List 1 to form the

3 This aspect of the model was based on the results from studies in other realizational theories of morphology, particularly that of Anderson (1992).

Numeration of a syntactic derivation, since this information is not relevant to the syntax. The only root-related features that are relevant to the syntactic computation, in Marantz's original conception, were features like [±count], [±animate], etc. Underspecified root terminal nodes occurred in List 1 which were bundled with such features, but that was the extent of the differentiation between root nodes. These abstract root nodes would then be subject to late insertion, exactly as for other terminal nodes. In principle, any root Vocabulary Item from List 2 which was consistent with the features of a given root node could be inserted into that node. That is, √dog and √cat were considered to be equally well suited to insertion at any [+count] root terminal node.[4]

This entailed that the List 2 Vocabulary Items which realize root nodes had one unique property in the model: their insertion was not subject to competition, as the insertion of functional Vocabulary Items was. Rather, at PF, the speaker had a choice as to which root VI to insert in any given node, based on the entire morphosyntactic derivation to that point, and their communicative intent. ((Marantz 1995b): 17) highlights this point, and notes the significant consequences this late differentiation of roots has for the semantic interpretation of a completed derivation:

> Late insertion involves making a specific claim about the connection between LF and semantic interpretation. LF can't by itself be the input to semantic interpretation. If "cat" is inserted in the phonology at a node at which "dog" could just as well have been inserted – and if, as we assume, the difference between "cat" and "dog" makes a difference in semantic interpretation – then the phonological representation, specifically the choice of Vocabulary items, must also be input to semantic interpretation.

This conception of the model thus required the interpretive interface to access *both* the PF and LF points of the derivation. This was necessary to prevent the possibility of a derivation in which the vocabulary item /kæt/ is inserted into a root node at PF, while the semantic content DOG is accessed at LF. Instead, both PF and LF were accessed simultaneously by the conceptual-intensional system, guaranteeing that the semantic information associated with the List 2 item /kæt/ was correctly introduced into the interpretation. The interpretation was thus constructed based on the outcome of the whole derivation, including both PF and LF.

In the next section, we turn to an argument against the concept of free-choice late insertion of root Vocabulary Items from List 2, showing that for a certain class

4 Cf Acquaviva's (2008) emphasis on the distinction between root-as-node and root-as-exponent: In early DM, the distinction became somewhat confused in terminology, since the root-as-exponent from List 2 contributed all of the information individuating roots in the model.

of cases, root Vocabulary Items must be in competition with each other. These cases involve root suppletion, and they indicate directly that the difference between different abstract root elements – the difference between "cat" and "dog" – must be present before Vocabulary Insertion. That is, the root node realized as /kæt/ and the root node realized as /dɑg/ must be distinct in List 1, as well as in lists 2 and 3. The cases we will consider also force the conclusion that roots are not individuated on the basis of their phonological content.

2.1 Roots are individuated in the narrow syntax: Root suppletion cross-linguistically

To recap: Because the phonological and encyclopedic distinctions between terms for cats and dogs are not relevant to the syntactic derivation, Marantz (1995) concluded that a root terminal node ultimately realized as 'cat' and one ultimately realized as 'dog' are not distinguished in the syntax: an abstract List 1 terminal node $\sqrt{}_{[+\text{count}]}$ could be realized either way. This, in turn, entailed that root insertion was governed by speaker choice, rather than by competition. The idea was that, if one wants to communicate the content of "The cat sat on the mat", one chooses /kæt/ and /mæt/ at Spell-Out and inserts them into the relevant root terminal nodes. On the other hand, if one wants to communicate "The dog sat on the log", one chooses /dɑg/ and /lɑg/ for insertion into the same nodes.

As pointed out by Marantz (1995, 1997) this view of root realization is unsustainable if there is true root suppletion. If a root can have two phonologically unrelated forms, one of which blocks the insertion of the other in a given morphosyntactic context, it would be evidence for competition-driven insertion of root Vocabulary Items, rather than free choice insertion.[5] Free-choice late insertion and root suppletion are incompatible.

5 Phonologically similar root forms which appear in different contexts, such as *goose* and *geese*, can be accomodated in the morphophonology in the DM model, rather than requiring root competition. A single Vocabularly Item, such as /gus/, realizes the root node, and subsequently a morphophonological 'rule of readjustment' applies to map /u/ to /i/ in the context of [+pl]. This morphophonological rule will apply regularly to a specially marked subclass of root vocabulary items. Such sub-phonologies ('co-phonologies') for particular morphological classes of elements are quite common cross-linguistically, and must be accomodated in any model, whether rule-based or optimality-theoretic (see Inkelas and Orgun 1995, Inkelas 1998, Antilla 2002, Inkelas and Zoll 2007, among many others). Of course, if root competition is admitted into the model, as I argue below it must be, this kind of root allomorphy can instead be taken care of with root competition, as in Siddiqi (2006, 2009) and Chung (2009), among others.

Marantz (1995b) illustrates this incompatibility with a thought experiment. He asks the reader to imagine that /dɑg/ has a special suppletive form /hawnd/ which necessarily appears in the context of [+pl], blocking the insertion of /dag/. In that case if root terminal nodes bear no features other than those relevant to the syntax, then the special suppletive form /hawnd/ will block not just /dɑg/, but also *any* other less-specified root Vocabulary Item from being inserted, by the Subset Principle. That is, the suppletive form /hawnd/ would also block the insertion of /kæt/ in the [+pl] context, since it is more highly specified than /kæt/ and would be compatible with the content of the root node.

The conclusion was that either free choice late insertion is incorrect and roots are fully specified, being distinguished in List 1, as well as Lists 2 and 3, or that true root suppletion does not exist. Marantz (1997) makes a plausible case for the latter position. It is well-known that word learners assume that novel phonological signs map to unknown meanings; this is known as the 'mutual exclusivity principle' (see, e.g. Markman et al. (2003) for an overview.[6] In the domain of roots, with a potentially infinite set of meanings to rule out, it would be reasonable for a word learner to consider this an inviolable principle. This would in turn prevent any phonologically wholly distinct sign from being assigned an identical meaning with another already learned sign, which is what would be required by true root suppletion.

In contrast, suppletion in functional categories appears to be quite common and relatively easily learned; children are well able to acquire morphologically conditioned allomorphs, for example of [+pl] in English (*-en* vs *-i* vs -Ø vs *-s*), or of [+past] (*-t* vs *-d* vs -Ø). This kind of learning follows the famous U-shaped learning curve for irregulars Marcus et al. (1992), showing that it is initially difficult for the learner to associate two distinct phonological exponents to a single underlying featural category. However, in the functional domain, it is clearly possible. Marantz pointed out that the search space for functional category meanings is fixed and limited, provided by UG. He argues that the learner, who may at first assume that *-s* and *-en* have distinct interpretations based on the mutual exclusivity preference, can perform reanalysis when they realize that *oxen* occurs in the same morphosyntactically and semantically plural contexts as *cows* (e.g. following *those*), and that the expected form **oxes* or **oxens* does not occur in these contexts. Because the learner is searching for the phonologi-

6 Note that the mutual exclusivity principle can be seen in operation in other species' learning of sign-symbol mappings. Even Chaser the word-learning dog obeys this principle; it's not specific to humans (Pilley & Reid 2011).

cal exponent of a UG-given feature, whose existence and content they can deduce from global properties of the structure, suppletive realizations of functional morphemes can be learned. In the domain of roots, however, whose meanings are in principle extremely variable and potentially arbitrary, it is plausible to think that such suppletion is in principle unlearnable. To maintain the free-choice model for late insertion of root Vocabulary Items into underspecified root nodes and keep the syntax free of extraneous phonological and encyclopedic information, Marantz suggested that root suppletion was in fact impossible to learn.

There is apparent root suppletion in English, however, in a few restricted cases, some of which are enumerated below.

(2) English:
   a. go ~ wen- 'GO ~ GO.pst'
   b. bad ~ worse 'BAD ~ BAD.Compar'
   c. person ~ people 'PERSON.sg ~ PERSON.pl'

Marantz's response to this problem of apparent suppletion in roots is to suggest that such cases in fact represent realizations of functional categories, such as the hypothetical categorizing heads *v*, *a* or *n*, rather than realizations of root terminal nodes. The meanings of the root-like elements that show suppletion in English are suitably 'light' in character, arguably encoding adjectival, verbal, and nominal universal features: *go/went* realize a 'light verb' functional category *v* (perhaps bearing a hypothetical universal feature [+Path]; *bad/worse* a 'light adjective' category *a* (perhaps bearing universal features [+Negative, +Evaluative]), and *person/people* a 'light noun' functional category *n* (perhaps bearing a universal feature [+human]). Their meanings in each case are suitably bleached and plausibly universal in character, and if English were the only case in which suppletive stems were known to exist, it's possible that the case against suppletion in root forms could be maintained.

However, when considering a broader cross-linguistic dataset, it becomes apparent that true root suppletion does exist after all: There are suppletive lexical items which cannot be considered to be instances of quasi-functional categories. Consider, for example, the following suppletive verbs of Hiaki[7], a Uto-Aztecan language spoken in Sonora and Arizona:

7 Also known as Yaqui and Yoeme.

(3) Hiaki:

a.	vuite ~ tenne	'run.sg ~ run.pl'
b.	siika ~ saka	'go.sg ~ go.pl'
c.	weama ~ rehte	'wander.sg ~ wander.pl'
d.	kivake ~ kiime	'enter.sg ~ enter.pl'
e.	vo'e ~ to'e	'lie.sg ~ lie.pl'
f.	weye ~ kaate	'walk.sg ~ walk.pl';
g.	mea ~ sua	'kill.sgObj ~ kill.plObj'

The above represents a selection from a set of about 14–15 total suppletive verbs in the language; the particular set varies somewhat across dialects, but the seven listed above are among those which are consistent. This is a typical Uto-Aztecan pattern; most Uto-Aztecan languages have at least a few suppletive verbs of this type, and some have more than Hiaki. Most of these verbs are clearly main verbs, not light verbs, in terms of both their semantically rich content and in terms of their behavior in the language.

Looking at suppletion across other language families produces a similar result. Veselinova (2003, 2006) surveys verbal suppletion in 193 languages, focussing particularly on suppletion conditioned by number and suppletion conditioned by tense/aspect. To address the question of what types of meanings are encoded by such suppletive verbs, she provides 'lexical type tables', which list and categorize the glosses of each suppletive verb from any language in her database. In (4) below, I reproduce her categorized lists of glosses for verbs exhibiting number-conditioned suppletion crosslinguistically Veselinova (2003: 222–224).[8] The macrocategories into which the glosses are grouped are those chosen by Veselinova; for our purposes, however, the key thing to focus on is the content of the glosses themselves. While the behavior of each verb in each language cannot be deduced from this list of glosses, and while grammaticalization from lexical verb to light verb can take many paths, I submit that the meanings reflected by many of these glosses are unlikely to be realizations of universal syntacticosemantic 'light verb' categories. I have bolded items in the lists below which to me seem to be particularly implausible candidates for light verb meanings.

8 Note that no one language contains this many suppletive verbs. This is the cumulative list of glosses of suppletive verbs from 193 languages. Each single language might have suppletive verbs corresponding to only one or two, or a handful, of the verb glosses listed here (as noted above, Hiaki has slightly more than a dozen such verbs).

(4) *Glosses of suppletive verbs whose suppletion is conditioned by number cross-linguistically (Veselinova 2003: 222–224):*
   a. *Motion, intransitive*: go, fall, come, run, arrive, enter, start, get.up, return, rise, walk, **fall.in.water**, fly, go.about, **go.around.something.out.of.sight**, jump, move, stampede, **swim**, visit, walk
   b. *Motion, transitive*: put, throw, take, give, drive.out, get, grasp, pick.up, pull.out, release, remove, take.out
   c. *Position*: sit, lie, stand, hold, carry, store
   d. *Die/Injure*: beat, **bite.off**, cut, die.of.old.age.or.hunger, injure, kill, break, hit
   e. *Stative*: sleep, big, small, be.at, be lost, exist, long, short
   f. *Other*: eat, belong.to, **bet**, **make.netbag**, make.noise, not.like, say

Veselinova gives a similar list for tense/aspect suppletion, which again I reproduce below, again bolding those suppletive verb glosses that strike me as relatively non-functional in character:

(5) *Glosses of suppletive verbs whose suppletion is conditioned by aspect cross-linguistically (Veselinova 2003: 115–116):*
   come/go, be/exist, say/speak, do, take, see/watch, eat, give/lay, put, die, become, sit/stand/stay, carry, **catch**, get, have, **hear**, throw, **beat**, **become.cold**, become/happen/go, cry, **drink**, fall, live/move, run, stay/continue, **wake up**, walk

If true root suppletion exists, as suggested by the data above, it must be the case that the mutual exclusivity assumption is just a heuristic, rather than a hard-and-fast inviolable principle. Mutual exclusivity can guide the learner, but given enough evidence, over time a learner can conflate the lexical entry of two phonologically distinct root Vocabulary Items, producing true suppletion. Within any language where such suppletion exists, it must certainly be the case that the suppletive items must have a very high token frequency, or else the suppletive alternation would be effectively unlearnable. It is this necessarily high frequency which in turn accounts for the kind of semantic categories which end up developing suppletive forms. The set of the highest frequency verbs verbs in any language are likely to be light-verb-like and have a universal semantic flavor to them. People everywhere frequently speak of activities intrinsic to the human condition. High-frequency items are also those which are subject to grammaticalization, hence the overlap between suppletive verb meanings and light verb meanings. However, the exceptions noted in bold above show that

grammaticalization is not a necessary precondition for the development of suppletion. In Hiaki, it is clear that suppletion of a given verb is not sensitive to whether it has a 'light' verb function or not; these verbs supplete when used as main verbs. I conclude that these are indeed suppletive √ exponents, competing to realize a single √ position.

With that conclusion in mind, let us consider the derivation of Hiaki sentences like those in (6)

(6) a. Aapo aman vuite-k. (*Vempo aman vuite-k.)
3sg there run.sg-prf 3pl there run.sg-prf
"He ran over there."
b. Vempo aman tenne-k (*Aapo aman tenne-k.)
3pl there run.pl-prf 3sg there run.pl-prf
"They ran over there."

Following the syntactic derivation, a √ node in the verb phrase is competed for by the Vocabulary items √*vuite* and √*tenne* from List 2. The item √*tenne* wins just in case the morphosyntactic context contains a plural argument, while the item √*vuite* appears elsewhere. That is, √*tenne* blocks √*vuite*, in the morphological sense, in the same way that √*wen-* blocks √*go* in the past tense in English. This is summarized by the Vocabulary Item entries in (7) below:

(7) a. √ ←→ /tenne/ / [DP_{pl} ____]
b. √ ←→ /vuite/ Elsewhere

It is imperative that the List 1 √ node, on the left hand side of these Vocabulary Items – the target of competition – be identified as distinct from other intransitive verb roots. Otherwise, √*tenne* will block the insertion of *any* other non-suppleting intransitive verb with a plural subject, as in Marantz's thought experiment above. This is because √*tenne* represents a more highly specified match for the √ node, and by the Subset Principle, more highly specified matches always block the insertion of less-specified matches. Consequently, the √ on the left-hand side of the rule which may be realized as *tenne* or *vuite* must be distinguished from other √s, like a √ which may ultimately be realized as non-suppletive *bwiika*, 'sing', or non-suppletive *nooka*, 'talk'.

The correct result could be derived if root nodes were distinguished in List 1 according to semantic criteria. For example, if root nodes in List 1 were

Fodorian atomic concepts (see, e.g. (Fodor 1998)), the rules in (7) could look like this:

(8) a. $\sqrt{\text{RUN}} \longleftrightarrow$ /tenne/ / [DP_{pl} ____]
b. $\sqrt{\text{RUN}} \longleftrightarrow$ /vuite/ Elsewhere

We will see in the next section, however, that such a proposal is unsustainable: the individuation criterion for List 1 roots cannot be semantic in character.

2.2 Root individuation in the syntax is not phonological

Before turning to a discussion of semantic individuation, however, I wish to draw out more clearly a corollary of the above discussion. We have so far focussed on the idea that roots cannot be underspecified in the syntax, but rather must be individuated before spell-out, in order to allow for competition between suppletive vocabulary items competing for specific root terminal nodes. A secondary, and equally important point, which should be clear from the above but which merits explicit comment, is that the individuation criteria for roots in List 1 cannot be phonological in character. That is, the existence of suppletive root competition proves that root terminal nodes are subject to late insertion, just like all other terminal nodes, as pointed out in Marantz (1995b). It cannot be the case that elements in List 1 are specified for phonological content, like √*kæt* (contra, among others, Borer (2009)). If they did, root suppletion could not exist; it would be an incoherent notion.

Borer (2009) discusses exactly this consequence as part of developing a model in which roots are phonologically individuated in the syntax. She hypothesizes that "suppletive pairs such as *go/went* constitute two, rather than one, roots with phonological gaps." That is, in her model, *went* does not block **goed* in a morphological sense at all. On such a view it becomes a simple coincidence that the root √*go* has gaps in its past tense distribution, while √*went* has gaps in exactly the complementary slots in its present and participial distributions. More disturbingly, it becomes a coincidence that their semantic extensions are exactly and perfectly overlapping. In a model in which 'go' and 'went' are suppletive realizations of an identical underlying root, idioms formed with the root realized by √*go*, like 'go around the bend' or 'go for it', will have have past tenses formed with √*went* just as for other uses of *go*. In contrast, if √*go* is a separate root from √*went*, as in Borer's model, it is not clear why the idiomatic readings of one should have anything to do with those of the other. With Aronoff (2011), among

others, I take covariation in contextually-determined interpretations to be one ideal kind of evidence for the existence of suppletion – the other being, of course, the speakers' intuitions about morphological blocking, and the ill-formedness of **goed*.

To recap: if roots went into the syntax fully specified for their phonological shape, a suppletive form could not compete to realize a √ node postsyntactically, conditioned by the syntactic context created by the construction of the sentence. That would be equivalent to treating suppletion as a phonological rewriting, a postsyntactic readjustment rule that would overwrite /vuite/ with /tenne/, or /gow/ with /wɛnt/. The undesirability of such an enrichment of the phonological system has been extensively commented on by many more knowledgable than I, and I will not belabor it further here. Root terminal nodes cannot be distinguished on the basis of their phonological signatures.

Next we turn to consider the viability of the hypothesis instantiated by the vocabulary items formalized in (8): Might it be the case that roots in List 1 are individuated on the basis of conceptual information? Such an approach is proposed by, e.g. Siddiqi (2006). However, we will see that there are cases which pose an analogous problem for LF as root suppletion poses for PF: There are roots whose meaning clearly cannot be determined outside of a particular syntactic context. I call these *caboodle* items; they are perhaps more familiar under the name *cran*-morphs.

2.3 Root individuation in the syntax is not semantic

The special property of suppletive roots is that their phonological form is not identifiable prior to its appearance in a derived morphosyntacic context – until you have the broader syntactic context, you cannot know how to pronounce them. To show that root terminal nodes cannot be semantically individuated, then, we need to establish that there are roots whose semantic interpretation is not identifiable prior to its appearance in a derived morphosyntactic context. In fact, such cases are well-documented in the literature.

One well-known instance of the general phenomenon is provided by the consonantal roots of Hebrew, alluded to earlier. Aronoff (2007) among many others, provides extensive argumentation that Hebrew verb roots are individuated morphological entities whose properties bear little or no relationship to meaning. Below I reproduce Aronoff's Table 6 (Aronoff 2007: 822), which illustrates the diverse range of meanings expressed by the root √*kbʃ* in different morphological contexts – in different binyanim, and with different affixes:

(9) Morphologically real root without clear semantic individuation: Aronoff 2007

Root = kbʃ ~ 'press'	Synchronic meaning:
Nouns	
keveʃ	'gangway, step, degree, pickled fruit'
kviʃ	'paved road, highway'
kviʃa	'compression'
kivʃan	'furnace, kiln'
maxbeʃ	'press, road roller'
mixbaʃa	pickling shop
Verbs	
kavaʃ	'to conquer, subdue, press, pave, pickle, preserve, store, hide'
kibeʃ	'to conquer, subdue, press, pave, pickle, preserve'
hixbiʃ	'subdue, subjugate'
Adjectives	
kavuʃ	'subdued, conquered, preserved, pressed, paved'
kvuʃim	'conserves, preserves'
mexubaʃ	'pressed, full'

In Aronoff's words, "trying to find a common meaning shared by pickles and highways brings one close to empirical emptiness".[9] And yet, the entity √*kbʃ* is a morphologically real and stored element of the synchronic Hebrew grammatical system in all of these uses.[10] This is not simply a bunch of distinct words containing a homophonous set of consonants, not related in the synchronic grammar. Aronoff is able to prove that this is the case by showing that every Hebrew root,

9 Again, the analogy to the phonological situation is very close to complete: trying to find a common phonological form shared by *go* and *went*, as required by a system in which roots are identified by a phonological form, also brings one close to empirical emptiness – and for the same reason.

10 Or at least most of them; Edit Doron (p.c.) points out that the 'gangway' meaning derives from an Aramaic root meaning 'descend', while all the others from a homophonous Akkadian root meaning 'press, tread on'. Eliminating the 'gangway' meaning from consideration as a potentially homophonous confound does not substantially change the overall picture, however; pickles and highways still are semantically disparate enough to make Aronoff's point with this example. See Moscoso et al. (2005) for some psycholinguistic evidence concerning the (non)-identity of some such cases.

regardless of its interpretive variation, belongs to a morphological alternation class which predicts its distribution and interaction with other morphological formatives of Hebrew grammar. Phonologically similar triconsonantal roots can belong to different alternation classes, as Aronoff illustrates for √*npl* '~fall' and √*npk* '~issue'; the former belongs to a marked class of roots which lose their initial consonant when a prefix is attached; the latter is a regular root which retains its initial consonant under prefixation. The class of initial-consonant-deleting roots is heterogenous, including roots beginning with *n, y* and *l*, and its members can often be phonologicaly similar to, or even identical with, roots whose behavior is completely regular. Because the deletion pattern cannot be derived from general properties of the phonological system, the consonant-deleting roots constitute an irregular morphological alternation class (the form √*npl* alternating with the trucated form √*pl* under prefixation). The alternation has become a stored property associated with particular roots, whose class membership identified as a property of the root, *regardless of which meanings it receives in which contexts*. The fact of alternation class membership thus proves the integrity of the root as an individual listed item in the mind of the speaker, across all of its different semantic interpretations, since it participates in the alternation regardless of which meaning it is carrying at the time.

A completely analogous case can be made from a set of patterns in English whose significance for morphological analysis in this regard was also first pointed out by Aronoff, in 1976. There is a well-known class of identifiable roots in English which are entirely meaningless outside of their morphosyntactic context:

(10) a. *-ceive*
deceive, receive, conceive, perceive
b. *-here*
adhere, inhere
c. *-port*
comport, deport, report, import, support
d. *-pose*
suppose, depose, compose, repose, propose
. . . etc

Despite their semanticaly underdetermined nature, these are clearly diagnosable as root elements of English by an acquiring child or linguist. Besides their phonological identity across contexts, their special prosodic properties and occasionally their special phonotactic properties (see Harley (2009) for a review), they can also show contextual allomorphy and impose morphological selectional restrictions regardless of the lexical item they appear in:

(11) a. *-ceive ~ -cept + ion*
deception, reception, conception, perception
b. *-pose ~ -pos + ition* (not *-ation* or *-ion . . .*)
composition, supposition, proposition, deposition

These roots, therefore, are clearly individuated elements in the grammar of English. It would be redundant to list allomorphs for *deceive ~ deception, receive ~ reception, perceive ~ perception* individually; the *ceive ~ cept* alternation is a property of the *-ceive* root itself, which is why it behave the same way across lexical items and in imaginary nonce items formed from *-ceive* (*#acceive, #acception*).

Even though they are listed individual elements, *ceive*-type items are meaningless outside particular morphosyntactic contexts.[11] Ergo, they are not individuated by their meanings. As noted by Marantz (1995b) this conclusion concerning the interpretation of bound roots is surprising only from the perspective of speakers of relatively isolating languages like English; it is almost self-evident when looking at languages whose roots are typically morphologically bound, as in Hebrew.

There are also roots whose interpretation is wholly dependent on occurrence in a particular purely *syntactic* frame – not, as in the case of the *-ceive* items, dependent on a word-internal morphological frame, but an entire idiomatic phrasal constituent. Consider the following English cases:

(12) a. kit and **caboodle** ‘everything’
b. run the **gamut** ‘includes a whole range’
c. by **dint** of[12] ‘by means of’
d. in **cahoot**s ‘conspiring’
e. **vim** and vigor ‘vitality’
f. high **jink**s ‘mischief’
g. **kith** and kin ‘friends and relations’

Indeed, in the grammar of any given speaker, it is likely that there are several undetected examples of such *caboodle* items, where the speaker has learned an expression and its meaning as a phrase without having yet learned an independent meaning for each of the individual items contained within it which would allow them to be recombined compositionally in other contexts. This kind of ‘semantic chunking’ does not entail syntactic or morphological chunking; *high*

11 See also Baeskow (2006) for additional discussion.
12 Example from Nunberg et al. (1994).

jinks, for example, is morphologically plural (cf. *I don't care for these/*this high jinks*), despite the unproductivity of *jink* outside the context of [+pl] and the adjective *high*. The syntax of such expressions is completely unremarkable, and functional units within them do the morphological job which they typically do. The only special property has to do with the context-dependence of the List 3 interpretation of the root.[13]

In short, just as one does not know how to pronounce a suppletive root outside a morphosyntactic context, one also does not know how to interpret a *caboodle* root outside a morphosyntactic context. The necessary conclusion is that syntactic roots are not interpretively individuated, either. The notion that the numeration contains roots identified by their atomic conceptual content, as speculated in (8) above, can't be right: There's no such item as √RUN in List 1.

The elements of List 1 of category √, therefore, must be individuated, but no single type of independent interface property can be taken to individuate them. They are simply units of morphosyntactic computation – abstract morphemes in the truest sense. We cannot individuate them by their phonological properties, which may depend on the derived morphosyntactic context; neither can we individuate them by their interpretive properties, for the same reason. In the next section, a sketch of the system whose shape emerges from the above discussion is provided.

2.4 Identity criteria: Nonsemantic, nonphonological

Above we concluded that roots from List 1 – the roots which are manipulated by the syntactic derivation – must have individuation criteria that do not depend on semantic or phonological content. They are individual units of morphosyntactic computation. We can identify these roots using an index notation, as proposed by Pfau (2000, 2009) and Acquaviva (2008).

Root vocabulary item competition can then be defined with respect to these indices, as can semantic interpretation. The identification of the correct interpretation of a given root in context, will work a lot like the identification of the correct vocabulary item for a root in context.[14]

The root terminal node elements ocurring in List 1 can thus be notated as $\sqrt_{279}$, $\sqrt_{322}$, $\sqrt_{2588}$, etc. List 2 consists of instructions for spelling out each of these entities

13 The extension to syntactic contexts, as well as morphological ones, is the reason I have chosen to rename these *caboodle* items, rather than simply use the more familiar term '*cran*-morph'.
14 On this view, production and parsing would be mirror images of each other, working 'forwards' from a semantic representation or 'backwards' from a phonological representation.

in a given morphosyntactic context. List 3 consists of instructions for interpreting these entities in a given morphosyntactic context.

Below, I give examples of List 2 and List 3 entries which might be accessed in response to a given root terminal node in the output of a syntactic derivation. The interpretive instructions given as the List 3 entry ("tape" etc) should be construed as shorthand for a meaning expressed in model-theoretic terms, as proposed in Doron (2003). I assume that these meanings exploit a basic ontology of conceptual entities, as proposed in Harley (2005a). That is, the various √ items may be have interpretations as predicates of entities, (e.g. the interpretation of the root of *calve* or *saddle*), predicates of properties (e.g. the interpretation of the root of *open* or *melt*) or predicates of events (e.g. the interpretation of the root of *run* or *dance*).[15] This is consistent with the observations of Marantz (2001, 2008). He observes that since some category-forming morphemes can attach both to roots (e.g. *atroc-ity*, from $\sqrt{atroc}$- + -ity_n) and to derived (already categorized) forms (e.g. *electr-ic-ity,* from $[\sqrt{electr\text{-}ic}_a]_{aP}$ + -ity_n), at least some root interpretations must be similar to the interpretations of derived nPs, aPs and vPs – by hypothesis, predicates of entities, properties and events, respectively.

In an idealized basic case, a root will have an invariant pronunciation across different contexts, and an invariant interpretation as well. Such a root would be a perfect Saussurean sign, giving the appearance of a straightforward linkage of sound and meaning. A potential example of such a case in English is given in (13). The phonological instructions on the left are contained in List 2, the list of Vocabulary Items; on the right, the interpretive instructions are contained in List 3, accessed when it is time to provide a syntactic structure with a compositional interpretation:

(13) Basic case: Interface instructions for a root node that is a Saussurean sign

PF instructions (List 2)	LF instructions (List 3)
$\sqrt{}_{279} \longleftrightarrow$ /tejp/	$\sqrt{}_{279} \longleftrightarrow$ "tape"

As noted above, the instructions on the LF side as I present them above are promissory notes only: informal representations of model-theoretic interpretations along the lines proposed by Doron (2003); "tape" here stands for whatever function will produce the correct predicate of entities in a nominal syntactic environment, e.g. one whose truth conditions involve something like "flexible thin flat material used to attach or bind, usually with a sticky side."

15 Thanks to Elena Anagnostopoulou for helpful discussion on this point.

An example of the interface instructions for the suppletive Hiaki roots described in section 2.1 above is given in (3); again "run" on the right hand side of the LF instruction entry is shorthand for an appropriate model-theoretic formula:[16]

(14) Interface instructions for a Hiaki suppletive root node

PF instructions (List 2)	LF instructions (List 3)
$\sqrt{}_{322} \longleftrightarrow$ /vuite/ / [[DP$_{\text{-pl}}$] ____$_{\sqrt}$]	$\sqrt{}_{322} \longleftrightarrow$ "run"
$\longleftrightarrow$ /tenne/ elsewhere	

The analogous situation in List 3 is the case of idioms, where a List 1 root terminal node has only one set of instructions on the PF side, but multiple interpretations are available on the LF side.

(15) Interface instructions for a root node with idiomatic interpretations in English

PF instructions (List 2)	LF instructions (List 3)
$\sqrt{}_{77} \longleftrightarrow$ /θrow/	$\sqrt{}_{77} \longleftrightarrow$ "vomit" /[v [[__]$_{\sqrt}$ [up]$_{\text{P}}$]]$_{\text{vP}}$
	$\longleftrightarrow$ "a light blanket" / [n [__]$_{\sqrt}$]
	{. . . other meanings in other contexts . . .}
	$\longleftrightarrow$ "throw" elsewhere

A *caboodle* item will have the special property of lacking 'elsewhere' interpretive instructions on the LF side, as illustrated in (16):

(16) Interface instructions for the root node for *cahoot* (a cran-morph, from the list in (12)) in English

PF instructions (List 2)	LF instructions (List 3)
$\sqrt{}_{548} \longleftrightarrow$ /kəhut/	$\sqrt{}_{548} \longleftrightarrow$ "a conspiracy" / [in [[____$_{\sqrt}$ n]$_{\text{nP}}$ -PL]$_{\text{DP}}$]$_{\text{PP}}$
	no Elsewhere interp

(See the discussion below for further commentary on whether 'competition' is relevant for interpretation at LF.)

16 It can be shown that *tenne* is truly an Elsewhere form, not just an allomorph inserted in the environment of a [+pl] nominal. When the argument of *vuite/tenne* is syntactically absent and consequently unspecified for number, as in the Hiaki impersonal passive, the root must surface as *tenne*, not as singular *vuite*. Also, see futher discussion of the structure of the conditioning context for *vuite* in section 3.3 below.

If *ceive/cept* type alternations are cases of suppletion, rather than simple morphophonological readjustment (Siddiqi 2006, 2009, Chung 2009) then these VIs represent the maximally complex case, an entity with contextually dependent interpretations both at PF and at LF. The List 2 and List 3 items which would provide the interface interpretations for the *ceive* root at PF and LF would then look like this:[17]

(17) Interface instructions for the root node for *-ceive*:

PF instructions (List 2)	LF instructions (List 3)
$\sqrt{}_{683} \longleftrightarrow$ *cept* / [. . .[____] n_{event}]	$\sqrt{}_{683} \longleftrightarrow$ "think"[18] / [v [[con-]$_P$ [__]$_{\sqrt{}}$]]$_{vP}$
$\longleftrightarrow$ *ceive* elsewhere	$\longleftrightarrow$ "fake" / [v [[de-]$_P$ [__]$_{\sqrt{}}$]]$_{vP}$
	{. . . other meanings with *re-*, *per-*, etc . . .}
	no Elsewhere interp

One note on the notion of 'elsewhere' in relation to LF interpretation is in order. The one significant formal difference between the LF instructions provided in List 3 and the PF instructions provided in List 2 is that the PF instructions include a form to be used 'elsewhere' – a least-specified form which wins the competition

17 P. Svenonius (p.c.) brings up cases where the two suppletive variants of a particular root, while remaining in a productive alternation in the main, have developed independent particular idiosyncratic meanings. For example, each member of the plural/singular *people* ~ *person* alternation occurs in particular contexts where the alternation is not productive. When this root occurs as a denominal locatum verb, for example, it's always *people*: *to people/*person the planet*. In contrast, in the context of official search-and-rescue operations, we always have *person*, even in the plural, losing its idiosyncratic plural: *The Missing Persons Bureau*. For this case, I suggest that *people* is the elsewhere form, *person* being specified to occur in the context of a [+sg] Num° head; thus *people* appears in the verbal as well as the nominal environment. In the special context of search-and-rescue (or other contexts where the individual's particular body is salient), we are dealing with a separate, half-homophonous root, realized by *person*. There are similar cases in the domain of Latinate verbs; consider, for example, the verb *to self-destruct*, which in undergoing backformation from *self-destruction* lost its identity with the root exhibiting *-stroy* ~ *-struct* alternations: **to self-destroy*.

18 This stands for a function that ultimately yields a predicate of events after composing with *con-*, as in the phrasal verb *think up* – 'con' contributes its (telic) content compositionally. Similary in the context of *de-*, the interpretation given, "fake", stands for a predicate of events like that in *fake out* – the P realized by *de-* contributes its telic content compositionally. Note also that *-ceive* may be specified as generating a second meaning in the context of *con-* as well, to do with pregnancy, when it composes with an object DP denoting a person. See section 3.2 below for further discussion of the conditioning of special meanings in syntactic contexts, following the treatment of verb-object idioms put forward in Kratzer (1996).

to realize the node when the node appears in any context other than one eligible for realization by a more highly specified competitor for that node. Nodes with an elsewhere realization will never suffer from a paradigm gap; there will always be a form which can be inserted to represent that node's content.

In contrast, it is not clear that the concept of an 'elsewhere' interpretation is coherent as part of the LF interpretive instructions which make up List 3. Empirically, it seems clear that some items must lack such an 'elsewhere'; that is the fact of the matter for *caboodle* roots, which can only appear in a single context. What about more typical roots, which are free to compose productively in syntax? In (15) the 'literal' meaning is listed as the elsewhere interpretation for what I have labelled root $\sqrt{}_{77}$, 'throw', but in fact, the nature of the model entails that this is most likely not correct. Model-theoretic interpretations must compose with the interpretations of other elements in their syntactic environment using one of a limited number of composition operations, most commonly function application (see, e.g. Heim & Kratzer (1998) for discussion). Even the 'literal' meaning of a root is only well-formed if its type-theoretic restrictions are satisfied by the interpretations of entities with which it is merged. If a root is contained in a syntactic environment in which its sister's interpretation is type-theoretically incompatible with any of the interpretations specified for the root in List 3, the resulting type-clash produces an ill-formed LF representation for the constituent. That is, it is formally impossible to specify a truly 'elsewhere' interpretation in the domain of roots, since any interpretation must be able to compose with the type of its sister. No interpretation provided by List 3 can provide a well-formed expression that will compose in all imaginable syntactic environments, which is what a truly 'elsewhere' interpretation would have to be. Consequently, the fact that the parallel between List 2 instructions and List 3 instructions breaks down at the concept of 'elsewhere' is expected, given the nature of the LF interface.[19]

19 Type clash can sometimes be resolved by coercion, as when a mass noun appears in a count syntax or vice versa, and also, I assume, in cases like those discussed by Gleitman (1990), where verbal roots which normally take clausal complements appear in a ditransitive syntactic enviroment: examples like *I thought the book to Mary* are interpreted as telekenesis or telepathic transmission. Such coercion operations, however, must be sharply constrained and limited in scope, and cannot rescue just any structure in which type-clash arises, cf Lidz et al.'s (2001) examples like *#The giraffe fell that the money was sick*. It is to be hoped that a full understanding of available coercion operations, in combination with a fully worked-out theory of possible root interpretations, can provide a predictive account of the significantly varied patterns of flexibility in root interpretation.

It is possible to use a *caboodle* item outside the context in which it canonically appears, in language play or other conscious manipulations (e.g. in poetic contexts). Given the framework above, I speculate that such uses will respect the type-theoretic constraints of the interpretation specified in the usual (more constrained) use, but require a nonce reinterpretation of its truth conditions in such a way that they are no longer dependent on the particular meanings contributed by other items. So, for example, the root √jink in *high jinks* has a meaning that normally requires it to compose with the plural morpheme, so its type is compatible with count noun contexts. A independent nonce usage of √jink without *high* can then be derived, as long as it occurs in a count context, as in the following arch and clearly playful passage from Dickens – note the plurality of *jinks* is signalled by the demonstrative *those* and the verb *are*:

> "It is quite time that I think I should explain to you why there should be high jinks at Christoffsky to night (the height of those jinks is the cause of our samovarising, this twenty-first of June, so late or early), where Christoffsky itself is, and what the jinks I have entitled high, are like." (Dickens 1857: 119)

Having established a general picture of root individuation in the syntax, and interpretation at the two interfaces, we next turn to a consideration of the syntactic distribution of root nodes. It is argued that their syntactic properties are unexceptional: they can undergo Merge with other XPs and project, just like any other syntactic category. The complements of roots are shown to condition both their phonological and semantic interpretation, and the complement-taking ability of roots is shown to permit an updating of the standard syntactic account of the distribution of English *one*-replacement in the Bare Phrase Structure framework.

3 Roots and their complements: Syntactic, semantic and morphological evidence

Two recent lines of research on syntactic root nodes have converged on the conclusion that root nodes are radically syntactically deficient. Roots, it is claimed, cannot take complements, cannot head phrasal constituents, and do not impose selectional requirements on structure. It should already be clear that the proposal here is incompatible with at least the last of these conclusions. Below, arguments are laid out whose implications run counter to other two, as well.

Borer (2003, 2009), developing a extensive line of work on the relationship of event and argument structure, argues that roots have neither internal grammatical

structure nor syntactic properties: they are acategorial, monomorphemic, and lack argument structure. This conception of roots is used to address several important problems in morphological and syntactic analysis. Borer argues that it explains why roots must always appear in the context of a categorizer: since a root is an entity which does not have any syntactic properties of its own, it cannot occur in a linguistic context without combining with at least one functional head. Other salient properties of roots are also shown to follow from the approach. In particular, the flexible valence of many verbs in English can be easily understood if roots are radically underspecified for argument structure. It also provides an account of the necessarily verbal nature of true argument-structure nominals: Since argument structure is derived by the projection of additional syntactic structure, rather than being a property of roots themselves, the categorial consequences of the necessary additional structure must be present whenever the arguments are present. Borer argues that since argument structure projections are verbal in character, the necessarily deverbal quality of true argument structure nominals follows.

De Belder and van Craenenbroeck (2011) and De Belder (2011) propose to derive the extreme underspecification of root nodes posited by Borer from independent properties of the syntactic computation. Root nodes, they claim, are structurally an epiphenomenon derived from the special properties of the first Merge operation in a given derivation. In Chomsky's (1994) original description of Bare Phrase Structure, all instances of Merge except the first in any derivation involve drawing a single element from the Numeration and Merging it with an entity already in the workspace. At the point of the first Merge operation, however, there is no element in the workspace. Chomsky proposes that in this one case, not one but two elements are drawn from the Numeration and then Merged. Van Craenenbroek and de Belder rightly observe that this gives the initial Merge operation a different character than any other, and propose instead that the first Merge operation involves drawing a single element from the Numeration and merging it with the empty set representing the empty workspace. The resulting maximally empty, completely featureless node, they argue, is the locus of root insertion. This empty node, a necessary byproduct of the Merge operation, is co-opted to serve as the interface between the narrow syntactic component and the broader cognitive system – exactly the role that lexical roots in general are taken to play, conceptually speaking. There are thus no root feature bundles in List 1 and consequently no roots situated in the numeration, awaiting insertion. Instead, List 1 is composed entirely of functional elements.

It follows from this approach, of course, that root nodes can never project, nor take a complement. All phrasal projection is the projection of functional elements. Before a root could merge with a complement DP, it would first have to

be categorized, presumably by the first Merge of a categorizing head such as *n*, *a* or *v*. The complement DP, having been built in a separate syntactic workspace, would then undergo Merge with the resulting nP or vP.

It seems to me that despite the conceptual appeal of these proposals, they face several empirical hurdles, in that there are phenomena whose analysis requires as a precondition that roots behave like normal syntactic elements, participating in Merge like any other element of the Numeration, even to the point of having arguments as sisters and projecting to the √P category. Below, three analyses are given which suggest that roots can indeed take complements directly. In the first subsection below, a proposal about the syntactic distribution of *one*-replacement is presented Harley (2005c) which suggests that in fact roots do merge with their complements and project to √P before the categorizing head is merged. We then briefly review a proposal of Kratzer's concerning the interpretation of verb-object idioms which is dependent on the same assumption. Finally, evidence that root suppletion in Hiaki is always conditioned by internal arguments is presented, again suggesting that roots and their complements are in a maximally local structural relationship.

3.1 Syntactic evidence: *One*-replacement, roots and objects (Harley 2005c)

One of the most familiar arguments in syntactic theory concerns the behavior of the English N′ anaphor *one*, as initially analyzed by Jackendoff (1977). Within deverbal nominals, arguments and adjuncts behave differently with respect to the *one*-replacement constituent test: Selected arguments, such as *of physics* in (18a) below, cannot be stranded under *one*-replacement of the nominal which selects, while nominal adjuncts, such as *with long hair*, can be, as in (18b).

(18) a. *This $[\text{student}]_{N}$ $[\text{of chemistry}]_{PP}$ and
that $[\text{one}]_{N}$ $[\text{of physics}]_{PP}$ sit together
b. That $[\text{student}]_{N}$ $[\text{with short hair}]_{PP}$ and
this $[\text{one}]_{N}$ $[\text{with long hair}]_{PP}$ sit together

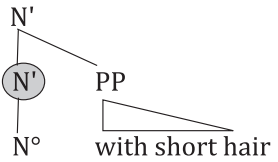
In Jackendoff's original account, phrases such as *This student of chemistry* were treated as NP projections of N°. To account for the difference between argument PPs within NP (like *of phsyics*) and adjunct PPs (like *with long hair*), Jackendoff proposed that *one* was anaphoric to an N′ projection. Arguments such as *of physics* in (18a), being selected by their head nouns, were sisters to N° under N′,

so *one*, targeting N′, rather than N°, could not strand them. In contrast, adjuncts were analysed as sisters and daughters of (potentially recursive) bar-level projections, so *with short hair* in (18b) was sister to N′, daughter of N′. Consequently, *one*-replacement can optionally include an adjunct (when it takes its mother N′ node as its antecedent), or strand it (when it takes the adjunct's sister N′ node). This original analysis is represented in bracket notation in (19) below; the bolded constituent in (19a) represents the only potential antecedent for *one*-replacement in each complex NP, while in (19b), two potential antecedents for *one* exist, identified via bolding in i) and ii):

(19) a. NP with argument PP, sister to N°, daughter of N′: only one antecedent for *one*
[$_{NP}$ That **[$_{N'}$ [$_{N}$ student] [$_{PP}$ of chemistry]]**]
b. NP with adjunct PP, sister to N′, daughter of N′: Two possible antecedents for *one*
i) [$_{NP}$ That **[$_{N'}$ [$_{N'}$ [$_{N}$ student]] [$_{PP}$ with short hair]]**]
ii) [$_{NP}$ That [$_{N'}$ **[$_{N'}$ [$_{N}$ student]]** [$_{PP}$ with short hair]]]

Jackendoff's proposal, however, cannot be implemented in Bare Phrase Structure theory (Chomsky 1994), or its antecedents, (Speas 1986, 1990), because it requires the projection of a nonbranching N′ node in *student with short hair* in the structures in (19b). To make it easier to see this nonbranching node, I provide a tree diagram of the structure in (19b) below:

(20) Nonbranching N′ projection:



Without the mandatory projection of an N′ level above every N, it would remain a mystery that *one* can be anteceded by a constituent consisting only of *student* in (18b), but not by a constituent consisting only of *student* in (18a). However, in Bare Phrase Structure, in which every projection is the result of a Merge operation, the projection of nonbranching structure is impossible, leaving this classic distributional difference between arguments and adjuncts without an analysis.

In Harley (2005c) I show how the proposal of an acategorial root node in DM can resolve this problem for Bare Phrase structure, on the assumption that roots

themselves select for arguments and project a √P constituent. This is suggested already by the fact that deverbal nominals have the same argument-selectional properties as their verbal constituents:

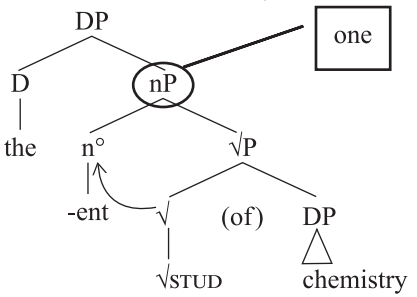
(21) a. John studied physics
b. John is a student of physics

If both verbal *study* and nominal *student* share the same root (realized as *stud-*), and if the semantic interpretive properties of that root are responsible for imposing selectional restrictions on its sister DP, the identical argument-selection properties of the related noun and verb can be captured at the root level, below n° or v°[20]. This makes sense, in that encyclopedic truth-conditional content is associated with root interpretation.

If the argument *of physics* is the sister of √, which projects to √P, and the resulting complex structure is nominalized by the addition of an n° (here realized by *-ent*), it becomes very easy to characterize *one*-replacement: *one* is an nP anaphor, not a √P anaphor.[21] All we need to complete the picture is to assume that adjunct PPs adjoin to nP, not √P, and the distribution of *one*-replacement is transparently derived, in exactly the spirit of Jackendoff's original proposal.

The structures of *student of chemistry* and *student with long hair*, under this analysis, are illustrated in (22) and (23) below. The nPs in each structure which can serve as potential antecedents for *one*-replacement are circled. Notice the different structural positions of the argument PP *of chemistry* and the adjunct PP *with long hair*:

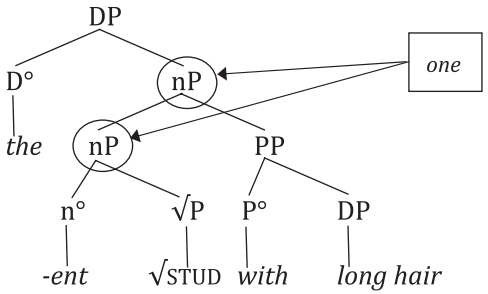
(22) The student of chemistry



20 Though cf. Panagiotidis (2005).

21 In the same way, *do* is a vP anaphor, as proposed by Merchant (2008).

(23) The student with long hair
(compare #*He studies with long hair*)



An argument in favor of the notion that *with long hair* is adjoined to nP rather than to √P is the fact that this modifier produces a distinctly odd stage-level depictive reading in the verbal context: *He studies chemistry with long hair*. This difference is captured, on this analysis, by the fact that nPs are predicates of entities, while vPs are predicates of events (see, e.g. Pylkkänen (2002, 2008) for discusssion); constituents which are appropriate modifiers of nPs, then, may not be appropriate modifiers of vPs.

To summarize: If *one* is of category nP, then we expect that nP modifiers can attach to it (predicting the grammaticality of *one with long hair*), but we expect that it cannot select argument PPs, as only roots can do that. On this account, the nonbranching projection problem for BPS posed by *one*-replacement is resolved.

The reason that this argument is relevant to the current discussion is that the whole proposal is predicated on the notion that internal arguments are sisters of root nodes, not sisters of nP or vP. Insofar as the analysis provides a successful resolution of an empirical problem for Bare Phrase Structure theory, then, it consistutes an argument in favor of the notion that root nodes can select for sister constituents, and subsequently project as the head of a phrasal category, just like a run-of-the-mill syntactic terminal node.[22]

22 A reviewer rightly points out that the more functional projections one assumes within the NP/DP domain, the more options exist for saving the *one*-replacement analysis without recourse to an acategorial root. On a cartographic approach to DP, for example, one could assume that *one*-replacement targets a relatively high node in the hierarchy, say NumP, that PP adjuncts to N′ in Jackendoff's analysis are in fact modifiers of NumP and that argument PP sisters to N are sisters to NP, rather than to an acategorial root (or any other relatively low functional projection). However, the fact that selectional restrictions remain in force across the nominal/verbal divide (*study chemistry/student of chemistry*) suggests that whatever low category is sister to the internal

In the next section, I recap a proposal from Kratzer (1994, 1996) which I contend has the same consequence as the *one*-replacement analysis above: Roots and their objects must be sisters, undergoing Merge directly and projecting to a √P constiutent.

3.2 Semantic evidence: Verb-object idioms (Kratzer 1996)

In the present model, the LF interpretations contributed by √ nodes provide Encyclopedic, truth conditions, whose evaluation draws on extralinguistic cognitive resources outside the linguistic system. Access to these interpretations, as discussed in section 2.4 above, can be contingent upon the category and content of other nodes in the local syntactic environment.

In particular, it seems clear that certain configurations of √s and other constituents are susceptible to the development of contingent truth conditions – i.e., susceptible to idiomatization – while other configurations are not subject to this tendency. Marantz (1984) observes that, while object-verb combinations frequently receive idiomatic interpretations while composing freely with their subject, subject-verb combinations rarely do so while composing freely with their object.[23] Indeed, these special interpretations are not restricted to idioms *per se*, but can arise whenever the denotation of the object has a particular semantic property, e.g. when the object denotes a beverage (*kill the beer/wine/soda*), or a time span (*kill an hour/day/evening*). Below, one familiar set of examples from (Marantz 1984) are repeated, and an additional set involving *pass DP* are provided to illustrate the point.

argument is not specific to the nominal extended projection. The acategorial root meets this description perfectly. See Punske and Schildmier Stone (2013) for discussion of idiomatic interpretations in deverbal nominals.

23 Nunberg et al. (1994) contend that this tendency is explicable as a conspiracy of independent factors, involving conceptual difficulty in ascribing abstract or metaphorical interpretations to DPs referring to animate entities and the tendency for animate entities to occur in subject position, and adduce a few counterexamples to Marantz's claim. Horvath and Siloni (2002) build the case against the pattern further, producing an additional class of counterexamples from different languages. Harley and Stone (2013) argue that the interpretations of the counterexamples in fact involve experiencer predicates, in which the purported idiomatic agent argument must be base-generated VP-internally, and so do not count as true counterexamples. Here, we set this debate aside for the moment, and take Marantz's generalization to be a true characterization of a constraint on special interpretations.

(24) a. kill a bug "cause the bug to croak"
kill a conversation "cause the conversation to end"
kill an evening "while away the time span of the evening"
kill a bottle "empty the bottle"
kill an audience "entertain the audience to an extreme degree"
b. pass judgement "evaluate"
pass thirty "get older than thirty"
pass a law "enact legislation"
pass a test "meet a standard of evaluaton"
pass a kidney stone "excrete a kidney stone"
pass the hat "solicit contributions"

Kratzer (1994, 1996) takes up the problem of explaining why verbs' truth conditions should be so frequently sensitive to the semantic content of their objects, but should be effectively indifferent to the content of their external arguments. In a model-theoretic approach in which a transitive verb directly composes with both its internal and external arguments, there is no technical barrier to imposing a constraint on a verb's truth conditions which depends on the content of the external argument, in the same way that it is clearly possible to do with internal arguments.

The idea is that a predicate can specify a particular set of truth conditions to employ if one or more of the predicate's arguments meets certain criteria. For example, the different meanings in (24) could arise if *kill* imposes a disjunctive set of truth conditions along the following lines:

(25) $[\![kill(y_{obj})(x_{subj})]\!] = 1$ iff
y is a period of time and period of time is over
y is a consumable and consumable is fully consumed
y is . . . and . . .

If this is the correct approach to special meanings for particular verb-object combinations, however, the apparently categorical absence of special meanings for particular subject-verb combinations becomes mysterious. If *x*, an external argument, composes directly with *kill*, the truth conditions of *kill* could just as easily be contingent upon the identity of *x*, instead. There must be some principled reason why it seems to be impossible to specify particular truth-conditions based on the content of the agentive subject.

Kratzer's proposal is to, in her words 'sever the external argument from the verb'. In fact, she concludes, the verb does not compose with its external argument at all. She argues for a (semi) neo-Davidsonian approach, in which transi-

tive predicates like *kill* in fact select for only one DP argument. Their external arguments are introduced into the derivation, and assigned their Agent interpretation, by a separate predicate entirely, the Voice head. This predicate and its argument are conjoined with the verb and its internal arguments by a special composition operation entitled Event Identification. Since the verb itself does not compose with an external argument, but only with internal ones, the truth conditions contributed by the verb can only be conditioned by the content of their internal argument, not by the content of the external argument.

The type of truth conditions which are at issue are Encyclopedic ones, that is, the truth conditions introduced by the interpretation of a root node. In the DM framework, then, the analogue to Kratzer's lexical V projection is √. The choice of disjunctive truth conditions is determined by the root when it composes directly with its object DP. Kratzer's proposal requires that roots, as the introducers of idiosyncratic truth conditions, compose by function application with their object arguments. The analysis is, I think, not compatible with the idea that objects are introduced by a separate verbal functional head, nor with the notion that roots do not compose directly with their internal arguments. Roots, or more precisely, the interpretations introduced by roots, must have an argument structure – an argument structure which includes the internal argument, but not the external one.[24]

3.3 Morphological Ergative Splits, Case and Agreement

A final suggestive piece of evidence indicating the close interaction of roots and their complements involves the triggering environment for root suppletion in languages like Hiaki, where the number of one of the arguments of the verb conditions the choice of suppletive root. This suppletive form of agreement, in Hiaki and all the other Uto-Aztecan languages with suppletive verbs, follows an ergative-absolutive distribution: Intransitive suppletive verbs are conditioned by the number of their subject argument (their only argument), while transitive

24 It remains an open question whether there are syntactic argument structures as well as semantic ones. That is, can a root bear a feature which requires that it be syntactically Merged with an argument DP, as well as introduce a function which seeks to compose with such an argument? The differing abilities that transitive predicates have to undergo object drop (the difference between *John ate* and *#John patted*, for example) is potentially relevant here, but a full consideration of these issues will have to wait for a future occasion.

suppletive verbs are conditioned by the number of their object argument. This is illustrated by the examples in (26) and (27) below:

(26) Hiaki verb suppletion: Intransitives controlled by subject number:
a. Aapo weye
3**sg** walk.**sg**
'He/she/it is walking.'
b. Vempo kaate
3**pl** walk.**pl**
'They are walking.'

(27) Hiaki verb suppletion: Transitives controlled by object number:
a. Aapo/Vempo uka koowi-ta mea-k
3sg/3pl the.**sg** pig-**ACC.sg** **kill.sg**-PRF
'He/They killed the pig.'
b. Aapo/Vempo ume kowi-m sua-k
3sg/3pl the.**pl** pig-**pl** **kill.pl**-PRF
'He/They killed the pigs.'

This pattern, if it represents true verbal agreement, poses a serious challenge to an otherwise robust typological generalization concerning agreement, described by Bobaljik (2008): If a verb agrees with just one argument in the clause, it is the argument bearing morphologically unmarked case.

In order to undersand this generalization, and why the Hiaki agreement pattern represents a challenge to it, we will briefly review the theory of Dependent Case advanced by Marantz (1991), in which the notion of 'unmarked case' is defined.

Languages typically exhibit one of two morphological case-marking patterns, if any: Nom-Acc, in which the subjects of intransitive verbs receive Nominative case, the same as the subjects of transitive verbs, or Erg-Abs, in which the subjects of intransitive verbs receive Absolutive case, the same as the objects of transitive verbs. (We set aside the more complex cases of split and mixed Case systems for ease of exposition here, though of course their relevance is not disputed.) Marantz proposed to account for this split in the morphological component. In the syntax, in both types of languages, DPs are case-licensed with either theta-dependent lexical case features, or by checking a structural case feature against a structural case-assigning head. In the morphological component, these structurally case-marked DPs' case features are subsequently spelled out as m-case marking – morphological case.

In Marantz's theory, languages have an unmarked m-case form, and a dependent m-case form. Unmarked m-case is used to realize the structural case fea-

ture of the single DP in an intransitive clause. In a transitive clause, Unmarked m-case will realize one of the structural case features present, and Dependent m-case will realize the other.[25] The difference between Nom/Acc systems and Erg/Abs systems is simply the locus of realization of Dependent case. In Nom/Acc systems, the Dependent case (Acc) is assigned to the object of transitive clauses, while in Erg/Abs systems, the Dependent case (Erg) is assigned to the subject.

Bobaljik 2008 points out that this provides a very straightforward characterization of the typological generalization concerning the relationship between case and agreement: Agreement, when present, depends on the argument bearing unmarked m-case.[26] In Nom/Acc languages, Nominative case is unmarked, and agreement is always with the nominative argument. In Erg/Abs languages, Absolutive case is unmarked, and agreement is always with the absolutive argument.[27]

One classic example illustrating the relevance of m-case, rather than syntactic position, in determining agreement is provided by Icelandic Dat-Nom constructions. In these constructions, the subject is marked with dative case, and the object bears nominative. One verb exhibiting this pattern is *líka*, 'like'; the NP bearing the role of 'liker' is marked with dative case, while the liked item is nominative. We can tell that the dative argument is the true subject of the construction because it must be realized as PRO in infinitive clauses, as in (28a), a property of subjects. (Note that if there were a stranded participle modifying the PRO argument it would agree with the null subject in exhibiting dative case, confirming that the null argument here bears morphological dative.) Although Icelandic agreement is usually with the subject argument, this is not the case with verbs like *líka* that take dative subjects. Here, agreement is with the nominative object argument, rather than the dative subject argument (28b, c). The point is that when the subject grammatical function and the unmarked nominative case

25 The basic idea is similar to the 'Case in Tiers' proposal of (Yip et al. 1987).

26 Note that it may *also* vary with the dependent-case argument, in systems where both subject and object agreement are marked, but the claim is that it is *at least* sensitive to the unmarked-case argument. P. Svenonius (p.c.) notes that systems in which agreement appears to track grammatical function, rather than *m*-case, do exist, though apparently rarely (Nepali and Burushaski are two such cases, discussed by Bobaljik 2008 and Baker 2010 respectively).

27 As noted by Bobaljik, this robust typological generalization runs counter to (Moravcsik 1974)'s agreement hierarchy, according to which agreement is characterized as tracking grammatical functions according to the usual hierarchy of subject > object > indirect object. In Ergative/Absolutive languages, agreement tracks the absolutive argument, even when the absolutive argument is the object of the verb, in transitive clauses.

diverge, agreement tracks the argument bearing unmarked case, rather than the argument bearing the grammatical function 'subject'.[28]

(28) a. Jón vonast til [að __ líka þessi bók]
Jon.nom hopes for [to PRO.dat like.inf this.nom book.nom
'Jon hopes to like this book.'
b. *Morgum studentum líka verkið
many students.dat like.3pl job.nom
c. Henni líkuðu þeir
her.dat like.pst.3pl they.nom
'She liked them'

With this theory of agreement and m-case in mind, let us revisit the Hiaki data presented in (26) and (27). Suppletive verb agreement is clearly tracking the subject of intransitive clauses and the object of transitive clauses – an Erg/Abs pattern. But Hiaki is not an Erg/Abs language.

Case-marking in Hiaki is very straightforwardly Nom/Acc, as illustrated in (29). Objects of transitive verbs are marked with accusative case, which is clearly structural in character, as it becomes nominative under passivization:

(29) Hiaki Case: Nom/Acc
a. Hoan Maria-ta vicha-k
Juan.nom Maria-acc see-prf
"Juan saw Maria"
b. Maria aman vicha-wa-k
Maria.nom there see-pass-prf
"Maria was seen there"

In Hiaki, then, Acc is the Dependent case, Nom the Unmarked case. Agreement, according to Bobaljik's typological generalization, should necessarily track the nominative argument. But transitive suppletive verbs agree with their accusative object, not their nominative subject, as shown in (27). If number-conditioned

28 Note that 'unmarked' here refers to the morphological category Nominative, which is unmarked in the sense of not being dependent on the realization of another case, in Marantz's system. 'Unmarked' is not intended here in the the morphophonological sense; the nominative case has both overtly marked and unmarked (zero) morphophonological allomorphic realizations.

suppletion is true Agreement, then Hiaki represents a counterexample to the typological generalization.[29]

What is the relationship between a suppletive root and the argument which conditions its insertion? Let us consider the hypothesis that it is not agreement, per se. Instead, it is simply context-conditioned root Vocabulary Item competition, as outlined in section 2.4 above. Why should transitive verbs be conditioned by object number, rather than subject number, as for the intransitive verbs?

One hypothesis has to do with the idea that vocabulary insertion is subject to a locality restriction. At the point when the phonological exponents of these roots are inserted, the local environment contains an internal argument, marked for number. Only elements in this local environment can condition a choice of root allomorph.[30]

This proposal concerning locality of conditioning makes a prediction: the intransitive suppletive verbs of Hiaki should be unaccusative. Their conditioning argument, although it ends up as a surface subject, must be base-generated in the immediately local environment, in object position, to trigger the insertion of the appropriate suppletive allomorph of the verb root.

In fact, language-internal evidence suggests that the intransitive suppletive verbs are indeed unaccusative. One test which indicates this is the inability of these verbs to combine with an applicative, as argued in Haugen et al. (2009).

Hiaki has a very productive applicative construction, which usually has a benefactive reading. It corresponds to a 'high' applicative in the terminology of Pylkkänen (2002, 2008) since it can apply to intransitive unergative verbs as well as to transitive verbs.[31]

---

29 Svenonius (p.c.) notes that the Hiaki pattern, if it constituted true Agreement, would also violate another typological generalization: Although 'split' systems exhibiting nom-acc agreement patterns with erg-abs case marking patterns are attested (and are potentially problematic for Bobaljik's generalization), the reverse – erg-abs agreement with nom-acc case marking – are not (Anderson 1977, Comrie 1978, Moravcsik 1978, as described in Woolford 2006). The conclusion here, that the Hiaki pattern does not constitute true Agreement, thus is consistent with that typological claim as well.

30 Indeed, if phase theory (Chomsky 1995, 1999) is correct, only internal arguments *could* be present in the immediately local environment of the verb root at Spell-Out, since external arguments are generated in a separate phase.

31 The applicative is formed by suffixing *-ria* to the verb, and introduces a benefactee argument. The benefactee, which must be animate, is marked with accusative case and c-commands any other internal arguments. The applicative argument, and not the erstwhile direct object, becomes the subject under passivization and can bind an anaphoric object of the verb, as shown for Hiaki in Rude (1996).

(30) a. U'u maaso uusi-m yi'i-ria-k
The deer.dancer children-pl dance-APPL-PRF
"The deer dancer danced for the children."
b. Inepo Hose-ta pueta-ta eta-ria-k
1sg Jose-ACC door-ACC close-APPL-PRF
"I closed the door for Jose"

The applicative cannot, however, co-occur with run-of-the-mill unaccusative intransitive verbs, as shown in (31)

(31) *Uu tasa Maria-ta hamte-ria-k
The cup.nom Maria-ACC break.intr-APPL-PRF
"The cup broke for/on Maria"

We see, then that unaccusative verbs are incompatible with an Applicative head, probably because the semantics of the Applicative require it to compose with a causative/agentive v°, and it cannot compose with the Agentless unaccusative v°. It is well-formed when attached to an unergative intransitive like *bwiika* 'sing', however. The applicative is thus a test for unergativity, since it can only apply to intransitive verbs whose subjects are intentional and agentive

Crucially, the applicative cannot apply to any of the suppletive intransitive verbs, even though the meaning of some of them seems to be fairly agentive, judging from their English translation equivalents (e.g. *vuite* ~ *tenne* 'run'; *weye* ~ *kate* 'walk'). The incompatibility of the applicative suffix with suppletive intransitive verbs is illustrated in (32a) below. Instead, to express an applicative meaning with a suppletive intransitive, a Hiaki speaker uses the periphrastic construction with the postposition *vechi'ivo*, 'for', as shown in (32b), which is compatible with verbs of all classes.[32]

(32) a. *Santos Maria-ta San Xavierle-u weye-ria
Santos Maria-ACC San Xavier-to go-APPL
"Santos is going/walking to San Xavier for Maria"
(e.g. carrying out a vow she had made for a pilgrimage)
b. Santos Maria-ta vetchi'ivo San Xavierle-u weye
Santos Maria-ACC for San Xavier-to go.
"Santos is going/walking to San Xavier for Maria"

32 Note that adding a Benefactee argument periphrastically is otherwise usually interchangeable with the applicative – when both are possible our consultants feel them to be synonymous. The activity described by the suppletive verb weye 'walk' is thus semantically compatible with a benefactive semantics.

This is a general property of all the suppletive intransitive verbs. All the verbs listed in (33) are ungrammatical with *-ria,* and all but one (*vo'e ~ to'e*) are compatible with *vechi'ivo* PPs instead:[33]

(33) a. vuite ~ tenne 'run.sg ~ run.pl'
b. siika ~ saka 'go.sg ~ go.pl'
c. weama ~ rehte 'wander.sg ~ wander.pl'
d. kivake ~ kiime 'enter.sg ~ enter.pl'
e. vo'e ~ to'e 'lie.sg ~ lie.pl'

Despite the agentive translations of some of these (*run, wander*), it is plausible on semantic grounds to consider these good candidates for unaccusativity, as they are all verbs of body posture or directed motion. This semantic class exhibits unaccusative behavior in some Indo-European languages (see, e.g., Hoekstra & Mulder (1990) on Dutch), and cross-linguistically exhibit special morphological behavior that distinguishes them from non-motion intransitive activity verbs.[34]

We thus conclude that suppletive verbs, whether transitive or intransitive, agree in number with elements generated as their complement – their 'deep objects' – regardless of their surface position. (This was anticipated in the conditioning context provided for suppletive root insertion in (14) above. This conclusion is already prefigured in Baker's (1985) discussion of the same phenomenon in the related Uto-Aztecan language Huichol.)

This, then, is not a real Agreement operation, which depends on case-marking and would not distinguish unergative and unaccusative intransitives. Rather, it reflects root competition conditioned by the local environment at the

33 Note: the problem with *-ria*-affixation is not about suppletion, per se. It is fine to add an applicative affix to suppletive transitive verbs, such as *mea~sua* 'kill':

i) Santos Hose-ta koowi-ta/koowi-m mea/sua-ria-k.
Santos Jose-ACC pig-ACC/pig-PL kill.sg/kill.pl-APPL-PRF
"Santos killed a pig/pigs for Jose."

It is also worth noting that although a new object argument, *Jose-ta*, has been added to the clause, verbal suppletion still depends on the number of the verb's thematic object, rather than the structural object introduced by the applicative, again suggesting that suppletive agreement is not structurally implemented.

34 (Guerrero 2004) within the context of Role and Reference grammar, argues on semantic grounds that these intransitive Hiaki verbs all assign a single Undergoer thematic role, rather than an Agent thematic role. This translates naturally within the present syntacticocentric framework to an unaccusative analysis for these verbs, since unaccusative status is importantly connected to the lexical semantics of the verbs involved Levin & Rappaport Hovav (1995).

point at which roots are inserted. This is consistent with a cyclic, bottom-up approach to vocabulary insertion (Bobaljik 2000) and strong locality conditions on spell-out domains. The strongest and most interesting hypothesis concerning the relevant locality condition is that the triggering DP is base-generated in a maximally local configuration with the suppletive root, i.e. as its sister. If this is the case, roots must take complements.[35] See Harley et al. (to appear), Bobaljik and Harley (to appear) for further discussion.

True agentive external arguments are never in such a local relationship with the verb root, and hence it would be surprising if they could trigger suppletion there. The account thus predicts that there should be no suppletive unergative verbs conditioned by subject number. This is certainly true for Hiaki; whether it is true for all languages exhibiting argument-conditioned verb suppletion remains, of course, an open empirical question. In any case, in Hiaki, it is clear that roots have a special relationship with their selected internal arguments. This, taken together with the other arguments for sisterhood of root and direct object presented in sections 3.1 and 3.2 above, suggests that roots do indeed merge directly with argument DPs in the syntax, and thereby project a √P.

Given the observation that the immediately local environment of a root can play a significant role in its phonological and semantic interpretation, we can now turn to our last question: Is the immediately local environment the *only* environment which can play such a conditioning role? A strong form of the hypothesis about locality constraints on root interpretation was suggested by Marantz (2001, 2008) and Arad (2003), who propose that the first categorizing node is a phase boundary. Since a given operation of Spell-Out cannot be affected by elements outside its phase edge, this effectively limits the domain that can condition special PF or LF interpretations for a given root to material within the first categorizing node. I argue that such a stringent locality condition is too strong, at least with respect to domains of idiomatic interpretation.

4 Locality domain for interpretation: Categorizing heads? Or VoiceP?

The discussion of idiomatic interpretations in section 3.2 above drew on the empirical observations from Marantz (1984) concerning the apparently special

35 Indeed, given the theory proposed in Marantz (2001, 2008) and Arad (2003), according to which the first categorizing head is a phase boundary, it would be impossible for root suppletion to be triggered by an argument generated in any other position than sister to the root node.

status of the external argument with respect to idiomatization. Kratzer (1994, 1996) established a semantic role for VoiceP as the external-argument-introducing head which provided an account of the special status of external arguments with respect to idiomatization, as reviewed above.

In the theoretical landscape of the late 1990s, the external-argument introducing head, called VoiceP by Kratzer, was labelled vP by Chomsky (1995), and identified with the external-argument introducer and causativizing V of Hale & Keyser (1993). In Harley (1995) and Marantz (1997), the additional, DM-specific connection between this external argument-introducing projection and the verb-creating categorizing head was laid out. In addition to introducing the external argument and defining a domain for idiomatic interpretation, then, the v° head furthermore created verbs from roots. The lower VP of Kratzer 1994, 1996 was identified with DM's √P headed by an uncategorized verb root. The vP, in composing with the √P, performed all three functions: it introduced the external argument, categorized the √, and provided a domain for special interpretation.

Subsequent work by Pylkkänen (2002), however, argued that the first two of these functions must be separated: VoiceP, in which the external argument is introduced, is distinct from verb-forming vP, below Voice. √P is lower still, the complement of v°. That is, the category-creating head and the external-argument-introducing heads are distinct.[36] Arguments to this effect are also given in Marantz (2001) and Doron (2003); a version of Marantz's argument is developed in detail in Harley (2007, 2013).

The question then arises as to whether the third function – defining a rigid domain limiting the potential for special interpretations – is properly linked to the first categorizing head. That is, are special interpretations limited to conditioning within categorizing vP, or can they extend up to include the external-argument introducing head, i.e. VoiceP? Couched in Minimalist syntactic terms, we can ask wheter it is vP or VoiceP which constitutes a phase boundary.

Marantz (2001) and Arad (2003) argue that the interpretive cycle occurs at the first categorizing node – that is, that categorizing nodes are syntactic phase heads, triggering Spell-Out and assigning phonological and semantic interpretations for the constituent dominated by the phase head. Root interpretation, then, is fixed with respect to that first categorizing node. Further derivation, outside the first categorizer, must then build on the interpretation defined at the first phase.

36 Or at least, may be distinct – see (Coon & Preminger n.d.) for a recent argument that both functions are indeed unified in a single head in Chol. Pylkkänen (2002) proposed a "Voice-bundling" parameter, according to which Voice may be unified with v in some languages, and distinct from it in others.

After the first phase, in other words, root interpretation is fixed and must figure compositionally in subsequent levels of derivation.

Arad 2003: 746 illustrates the prediction made by this claim with data from Hebrew. The different word-forming binyanim are analyzed by both Arad and Doron (2003) as realizations of categorizing heads, v°, n° and a°. As we have seen above in section 2.3, words derived when a binyan combines with a triconsonantal root exhibit great semantic variability. In contrast, Arad claims that verbs derived from applying a verbal binyan to an already-categorized noun (itself derived by combining an n° template with a triconsonantal root), have compositional semantics which must include the meaning established at the nP cycle.

(34) Root-derived words from √*sgr* exhibiting a range of idiosyncratic interpretations

a.	CaCaC (v)	sagar	v, 'close'
b.	hiCCiC (v)	hisgir	v, 'extradite'
c.	hitCaCCeC (v)	histager	v, 'cocoon oneself'
d.	CeCeC (n)	seger	n, 'closure'
e.	CoCCayim (n)	sograyim	n, 'parentheses'
f.	miCCeCet (n)	*misgeret*	n, 'frame'

(35) Noun-derived verb from (34)f, *misgeret*, n, 'frame'

CiCCeC	*misger*	'to frame'

The fact that *misger*, 'to frame', is derived from the noun *misgeret* is shown by the fact that the nominal augment *mi-* from the nominalizing template in (34)f is contained within the verbal form. The fact that the nP is contained within the verb *misger* also explains why the nominal semantics is contained within it as well: the meaning of the verb is built up from the meanings of its parts, including the meaning of the nP.

A parallel argument is given for English by Marantz (2001: 17). He points out that the meanings of root-drived *rot-or* and *don-or* are relatively idiosyncratic in character compared to the meanings of verb-derived *rotat-or* and *donat-or*.[37]

It is unsurprising that a complex constituent contained within a larger constituent can contribute its meaning to the meaning of the whole; that is standard

37 Although it is worth noting, as Marantz does, that while one would speak of a *blood donor* not a #*blood donator*, one equally refers to a *rotator cuff*, not a *#rotor cuff*. Yet both *blood donor* and *rotator cuff* strike me as involving the same amount of 'listedness' in their interpretations.

compositionality. The question at hand, however, is whether these particular subconstituents *must* do so. That is, is interpretation above the first categorizing head necessarily compositional in character? If categorizing heads are phases – domains at which interpretations are fixed with respect to all subsequent computation – then they must be.

With Borer (2009), I contend that the evidence of layered derivational affixes does not suggest a clear dividing line between productive, regular, compositional interpretation outside the first categorizing affix, compared to irregular, idiosyncratic, idiomatic interpretation within it. Obviously the interpretation assigned at the level of the first categorizing affix will be idiosyncratic, as the root never occurs without such superstructure, and cannot be interpreted in its absence. However, it seems clear that idiosyncratic semantics can also be assigned outside the first categorizer heads, on later cycles of derivation. Below I list a number of examples in which multiply derived words exhibit new senses that to me seem to lack the predicted compositional contribution of content from their inner constituents (underlined in the examples below). Indeed, in some cases, the compositional contribution of the contained substructure seems in fact to be unavailable, see particularly (36)c, e.

(36) a.	*edit*	*edit-or*	*editor-ial*	
			‘of or relating to the editor’	*compositional*
			‘opinion article’	*idiosyncratic*
b.	*nature*	*natur-al*	*natural-ized*	
			‘made natural’	*compositional*
			‘became a citizen by residing in a country’	*idiosyncratic*
c.	*class*	*class-ify*	*classifi-eds*	
			#‘things which have been classified’	*#compositional*
			‘small newpaper advertisements’	*idiosyncratic*
d.	*nation*	*nation-al*	*national-ize*	
			‘make national’	*compositional*
		Not antonym: *private)*	‘government takeover of business’ (Antonym: *privatize*)	*idiosyncratic*
e.	√*domin*	*domin-ate*	*dominat-rix*	
			#‘woman who dominates’	*#compositional*
			‘woman who performs ritualized sexual domination’	*idiosyncratic*

f.	*institute*	*institut-ion*	*institution-al*	*institutional-ize*	
				'make institutional'	*compositional*
				'commit someone to a care facility'	*idiosyncratic*

Some multiply affixed words seem particularly idiosyncratic in character, lacking a compositional reading altogether (like *dominatrix* and *classifieds*), although the stem for the final affix is clearly itself already clearly a categorized and independently meaningful word. Consider also *universe* ~ *university* (compare *universality*); *hospital* ~ *hospitality*, *sanitary* ~ *sanitarium,* and *auditory* ~ *auditorium*: in none of these cases do the entailments of the inner derived word contribute compositionally to the meaning of the outer one. Other cases are not hard to come by, though further discussion might perhaps be warranted; does *conserve* contribute its content compositionally to *conservation* (or *conservative*)? In the triad *relate* ~ *relation* ~ *relationship*, the idiosyncratic meaning of *relation* does not seem to contribute compositionally to the most salient meaning of *relationship*; indeed, one is presumably not likely to enter into a relationship (on its idiosyncratic meaning) with one's relations. Similarly, a *protectorate* is not just any old entity which has a *protector*, and the relationship between *economic* and *economic-al* is also a little tough to understand compositionally.

That is not to say that derived words cannot be interpreted compositionally. As noted in Marantz (1995a) *transmission* has both an idiomatic and a compositional reading, just like phrasal idioms such as *kick the bucket* do. It seems to me to be one of the strongest arguments for the syntax-all-the-way-down hypothesis, that semantic idiosyncracy crops up at both the phrase and word level in more or less the same continuum of variability. Furthermore, Marantz (1997) argued that extensive internal structure matters at the word level just as at the sentence level: *blick* can't mean what *nationalization* can mean (though see Marantz 2013 for an argument that the reverse is not true).

The slippery and gradient judgements concerning differences in compositionality between cases like *rotor* vs *rotator*, to my mind, are quite distinct from the classic examples of 'inner' vs 'outer' derivational morphology with which we are familiar from the past 15 years of research on the topic (or, indeed, 30–35 years, considering that Shigeru Miyagawa extensively documented this very point for Japanese lexical vs syntactic causative constructions in the early 1980s (Miyagawa 1984) and Tom Wasow did the same for stative vs. eventive English passives in 1977 (Wasow 1977). The distinction between 'inner' vs. 'outer', 'lexical' vs. 'productive', occurences of the very same affixes keeps appearing robustly in

language after language, and the high/low affixation analysis incorporating the concept of an Elsewhere form pioneered by Miyagawa (1994, 1998) strikes me as one of the great insights of the syntacticocentric approach to morphological analysis. It has been independently discovered and productively employed again and again: Kratzer (1994, 1996) for *of*-ing and *acc*-ing forms in English Marantz (1997) for Dubinsky and Simango's (1996) Chichewa statives and passives, Travis (2000) for Malagasy lexical and syntactic causatives, Sugioka (2001, 2002) on Japanese nominalizations, Embick (2003, 2004) for stative, resultative, and passive participles in English, Fortin (2004) for Minnangkabu causatives, Svenonius (2004) for lexical and superlexical prefixes in Slavic, Jackson (2005) for statives and resultatives in Pima, Alexiadou & Anagnostopoulou (2008) for adjectival and verbal participles in Greek, Svenonius (2005) for high/low treatments for causatives in several languages, Killimangalam & Michaels (2006) on causatives in Malayalam, Serratos (2008) on causatives in Chemehuevi, and doubtless others I am unaware of.

The question is not whether the inner vs. outer insight is correct; it seems incontrovertible that it is. The question is what kind of constituent demarcates the boundary between 'inner' attachment and 'outer' attachment. Is it, as Marantz (2001) proposes, the first categorizing head? Or is it instead, as Marantz (1997) proposes, whatever head is responsible for introducing the external argument into the semantic and syntactic derivation? Or is it some third domain-creating functional projection which is crucial in introducing eventiveness into the derivation?

If Marantz (1997) was correct, and it is in fact the external-argument-introducing head which delimits the domain for special interpretations, then his own generalization concerning the exclusion of true external arguments from idiomatic interpretations from Marantz 1984 falls into place as another reflection of the interpretive boundary between the 'inner' and 'outer' domains. Hale & Keyser's (1993, 2002) vision of 'l-syntax' involved a limit imposed by the introduction of an agent-introducing head; so too does Ramchand's (2008) framework of First Phase Syntax. Because it is the external-argument head that demarcates the phasal domain, it is morphology that references external arguments that exhibits compositional, high-attachment behavior: syntactic causatives, eventive ('verbal') passives and participles, eventive nominalizations, *-able* formations, and so on. Voice is the phase head, not v.

This view allows for the occurrence of genuinely idiomatically interpreted phrasal constituents in languages like Persian, in which meanings which would translate as simple verbs in English must be represented by a complex predicate construction involving at least two fully categorized heads (see, among many others, Folli et al. 2005). It also allows for the existence of *caboodle* items, which

are clear cases of categorized roots whose meanings are wholly dependent on occurrence in a bigger conditioning context.

Marantz (1995b: 10–11) put the case very clearly:

> Constructions in English with "do" "take" "give" and other light verbs have the semantics of single verbs and call into question the notion that the phonological word is the distinguished locus of idosyncratic meaning.
>
> (5) a. Take a leap
> b. Take a leak
> c. Take a piss
> d. Take a break
>
> Although light verb constructions and idioms show that the domain of specialized meanings is not the phonological word, there do seem to be locality constaints on the contexual determination of specialized meaning. Note that in light verb constructions/idioms with "make," for example, a lower verb cannot be agentive.
>
> (6) a. Make X ready
> b. Make X over
> c. Make ends meet
> d. *Make X swim/fly a kite/etc.
> (only pure causative meaning on top of independent reading of lower VP)
> e. Marie a laissé tomber Luc.[38]
> Marie has let fall Luc
> 'Marie dropped Luc like a hot potato', Lit "Marie let Luc fall"
> f. On lui fera passer le goût du pain.
> One to.him will.make pass the taste of bread
> 'They'll kill him', Lit. 'They'll make the taste of bread pass from him'.
> g. *Marie a laissé/fait V (NP) (à) NP*
> Marie has let/made . . .
> with special meaning of "V" that is not available outside the causative construction and where NP* is an agent

What, then, is the status of the observation that root-derived words are more idiosyncratic in character than word-derived words? It is one of degree, not kind; idiosyncratic noncompositionality just becomes less frequent the more structure is involved. As noted above, the first combination of a root with a categorizer will *have* to be 'idiosyncratic'; roots don't occur in isolation, so all root meanings will have to be context-dependent. The main point is that interpretations of derivations even after the first categorizer *can* still be idiosyncratic, not necessarily containing the meaning specified at the first categorizer as a proper subpart – as long as the conditioning environment for the idiosyncratic interpretation (the

38 Marantz's French examples are from Ruwet (1991); I have added the gloss lines.

en-search domain, in Borer's (2009) terms) doesn't extend beyond the real first phase head – VoiceP, if the discussion above is on the right track.

5 Conclusion

The points I have tried to establish in the above discussion have focussed on three topics. First, what are roots, and what are they like? In the Distributed Morphology model, this needs to be addresssed in at least three domains, corresponding to the three types of lexicon-like listings in the model.

List 1 'roots' are root terminal nodes, manipulated by the syntax; I have argued that they are not underspecified, but rather must be individuated even in the narrow syntax. Their individuation cannot be semantic or phonological in character, however; I adopt Pfau's and Acquaviva's index notation to indicate the distinctions between roots in List 1.

List 2 'roots' are the phonological exponents which compete to realize particular root terminal nodes provided to the syntactic derivation by List 1. These exponents can compete with each other for insertion into appropriate positions, like other Vocabulary Items from List 2, and this competition can be conditioned by the content and structure of the local syntactic environment.

List 3 'roots' are interpretations, instructions for the interpretation of particular root terminal nodes provided to the syntactic derivation by List 1. These interpretations can also be conditioned by the content and structure of the local syntactic environment; it is such conditioning which creates idiomatic interpretations and allows for the existence of *caboodle* items.

The second question addressed above involved syntactic behavior of root terminal nodes from List 1. Do such terminal nodes behave like other syntactic feature bundles drawn from the Numeration, once introduced into the syntax? In particular, can they undergo Merge with phrasal constituents and themselves project? Based on a particular analysis of the distribution of *one*-replacement in argument structure nominals, it is argued that roots can indeed combine with internal arguments directly, without the need for mediation by a functional category of any kind. Circumstantial evidence from an analysis of special internal-argument conditioned meanings (verb-object idioms) and internal-argument conditioned pronunciations (suppletive forms of Hiaki verb roots) was taken to bolster this position.

Finally, the debate concerning the syntactic identity of a demarcating domain for special interpretation was reviewed. Is the domain of idiosyncratic interpretations for a given root restricted to the first categorizing node above a given root? Or can the conditioning environment of idiosyncracy involve structures

outside this domain? After a review of the arguments and evidence presented in favor of both positions, the boundary domain – the first phase head – is identified as VoiceP, not nP, aP, or vP.

The conclusions here obviously cry out for further refinement and testing against a broader range of data from as many languages as possible. Assuming for the moment that they represent a solid basis for future research, there are many pressing questions that arise.

What, for example, is the reason for the relationship between agency, VoiceP, and eventiveness? The morphological phenomena which reveal the properties of inner vs outer attachment implicate eventiveness as well as agency: inner attachment involves only a single event, or stativity, while outer attachment always entails at least one event, and often (as in the case of productive causativization) two. However, VoiceP is not the locus of introduction of event arguments in the syntax; it is clear that the compositional semantics below VoiceP involves event arguments. Is the single-event limitation in idiosyncratic interpretation simply an accident of the locus of projection of VoiceP? Or is it a necessary consequence of the semantic operations required to introduce external arguments?

Similarly, further crosslingusitic investigation of the locality domains for morphophonological conditioning are called for. Phase theory predicts that VoiceP should be a boundary for the conditioning environments described in List 2, just as for those describe in List 3. Embick (2010) has taken up this challenge and proposed an analysis whereby elements outside VoiceP can condition the morphological realization of elements inside VoiceP under certain particular conditions which reflect the linear nature of morphophonological representations. Absent such conditions, however, VoiceP should be a domain boundary for idiosyncratic morphology just as it is for idiosyncratic interpretation. Careful crosslinguistic work is needed to investigate this question.

Most pressingly, the promissory notes of section 2.4 above need to be cashed, and concrete model-theoretic interpretations both for roots and for derivational affixes worked out in detail. In particular, the interpretations of roots in larger idiomatic structures require attention, since idioms seem to require a conspiracy between the interpretations specified for different roots. If the root of *kick* in *kick the bucket* is given an idiomatic interpretation conditioned by the larger context, so too must the root of *bucket* be idiomatically interpreted, and the conditions must be made mutually dependent so that one entity can't receive an idiomatic interpretation unless the other does. The contributions of the functional categories involved in idiomatic interpretations must be explicitly factored in too, as a central claim of the framework is that the syntactic functional architecture within an idiom is unexceptional, behaving precisely as it does in the non-idiomatic context (Marantz 1995b, 1997, McGinnis 2002). *Kick the bucket* inflects and distributes

like any other verb phrase of English. How, then, does the semantic content of *the* participate in the whole phrase's idiomatic meaning?

Lastly, the discussion in this paper did not touch on one other point which I consider of central importance in the development of our understanding, which is the predictive value of the model developed here for the on-line processing and production of language in real time. Roots, or rather their individual instantiations in all three lists – all three mental 'lexicons' – are accorded a very special status in the model, and we should be able to find evidence for the proposals outlined here using standard psycholinguistic methodologies as argued by Barner & Bale (2002). Indeed, lexical priming work from Taft and Forster (1975) to Twist (2007) and Ussishkin and Twist (2009) supports the notion that even the most semantically underassociated elements from List 2 – *caboodle* roots like *-ceive* and √*sgr* – are accessed in real time during language processing. Pfau (2000, 2009) who was the first to argue that List 1 root nodes needed to be individuated in the narrow syntax within the DM model, argues on the basis of speech error data that the model has the potential to provide a comprehensive and predictive theory of language production. The overall model, and these specific proposals within it, should be evaluated also for their ability to incorprate, respond to, and make predictions about such an increasing range of types of evidence.

Acknowledgments: I owe a huge debt of thanks to many people for helping bring this paper into existence. First and foremost, profound thanks to Edit Doron who organized the conference at which this material was presented, gave me the opportunity to create this target article, provided extremely helpful feedback and guidance, and solicted the incredible group of commentaries that accompany it; I've never before had such an opportunity and cannot thank her enough. Many thanks to the commentators for their extremely stimulating and important discussion of the issues raised here, and also to the participants in the Approaches to the Lexicon workshop in Jerusalem in 2011. Portions of this material were also presented at the Roots workshop in Stuttgart, organized by Artemis Alexiadou, in June 2009, and at the Roots Bound workshop at USC, organized by Hagit Borer, in February 2009; many thanks are due to the organizers and participants at those workshops also. Finally, none of these thoughts would ever have been born without the enthusiasm and endless patience of Maria Florez Leyva and Santos Leyva, who have shared their knowledge of Hiaki with me for many years now. Thanks also to my collaborators on various aspects of the analyses presented in this paper, Mercedes Tubino Blanco, Jason Haugen, Jonathan Bobaljik, and Megan Stone. This research was supported in part by NSF grants BCS-0446663 and BCS-0843898. Needless to say, all of the (many) shortcomings of this work are entirely my responsibility.

References

Acquaviva, P., 2008. Roots and lexicality in Distributed Morphology. Ms.: University College Dublin/Universität Konstanz. Available at http://ling.auf.net/lingBuzz/000654.

Alexiadou, A. & Anagnostopoulou, E., 2008. Structuring participles. In *Proceedings of WCCFL*. pp. 33–41.

Anderson, S. R. 1977. On mechanisms by which languages become ergative. In *Mechanisms of syntactic change*, C. N. Li (ed.), 317–365. Austin: University of Texas Press.

Anderson, S.R. 1992. *A-morphous morphology*, Cambridge: Cambridge Universty Press.

Anttila, A. 2002. Morphologically conditioned phonological alternations. NLLT 20, 1–42

Arad, M., 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory*, 21(4), pp. 737–779.

Arad, M., 2005. *Roots and patterns: Hebrew morpho-syntax*, Dordrecht: Kluwer Academic Pub.

Aronoff, M., 2007. In the beginning was the word. *Language*, 83(4), pp. 803–830.

Aronoff, M., 2011. The roots of language. In Approaches to the Lexicon (Roots III). The Hebrew University of Jerusalem, pp. Paper presented June 15, 2011.

Aronoff, M., 1976. Word Formation in Generative Grammar (Linguistic Inquiry Monographs 1). *Cambridge/Ma*.

Baeskow, H., 2006. A revival of Romance roots. *Morphology*, 16(1), pp. 3–36.

Baker, M., 1985. The mirror principle and morphosyntactic explanation. *Linguistic inquiry*, 16(3), pp. 373–415.

Baker, M. 2010. Types of crosslinguistic variation in case assignment. Ms., Rutgers University. Paper presented at Workshop on Variation in the Minimalist Program, Universitat Autònoma de Barcelona, January 14, 2010, and scheduled to appear in a volume edited by S. M. Ferreirro and C. Picallo.

Barner, D. & Bale, A., 2002. No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification. *Lingua*, 112(10), pp. 771–791.

De Belder, M., 2011. *Roots and affixes: Eliminating lexical categories from syntax*. PhD. Brussels: HU Brussel / Utrecht University.

De Belder, M. & van Craenenbroeck, J., 2011. How to merge a root. http://www.crissp.be/wp-content/uploads/papermjroots.pdf.

Bobaljik, J.D., 2000. The ins and outs of contextual allomorphy. *University of Maryland working papers in linguistics*, 10, pp. 35–71.

Bobaljik, J.D., 2008. Where's phi? Agreement as a post-syntactic operation. *Phi-Theory: Phi features across interfaces and modules*, pp. 295–328.

Bobaljik, J.D. and H. Harley. To appear. Suppletion is local: Evidence from Hiaki. In H. Goad, M. Noonan, G. Piggott, L. Travis, eds, "Word Structure," Oxford University Press.

Borer, H., 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. *The nature of explanation in linguistic theory*, pp. 31–67.

Borer, H., 2009. Roots and categories. In Talk given at the 19th Colloquium on Generative Grammar, University of the Basque Country, Vitoria-Gasteiz. Available at: http://www-rcf.usc.edu/~borer/rootscategories.pdf.

Chomsky, N., 1994. Bare Phrase Structure. *MIT Occasional Papers in Linguistics*, 5, pp. 1–48.

Chomsky, N., 1999. Derivation by Phase. *MIT Occasional Papers in Linguistics*, 18, pp. 1–43.

Chomsky, N., 1995. *The Minimalist Program*, Cambridge, MA: Massachusetts Institute of Technology.

Chung, I., 2009. Suppletive verbal morphology in Korean and the mechanism of vocabulary insertion. *Journal of Linguistics*, 45(03), pp. 533–567.
Comrie, B. 1978. Ergativity. In Syntactic Typology, W. Lehmann ed., 329–394. Austin TX: The University of Texas Press.
Coon, J. & Preminger, O., 2013. Transitivity in Chol: A New Argument for the Split-VP Hypothesis. In *Proceedings of NELS* 36, ed by E. Elfner and M. Walkow. Amherst, MA: GLSA pp. 1–14.
Dickens, C., 1857. *Household words*, Vol. 15 Bradbury & Evans.
Doron, E., 2003. Agency and voice: The semantics of the Semitic templates. *Natural Language Semantics*, 11(1), pp. 1–67.
Dubinsky, S. & Simango, S.R., 1996. Passive and stative in Chichewa: Evidence for modular distinctions in grammar. *Language*, 72(4), pp. 749–781.
Embick, D., 2010. *Localism versus Globalism in Morphology and Phonology*, Cambridge, MA: The MIT Press.
Embick, D., 2003. Locality, listedness, and morphological identity. *Studia Linguistica*, 57(3), pp. 143–169.
Embick, D., 2004. On the structure of resultative participles in English. *Linguistic Inquiry*, 35(3), pp. 355–392.
Fodor, J.A., 1998. *Concepts: Where cognitive science went wrong*, Oxford University Press, USA.
Folli, R., Harley, H. & Karimi, S., 2005. Determinants of event type in Persian complex predicates. *Lingua*, 115(10), pp. 1365–1401.
Fortin, C., 2004. Minangkabau Causatives: Evidence for the l-syntax/s-syntax division. Paper presented at The 11th annual meeting of the Austronesian Formal Linguistics Association (AFLA 11), Zentrum für Allgemeine Sprachwissenschaft, und Typologieforschung und Universalienforschung (ZAS), Berlin, Germany.
Gleitman, L., 1990. The structural sources of verb meanings. *Language acquisition*, 1(1), pp. 3–55.
Guerrero, L., 2004. *The syntax-semantics interface in Yaqui complex sentences: A Role and Reference Grammar analysis*. PhD, Buffalo, NY, University at Buffalo.
Hale, K. & Keyser, S.J., 1993. On argument structure and the lexical expression of syntactic relations. *The view from Building* 20, pp. 53–109.
Hale, K.L. & Keyser, S.J., 2002. *Prolegomenon to a theory of argument structure*, Cambridge, MA: The MIT Press.
Halle, M., 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics*, 30, pp. 425–449.
Halle, M. & Marantz, A., 1993. Distributed morphology and the pieces of inflection. *The view from Building* 20, pp. 111–176.
Harley, H., 1995. *Subjects, events and licensing*. PhD. Cambridge, MA: Massachusetts Institute of Technology.
Harley, H., 2005a. How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in English. *The syntax of aspect*, 1(9), pp. 42–65.
Harley, H., 2005b. Bare Phrase Structure, A-categorial roots, one-replacement and unaccusativity. In *Harvard Working Papers on Linguistics Vol. 11*, ed by S. Gorbachov and A. Nevins, pp. 59–78.
Harley, H., 2007. External arguments: On the independence of Voice° and v°. Talk presented at the 30th GLOW meeting, University of Tromsø, Tromsø, Norway, April 12–14, 2007.
Harley, H., 2009. A morphosyntactic account of the "Latinate" ban on dative shift in English.

Harley, H., 2013. External arguments and the Mirror Principle: On the independence of Voice and v. *Lingua* 125, pp. 34–57.
Harley, H. and M. Stone., 2013. The 'No Agent Idioms' hypothesis. In R. Folli, C. Sevdali and R. Truswell, eds., Syntax and its limits. pp. 251–273. Oxford: OUP.
Harley, H., M. Tubino-Blanco, and J. Haugen. To appear. Locality conditions on suppletive verbs in Hiaki. In V. Gribanova and S. Shih, to appear, The Morphosyntax-Phonology Interface. CSLI: Stanford, CA.
Haugen, J., Tubino Blanco, M. & Harley, H., 2009. Applicative constructions and suppletive verbs in Hiaki. In *Rice Working Papers in Linguistics Volume 1*, ed. by L. Lanz, A. Franklin, J. Hoecker, E. Gentry Brunner, M. Morrison, and C. Pace, pp. 42–51. Houston, TX: Rice University.
Heim, I. & Kratzer, A., 1998. *Semantics in generative grammar*, Wiley-Blackwell.
Hoekstra, T. & Mulder, R., 1990. Unergatives as copular verbs; locational and existential predication. *The Linguistic Review*, 7(1), pp. 1–80.
Horvath, J. & Siloni, T., 2002. Against the little-v hypothesis. *Rivista di Grammatica Generativa*, 27, pp. 107–122.
Inkelas, S. 1998. The theoretical status of morphologically conditioned phonology: a case study of dominance effects. *Yearbook of Morphology* 1997: 121–155.
Inkelas, S. and C. Orhan Orgun. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language* 71:4.
Inkelas, S. & C. Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45, 133–171.
Jackendoff, R., 1977. *X-bar syntax: A study of phrase structure*, Cambridge, MA: MIT Press.
Jackson, E., 2005. Derived statives in Pima. Available at: http://www.linguistics.ucla.edu/people/grads/ejackson/SSILA05PimaDerivedStatives.pdf.
Killimangalam, A. & Michaels, J., 2006. The Three "ikk"s in Malayalam. Ms., Cambridge, MA: MIT.
Kiparsky, P., 1973. "Elsewhere" in phonology. In *A Festschrift for Morris Halle*, ed by S. Anderson and P. Kiparsky. New York: Holt, pp. 93–106.
Kratzer, A., 1994. On external arguments. In *Functional projections*. UMass Working Papers in Lingusitics. Amherst, MA: GLSA, pp. 103–130.
Kratzer, A., 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed by J. Rooryck and L. Zaring. pp. 109–137.
Levin, B. & Hovav, M.R., 1995. *Unaccusativity: At the syntax-lexical semantics interface*, Cambridge, MA: The MIT Press.
Lidz, J., Gleitman, H. & Gleitman, L., 2001. Kidz in the 'hood: Syntactic bootstrapping and the mental lexicon. IRCS Technical Report No. 20, Philadelphia, PA: University of Pennsylvania.
Marantz, A., 1984. *On the nature of grammatical relations*, Cambridge, MA: The MIT Press.
Marantz, A., 1991. Case and licensing. In *Proceedings of ESCOL '91*. Eastern States Conference On Linguistics, ed. by GF Westphal, B. Ao & H.-R. Chae. pp. 234–253.
Marantz, A., 1995a. A late note on late insertion. *Explorations in generative grammar*, pp. 396–413.
Marantz, A., 1995b. "Cat" as a phrasal idiom: consequences of late insertion in Distributed Morphology.
Marantz, A., 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics*, 4(2), pp. 201–225.

Marantz, A., 2001. Words. *WCCFL XX Handout, USC*.
Marantz, A., 2008. Phases and words. In *Phases in the theory of grammar, ed by Dong In*. Seoul: Dong In, pp. 191–222. Available at: http://homepages.nyu.edu/~ma988/Phase_in_Words_Final.pdf.
Marantz, A. 2013. Locality domains for contextual allomorphy across the interfaces. In O. Matushansky and A. Marantz, eds, *Distributed Morphology Today*, 95–115. Cambridge, MA: MIT Press.
Marcus, G.F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, and F. Xu. 1992. Overregularization in language acquisition. *Monographs of the society for research in child development*, 57(4). Chicago, IL: University of Chicago Press.
Markman, E.M., Wasow, J.L. & Hansen, M.B., 2003. Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47(3), pp. 241–275.
McGinnis, M., 2002. On the systematic aspect of idioms. *Linguistic Inquiry*, 33(4), pp. 665–672.
Merchant, J., 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry*, 39(1), pp. 169–179.
Miyagawa, S., 1998. (S) ase as an elsewhere causative and the syntactic nature of words'. *Journal of Japanese Linguistics*, 16, pp. 67–110.
Miyagawa, S., 1984. Blocking and Japanese causatives* 1. *Lingua*, 64(2–3), pp. 177–207.
Miyagawa, S., 1994. Sase as an elsewhere causative. In *Program of the conference on theoretical linguistics and Japanese language teaching: Seventh Symposium on Japanese languages*. pp. 61–76. Tokyo: Tsuda University.
Moravcsik, E. A., 1974. Object-verb agreement. *Working Papers on Language Universals*, 15, pp. 25–140. Palo Alto, CA: Stanford University.
Moravcsik, E. A. 1978. On the distribution of ergative and accusative patterns. *Lingua* 45: 233–279.
Moscoso del Prado Martin, F., A. Deutsch, R. Frost, R. Schreuder, N. H. De Jong, and R. H. Baayen. 2005. Changing places: A cross-language perspective on frequency and family size in Dutch and Hebrew. *Journal of Memory and Language* 53: 496–512.
Nunberg, G., Sag, I. & Wasow, T., 1994. Idioms. *Language*, 70(3), pp. 491–538.
Panagiotidis, P., 2005. Against category-less roots in syntax and word learning: Objections to Barner and Bale (2002). *Lingua*, 115(9), pp. 1181–1194.
Pfau, R., 2000. *Features and categories in language production*, Inauguraldissertation zur Erlangung des Grades eines Doktors der Philosophie im Fachbereich Neuere Philologien der Johann Wolfgang Goethe-Universität zu Frankfurt am Main.
Pfau, R., 2009. *Grammar as processor: a distributed morphology account of spontaneous speech errors*, Amsterdam: John Benjamins Publishing Company.
Pilley, J.W. & Reid, A.K., 2011. Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86(2), pp. 184–195.
Punske, J. and M. Schildmier Stone (2013). Idiomatic expressions, passivization, and gerundization. Talk presented at the 2013 meeting of the Linguistic Society of America, Minneapolis, MN, Jan 3, 2013.
Pylkkänen, L., 2002. *Introducing arguments*. PhD. Cambridge, MA: Massachusetts Institute of Technology.
Pylkkänen, L., 2008. *Introducing arguments*, Cambridge, MA: MIT Press.
Ramchand, G., 2008. *Verb meaning and the lexicon: A first-phase syntax*, Cambridge University Press.

Rude, N., 1996. Objetos dobles y relaciones gramaticales: el caso del yaqui. In *III Encuentro de Lingüística en el Noroeste*, ed by Z. Estrada, M. Figueroa, and G. López, pp. 491–522. Hermosillo: UNISON.

Ruwet, N., 1991. On the use and abuse of idioms. *Syntax and Human Experience*, pp. 171–251.

Serratos, A. E., 2008. *Topics in Chemehuevi morphosyntax: Lexical categories, predication and causation*. PhD. The University of Arizona.

Siddiqi, D., 2006. *Minimize exponence: Economy effects on a model of the morphosyntactic component of the grammar*. PhD. The University of Arizona.

Siddiqi, D., 2009. *Syntax within the word: economy, allomorphy, and argument selection in Distributed Morphology*, Amsterdam: John Benjamins Publishing Company.

Speas, M., 1986. *Adjunctions and projections in syntax*. PhD. Cambridge, MA: MIT.

Speas, M., 1990. *Phrase structure in natural language*, New York; Springer.

Sugioka, Y., 2001. Event structure and adjuncts in Japanese deverbal compounds. *Journal of Japanese Linguistics*, 17, pp. 83–108.

Sugioka, Y., 2002. Incorporation vs. Modification in Japanese deverbal compounds. In *Japanese/Korean Linguistics 10*, ed. by Akatsuka and Strauss. Stanford: CSLI Publications, pp. 496–509.

Svenonius, P. 2004. Slavic prefixes and morphology, *Nordlyd* 32.2:177–204.

Svenonius, P. 2005. Two domains of causatives. *CASTL, University of Tromsoe, unpublished ms*.

Taft, M. & Forster, K.I., 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), pp. 638–647.

Travis, L., 2000. Event structure in syntax. *Events as grammatical objects: The converging perspectives of lexical semantics and syntax*, pp. 145–185.

Twist, A., 2007. *A psycholinguistic investigation of the verbal morphology of Maltese*. PhD. The niversity of Arizona.

Ussishkin, A. & Twist, A., 2009. Auditory and visual lexical decision in Maltese. In *Introducing Maltese linguistics: selected papers from the 1st International Conference on Maltese Linguistics, Bremen, 18–20 October, 2007*. ed. by B. Comrie, R. Fabri, E. Hume, M. Mifsud, T. Stolz and M. Vanhove. pp. 233–249.

Veselinova, L., 2003. *Suppletion in verb paradigms: bits and pieces of a puzzle.* PhD. Stockholm, Sweden: Stockholm University.

Veselinova, L., 2006. *Suppletion in verb paradigms: bits and pieces of the puzzle*, Amsterdam: John Benjamins Publishing Co.

Wasow, T., 1977. Transformations and the lexicon. In *Formal syntax*, ed. by A. Akmajian, P. Culicover and T. Wasow. New York: Academic Press, pp. 327–360.

Woolford, E. 2006. Case-agreement mismatches. In Cedric Boeckx ed., *Agreement systems*, 317–339. Amsterdam: John Benjamins.

Yip, M., Maling, J. & Jackendoff, R., 1987. Case in tiers. *Language*, 63(2), pp. 217–250.