



<b>Title</b>	Roots and Lexicality in Distributed Morphology
<b>Authors(s)</b>	Acquaviva, Paolo
<b>Publication date</b>	2009-05
<b>Publication information</b>	Acquaviva, Paolo. "Roots and Lexicality in Distributed Morphology." University of York. Department of Language and Linguistic Science, 2009.
<b>Conference details</b>	Fifth York-Essex Morphology Meeting (YEMM), 9th February and 10th February 2008, Department of Language and Linguistic Science, University of York
<b>Publisher</b>	University of York. Department of Language and Linguistic Science
<b>Item record/more information</b>	<a href="http://hdl.handle.net/10197/4148">http://hdl.handle.net/10197/4148</a>

Downloaded 2023-03-15T17:09:45Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd\_oa)



© Some rights reserved. For more information

*Abstract*

This paper examines the nature and content of morphological roots in relation to their syntactic context. A careful consideration of doublets, where the same root may take alternative noun-inherent features, leads to the claim that roots do not carry selectional features or class diacritics. Relying on the distinction between syntactic nodes and their exponents, central to a realizational model like Distributed Morphology, I argue that the syntactic atoms corresponding to root nodes are associated with open-class exponents but not with a specific meaning that might select a licensing syntactic context. "Lexical" meaning arises constructionally, and so do lexical properties like gender or class, which however emerge at Vocabulary insertion and may show selectional properties. Content and exponence of roots are thus dissociated, in line with the separationist character of Distributed Morphology. This predicts the existence of root-like elements with mixed status, namely open-class exponents used as grammatical morphemes (like auxiliaries or classifiers), or category-free root extensions below the innermost category-assigning head (like *de-* in *de-struction*).

*1. Introduction*

The notion of morphological root is based on the intuition that complete words may share a minimal "core", which remains invariant when all identifiable morphological formatives have been abstracted away. Distributed Morphology has incorporated this intuition in a realizational approach where morphology manipulates, linearizes, and spells out an input syntactic structure (Halle and Marantz 1993, Noyer 1997). Within this framework, a consensus has emerged to the effect that roots correspond to the non-grammatically definable part of a word; that part, that is, which does not arise from the spellout of morphosyntactic features hosted on functional heads (Marantz 1997). In recent years, this has taken the shape of a more definite hypothesis: all lexical categories are made up of category-less roots combined with category-assigning heads (Marantz 2001, Embick and Noyer 2007, Embick and Marantz 2008). For example, nouns and verbs are not syntactic atoms, but derive respectively from a structure [Root + n] and [Root + v], where [n] and [v] are syntactic functional heads whose grammatical content defines a nominal and a verbal domain.

My aim is to explore some of the implications of this hypothesis, and ask what properties roots must have if we take lexical decomposition seriously. The first part of this paper is a critique of the received view of roots: I will argue that roots are not marked for class diacritics (section 2), and, more radically, that there is no coherent sense in which a root has a meaning that may be isolated from the meaning of a larger structure (section 3). This raises fundamental questions about the representation of lexical information, which are addressed in the rest of the paper. Section 4 illustrates how root-conditioned phenomena may be analyzed without special marking on the root itself; section 5 develops Harley and Noyer's (2000) distinction between roots as exponents and the syntactic nodes hosting them, reaching novel conclusions. Finally, section 6 focuses on the question of what content roots can have at all, if they have no meaning in isolation; the answer is that they mark difference and relatedness between lexical words. Categorized words, not roots, are the minimal units of interpretation; but roots make

lexical words related or distinct, acting as indices with a purely differential value. Many open questions remain, of course, some of which are outlined in the conclusion.

## 2. *Roots have no diacritic features*

The words of a language are partitioned by morphosyntactic criteria into categories, which match more or less closely the traditional notions of nouns, verbs, adjectives, prepositions, adverbs, and pronouns. Across these categories, both grammatical and lexical words typically fall into one of several possible morphological classes, usually divided into nominal and verbal inflectional classes (declension and conjugation). Grammatical words, by definition, are entirely defined by their featural content, so it is the choice of features that determines class. In this sense, if they belong to a morphological class at all, the choice of class is deterministic. For lexical words, on the other hand, class usually depends on the choice of the root, and is not therefore grammar-driven. In most accounts, the representation of root-dependency is straightforward: roots are equipped with class diacritics. Embick and Halle (2005) state this solution in unambiguous terms:

Membership in one of the conjugation classes is an arbitrary property of the roots that appear in the Latin verbal system. The simplest implementation of this fact involves specifying each Root for a diacritic feature that encodes membership in a specific class:  $\sqrt{\text{AUD}}_{[\text{IV}]}$

(Embick and Halle 2005:46)

This is the approach taken in most Distributed-Morphological accounts, made popular by Oltra Massuet's (1999) influential analysis of thematic vowel selection in Catalan. Harris (1999) takes the same line, at least as a "default case":

The particular form-class morpheme or thematic suffix is selected by the closest non-null c-commanding morpheme; for example, by the stem-forming morphemes like *-ic-* and *-ar-* or by the root in the default case if the stem-licenser has no phonological content

(Harris 1999:54)

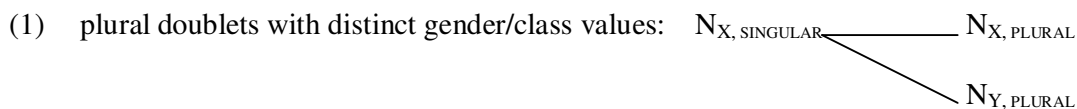
Outside Distributed Morphology, the direct encoding of diacritic features on roots is also argued for in analyses that share the assumption of lexical decomposition but not of distinct post-syntactic morphological operations: see Josefsson (2001) and Alexiadou and Müller (2008), who interpret in syntactic terms the agreement in category/declension feature between category head (probe) and category-less "stem" (goal).

Representing diacritics directly on roots is conceptually problematic, however. Quite simply, if a root has a feature that presupposes a category, then it is not really category-free. Positing an invisible class marker on a root in order to make sure that it ends up in the right nominal or verbal inflectional class simply states the observed correlations (if noun, class X, if verb, class Y), treating them as part of the root itself. But then the root has nominal or verbal information, which is precisely what the lexical decomposition hypothesis is meant to exclude.

One of the consequences of this simplistic solution is that the roots that appear in a same-category pair like the Italian verbs *arrossare* 'to make red' and *arrossire* 'to become red', which differ in conjugation but share the same root with the adjective *rosso* 'red', are different: either differently annotated, or accidentally homophonous. If they are differently an-

notated, say RED<sub>A</sub> and RED<sub>I</sub> (plus RED<sub>ADJ</sub> for the adjective), then there is an invariant core underlying the different features, and *that* is the root; otherwise, if the choice of diacritics is an integral part of the root as an unanalyzable atom, the formal identity of these three distinct roots has no formal expression and is therefore treated by the grammar as accidental.

A variant of this argument comes from the widely attested phenomenon of plural doublets, where one singular noun has two or more plural alternants with distinct diacritics. The examples in (2a–c), from unrelated language families, illustrate the abstract scheme in (1):



- (2) a. osso 'bone', masc.sg. ossi 'bones' masc.pl.  
 ossa 'bones' fem.pl.

(Italian; Acquaviva 2008)

- b. dáas 'shop', masc.sg. daas-ás 'shops', masc.pl.  
 daas-yó 'shops' fem.pl.

(Somali; Lecarme 2002:120)

- c. u-ta 'bow', class 11 nyu-ta 'bows', class 10  
 ma-ta 'bows and arrows, class 6

(Swahili; Contini-Morava 1999:8)

Again, if gender or class is constitutive of the identity of a root, then there must be two homophonous roots, one of which only surfaces in the plural. Both this restriction to a number value and the semantic overlap (or even identity) have no grammatical expression; as far as the system goes, they are entirely accidental. If on the other hand there is a shared core, the root must be identified with this featureless atom. The latter is the position I would like to argue for, in keeping with the lexical decomposition hypothesis (cf. also Harris 1996:105 note 15: "Roots have no morphosyntactic category, no gender, *and no form of class affiliation*"; emphasis mine).

In addition to these conceptual arguments, there is also empirical evidence that roots should not carry diacritics, in particular class diacritics like declension and conjugation class. Suppose they do: then, these diacritics must be visible in the syntax, in so far as roots occupy syntactic nodes (regardless of whether they are spelled out after syntax or enter syntax already specified for phonological form, as claimed by Embick and Halle 2005). However, this sort of information *never* enters into relations of syntactic agreement. Not only that: if roots had class features, we would expect that some heads attaching to them should be sensitive to the choice of features. This would lead, for example, to nominalizations only for roots that surface as 2nd declension verbs, or to adjectives corresponding only to nouns with irregular plural, or to causative verbs built only on the roots of *-a*-stem nouns. But this does not seem to happen. Apparently, lexical stems and some function words have diacritics; category-free roots, in so far as one accepts them as pieces of morphology, do not.

### 3. *Roots have no meaning by themselves*

The idea that roots have no diacritics is actually very natural given the lexical decomposition hypothesis: if roots are category-free, they are not marked for category-internal subdivisions.

But taking this hypothesis seriously we are led to the much more radical conclusion that roots, taken by themselves, have no meaning.

This probably sounds counterintuitive, and certainly runs counter to all that has been said in *Distributed Morphology* about the semantics of roots: for Marantz (1997:212–213), "things with special meanings are roots"; and Marantz (2001) specifies that roots can mean entities, states, or events. To see why this view is conceptually problematic, we must go back to the kind of evidence that motivated category-free roots in the first place.

In a broad-ranging critique to Hale and Keyser's (1992) structural decomposition of the lexical semantics of verbs, Kiparsky (1997) showed that only some noun-related verbs appear to be truly denominal, but not others. For instance, the activity described by *to paint* must involve the use of a "paint-like substance", but *to dump* does not in any way refer to a dump or dump-like thing. In Kiparsky's words, "Instrumental verbs likewise include both true denominal verbs, which contain the meaning of the noun, possibly in an attenuated version, and verbs which are related to nouns via a shared root, and which do not semantically incorporate the meaning of the noun". The opposition emerges when the sentence contains an instrumental adjunct that explicitly contradicts the instrument named in the verb:

- (3) a. #she taped the picture to the wall with pushpins
- b. he hammered the desk with his shoe

*To tape*, like *to paint*, seems to "contain the meaning of" the corresponding noun in a way that is not true of *to hammer* or *to dump*. At least in some cases, then, verbal derivation involves not a category-free root, but a real noun.

To account for these facts, Arad (2003) argued that roots may be affixed to by more than one category-assigning head. *To tape* "contains the meaning of the noun" because the projection headed by [v] literally contains the noun, while *to hammer* is a root verbalization:

- (4) a. to tape           [[[ tape ] n ] v ]
- b. to hammer       [[ hammer ] v ]

This straightforward solution models in structural terms the difference between denominal and root derivations. But it also brings into focus a latent question about root derivations: if *to hammer* and *to dump* do not "contain the meaning of" the nouns *hammer* and *dump*, what other semantic relation links the noun and the verb? A very tenuous one, it would appear. Kiparsky (1997) simply notes that "(M)orphologically, noun and verb could still be analyzed as related, but the relationship would be a matter of derivation from a common root". Indeed, positing a non-categorized root serves precisely to express this type of relation. But if this root is a piece of morphosyntactic structure, as in *Distributed Morphology*, a question arises about its meaning, not just its form. And this meaning seems too elusive to be pinned down. This is because the root HAMMER in (4b) is neither a noun nor a verb, nor any other category; something so radically underspecified cannot even convey the distinction between argument and predicate. What meaning can a root have that is not yet specified as an entity-, state- or process-referring expression?

To appreciate this point, it will be useful to consider how Harley and Noyer (2000) related the lexical-semantic properties of *to grow*, *to destroy*, *growth*, *destruction* to the semantic content of the respective roots: "The speaker knows that these roots [GROW, DESTROY] denote events that may occur spontaneously, like growing, or that may be truly externally caused, and hence it may not participate in a transitive nominalization [...] This knowledge is part of the real-world knowledge of the speaker about the meaning of the root ... " (Harley and Noyer 2000:365). But this knowledge is about the meaning of a *verb*, not of a root; it is verbs that describe events and the relations between their participants. The same point applies

to Alexiadou's (2001) treatment of nominalizations as consisting of a root dominated by an articulated functional structure:

Levin (1999) argues that verb-meanings have two components: one component which is provided by its event structure, and one component provided by the core meaning, i.e. the part of the meaning which is idiosyncratic to a specific verb. This second type is referred to as the 'constant'. Constants are then integrated into event structure templates to yield the various verb meanings. Let me equate what Levin refers to as constant with an unspecified root. (Alexiadou 2001: 66–67)

This extended quotation makes it unambiguously clear that the "core" meaning in question is verbal (and indeed, Marantz, Harley and Noyer, and Alexiadou all analyze nominalizations, not nouns). It seems plausible to speak of general 'growing' or 'destroying' concepts, but these concepts are of verbal nature: they involve events extended in time, with a specific argument structure and actional properties (note that the respective nominalizations *growth* and *destruction* have a transparently deverbal interpretation). If there is such a thing as the meaning of the root alone, it should be distinct from the meaning of the verb; but it is not clear how the interpretations outlined by Harley and Noyer differ from those of the corresponding verbs.<sup>1</sup> In fact, it is not clear how they could conceivably be different. It seems impossible to describe the semantic "core" of *grow* or *destroy* without implicitly treating them as verbs — or as deverbal nominalizations. In the same way, it's impossible to capture the meaning shared by *to hammer* and *hammer* without recourse to notions that are already categorized as entity-referring nouns or state- / process-referring verbs. In sum, when the lack of category label is taken seriously, and is not just a formal device to mask elements already interpreted as nouns or verbs, roots can be assigned no coherent meaning on their own, because meaning presupposes at least a categorization in semantic types, and this in turn presupposes a syntactic category.

In an approach like that of Kiparsky (1997), this leads to the conclusion that some root-related lexical items (nouns and verbs) can be related just in their form, and not in their lexical semantics. But in a version of Distributed Morphology that generalizes category-free roots, there are no lexical items corresponding to nouns or verbs, and no lexical semantics independent of the pieces that make them up. The meaning of *hammer* and *to hammer* must result from that of their syntactic parts — in particular, from the meaning of the root they share. But there does not seem to be an identifiable semantic core to match this shared morphological core. This might appear to be a decisive argument against reifying category-free roots as syntactic pieces. I believe instead that these objections represent a serious empirical challenge, but a challenge that can be met. Achieving this goal requires a deeper conception of roots, which is not only theoretically more coherent but also, I will argue, empirically superior to the current one.

#### 4. Lexical information is root-external

If the critique articulated in the preceding two sections is on the right track, roots do not select their syntactic environment, either by formal diacritics or in virtue of their meaning. But, of course, not all roots occur in all contexts. The challenge lies in representing root-related lexical properties as constructional. In this section I will offer three examples of how morphological and semantic information can be dependent on the choice of a root without being encoded on the root itself. The general approach is based on Harley and Noyer's (2000) idea, taken up and developed by Galani (2004), that root Vocabulary items are *licensed* in certain syntactic

environments. To say "a noun has gender X", for instance, means in this perspective "a root Vocabulary item is licensed in the context of [n] with gender X". This type of licensing is clearly linguistic knowledge, but it is not deterministically grammar-driven. Licensing statements that apply to lists of roots, by themselves, are not more (nor less) arbitrary than explicit specifications on each root; the difference is empirical, and I will argue that viewing lexical information as a property of constructs rather than roots is empirically more successful. The crucial difference from earlier approaches is that meaning arises in a construction, not in a root.

#### 4.1. Dutch conversion (Don 2004)

Two prototypical root-dependent properties of nouns and verbs are gender and conjugational class. In Dutch, nouns are either neuter or non-neuter:

- (5)    *het huis*                      *de weg*  
          'the.NEUT house'        'the.NON-NEUT road'

Verbs are either regular, in the sense of stem-invariant across the paradigm, or irregular, if they display stem-internal alternations (revowelling) across different cells of the paradigm:

- (6)
- |              | regular       | irregular   |
|--------------|---------------|-------------|
|              | 'to count'    | 'to fall'   |
| present(1sg) | <i>tel</i>    | <i>val</i>  |
| past (1sg)   | <i>tel-de</i> | <i>viel</i> |

It is common for one and the same root to appear as noun or verb without further affixes. Assuming nouns and verbs to be basic categories, this amounts to conversion, or zero-derivation from one category to the other. The assumption of category-free roots allows instead to view both as (zero) root derivations. However, Don (2004) points out that verbal irregularity and nominal neuter gender do not cross-classify completely for these conversion pairs. Of the four possible combinations, one is excluded:

- (7)
- |    |                                  |   |
|----|----------------------------------|---|
| a. | regular verb ~ non-neuter noun   | <i>(tel ~ de tel</i> 'to count' ~ 'the count')  |
| b. | regular verb ~ neuter noun       | <i>(werk ~ het werk</i> 'to work' ~ 'the work') |
| c. | irregular verb ~ non-neuter noun | <i>(loop ~ de loop</i> 'to walk' ~ 'the walk')  |
| d. | *irregular verb ~ neuter noun    | (Don 2004:940)                                  |

Irregular verbs may of course be turned into neuter nouns by affixal nominalization (e.g. *zaag* 'to saw' > *het zaag-sel* 'sawdust'), but an irregular verb with form X cannot appear as a neuter noun with the same form X, up to a few exceptional pairs (e.g. *sluit* 'to close' ~ *het slot* 'lock'; Don 2004: 954, note 6).

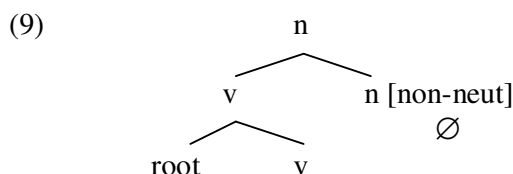
Don correctly observes that this distributional gap is unexpected if both nouns and verbs are constructed from a category-free root, unless some stipulation restates the observed incompatibility. He therefore proposes directional zero-derivations from one category to the other, specified as assigning to the output the unmarked value of the respective feature: nouns become regular verbs (N > V [regular]), and verbs become non-neuter nouns (V > N [non-neuter]). Since the output has always the unmarked value, neither derivation can turn an input with marked value (neuter or irregular) into an output with marked value.

What makes this analysis attractive is that it captures without stipulation the role of markedness: a zero-derivation may change category, but then it assigns it the unmarked feature value. This insight can be preserved while keeping to the lexical decomposition hypothesis, if the features of nouns and verbs are represented not on roots but on the category-assigning heads [v] and [n]. A neuter noun is one whose root is associated with a [n] marked [neuter], and an irregular verb is one whose root is associated with a [v] marked [irregular]. The idea is then that, in [[ROOT] n], a neuter [n] has a zero spellout only in the context of listed roots, as illustrated by the Vocabulary items in (11):

(8) a. [n: neuter] ↔ ∅ / {ROOT}\_\_\_\_\_ (ROOT = WERK, HUIS ...)

b. [n: non-neuter] ↔ ∅

In the absence of a neuter affix, a zero neuter nominalization is a marked phenomenon that requires specific mention of the root to which it applies. Since the root must be local to [n], the list of roots in the context for (8a) cannot contain irregular verbs, for the simple reason that irregular verbs are verbs, i.e. [ROOT + v]. A null [n] head may embed such a verbal complex, but in that case the root is separated from [n] by the intervening head [v]. Because the contextual restriction for (8a) is not met, for [n] to be spelled out as zero it must be non-neuter, as in (8b):



As a result, only root nominalizations are neuter (again, restricted to zero-derivation):

(10) a. (het) werk      [[WORK] n<sub>NEUT</sub>]      root zero-nominalization

b. (de) loop      [[[WALK] v<sub>IRREGULAR</sub>] n<sub>NON-NEUT</sub>]      verb zero-nominalization

This analysis does not rule out the pattern (7d) on principled grounds. Nothing bars neuter root nominalizations [[ROOT] n] where the corresponding verb [[ROOT] v], based on the same root, is irregular. However, this requires the same root to be listed both in the context of (8a) and in the context of [irregular] [v]. This option is open in principle, but is bound to have a restricted application; in addition, since both noun and verb are here root derivations, a certain amount of semantic opacity is expected. This is exactly what happens with the few exceptions mentioned by Don. So, the fact that the grammar does not actually rule out the fourth pattern in (7) is a strength of the analysis, once markedness relations make this option a peripheral one. By contrast, no structural solution would be forthcoming if [irregular] was a feature on roots, short of stipulating that neuter [n] cannot be zero in the context of ROOT<sub>IRREGULAR</sub>. This restatement of the generalization, however, would not account for the gender and the partial semantic unpredictability of exceptions like *het slot* 'a lock' (an object, not an event like 'a walk' or a result nominal like 'a piece of work').

This structural reinterpretation of Don's analysis shares with other approaches the insight that root-specific means root-local (Marantz 2001, Arad 2003, Embick 2003, Embick and Marantz 2008). In this perspective, the Dutch facts are closely related to the familiar regularization of English irregular verbs when they are reinterpreted as denominal. *To cost*, for instance, has the irregular past *cost* in its main sense, but *costed* in the denominal sense 'to



quantify the cost'. This pattern follows if [irregular] is a feature of [v] in the immediate context of listed roots, but not when a [n] intervenes:

- (11) a. the project cost me a lot of effort (irregular) [[COST]v] (root derivation)  
 b. I costed the project (regular) [[COST]n]v] (denominal)

This regularization pattern would not lend itself to such a straightforward structural explanation if the two verbs had differently annotated homophonous roots. If the source of this root-dependent feature lay in the root itself, there would be no reason why the regularized variant should systematically presuppose a noun (the relevance of this was highlighted by Kiparsky 1997, and noted by Don 2004).<sup>2</sup>

#### 4.2. *oxen, oxes*

The plural of *ox* is *oxen*, an irregularity that has become a standard example of the disjunctivity of inflectional morphology: the root-dependent exponent blocks the regular one. But it is not completely true that *oxes* is blocked by *oxen*. For many speakers *oxes* is preferred or at least possible when *ox* has the sense of 'large uncouth man' (cf. the discussion of similar cases in Pinker 1999). One may invoke the difference between literal and non-literal reading to justify two distinct nouns. But this seems unrevealing, to say the least. If metaphorical or metonymical extensions were to justify distinct lexical entries in all cases, there would be a very large mass of homophonous noun pairs lacking the morphological distinction of *oxen* – *oxes*. But even in the presence of such morphological distinctions, simply positing two distinct nouns would be unsatisfactory without somehow expressing that the two are related, and that the meaning of one is based on that of the other and not conversely. Finally, and decisively for present purposes, what must be captured is that the semantically derived version very often has a regularized morphology. This is not true in all cases (*brethren*, for instance, is arguably derived from the basic biological sense of 'brother', but has an irregular plural), but it seems to be consistently true when semantic extension amounts to a metaphorical application to human beings: if humans with the grace of an ox are likely to be *oxes* rather than *oxen*, human parasites are often *louses* rather than *lice*. This is not restricted to English: in Breton, Trépos (1957: 72–75) reports that *bioc'h* 'cow' has the suppletive plural *saout* 'cattle', but the regular suffixal plural *beuc'h-ed* 'cows' when applied as a derogatory term for women. All this suggests that the regularized form *oxes* is not an isolated accident, but is a consequence of the semantic extension of the basic sense of the noun to humans. The analytical problem is that the plural morpheme expresses purely grammatical information, yet its choice depends on the lexical semantics of the base. Speakers choose freely between words, roots, stems, and lexical 'senses'; but the choice of a plural morpheme should be deterministic and grammar-driven.

In a framework without roots, Halle (1990) and Halle and Marantz (1993) could simply posit a plural allomorph in the context of listed nouns. But assuming the number-inflected *oxen* to correspond to [OX]+[n]+[Number: Plural], we cannot simply stipulate an idiosyncratic choice of the plural morpheme in [Number] in the context of certain roots, since [n] intervenes. What is called for is a way to make Number noun-dependent, not just root-dependent. A two-step analysis suggests itself. First, we can posit a Fusion of [n] and [Number] for a plural value of the latter and for certain choices of listed roots:

- (12) Fusion (for root-dependent irregular plural nouns):

[n]+[Num: Pl] → [n, Num: Pl] / {OX, CHILD, BRETHR, GOOSE, MOUSE, ...}\_\_\_\_\_

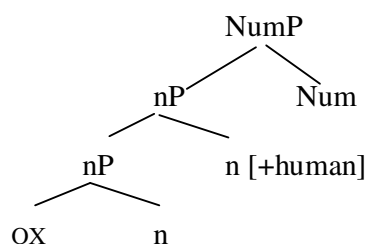
This root-triggered Fusion allows a root-dependent spellout of plurality, because it makes [Number] adjacent to root. This operation takes place for all cases of root-dependent idiosyncratic Number morphology, not just for plurals in *-en*. In fact, it is the formal expression of the fact that to know certain nouns involves knowing that they have a special plural form: the information about Number coalesces with the information about the root making up a noun. Given (12), the idiosyncratic ending *-en* represents a particular realization of the fused input [n, Num: Pl], for a subset of the relevant roots:

(13) Vocabulary insertion (restricted to *-en*)

[ n, Num: Pl ] ↔ *-en* / {OX, CHILD, BRETHR}\_\_\_\_\_

What of the regularized *oxes*? A revealing account must capture its derived status (formally and semantically) from the base noun, and it must circumvent the deterministic insertion of *-en*. One simple assumption achieves both goals: that the reading of *ox* as a human being involves an additional head [n], which hosts the feature [+human]. It is not crucial that this information is featurally represented; what is crucial is that it is introduced by an additional [n]:

(14) Regularized pattern (singular and plural): additional [n] with human interpretation



This structure does not present a context for Fusion in (12), because at no point does a sequence [n]+[Num] find itself linearly or structurally adjacent to the root. The realization scheme in (13) therefore does not apply, because the input is not there, and [Num: Pl] is regularly spelled out as *-s*. As in the previous case of verbs like *cost*, an independently justified more complex structure (here, the derived sense of *ox* contains the base one) blocks the realization of irregular morphology, by breaking the adjacency with the root. The solution is not ad-hoc, in so far as regularization typically accompanies semantic extension to humans, and this interpretation "contains" the meaning of the base noun like a true denominal verb "contains" the meaning of an instrument noun.

#### 4.3. *ossa, ossi* (Acquaviva 2002, 2008)

A small set of Italian plural doublets raises the same sort of questions caused by non-deterministic choice between plural forms, with the additional twist that the alternants have different gender:

(15) *osso* 'bone' (masc. sg.) } *ossi* 'bones' (unconnected wholes) masc. pl.  
} *ossa* 'bones' (connected parts) fem. pl.

The value of the gender feature, unlike that of number, is inherent in nouns: in the normal case, singular and plural are alternative values for one and the same noun, but masculine, fem-

inine, and neuter are values that correspond to distinct nouns (although they may share the same *stem*; cf. Alexiadou 2004 for some critical discussion of the non-interpretable nature of gender as a formal feature). Even before addressing the anomalous pattern schematized in (15), we should reconcile this intuitive observation with the assumption that nouns are in fact constructions. As outlined in the beginning of this section, I assume that gender and class diacritics are hosted in [n] (following in particular Kihm 2005 and Lowenstamm 2006). A statement like "a noun has gender X" means "a root is licensed in a syntactic context with [n] having gender X". Learning the gender of nouns amounts to learning the roots that go with a choice of gender on [n]. Gender is thus a constitutive property of every noun (in this sort of language), but no grammatical mechanism forces a root to select one and only one gender value. Masculine-feminine pairs that share the same root need no special provision, as they simply illustrate roots that are licensed in both contexts:

(16) *cugino* ~ *cugina* 'cousin (male)' ~ 'cousin (female)'

[n: masc] / { ... COUSIN ... } \_\_\_\_\_  
 [n: fem] / { ... COUSIN ... } \_\_\_\_\_

In particular, there is no need to stipulate homophonous or doubly-marked roots, as would be the case if gender was encoded directly on the root itself.

Under these assumptions, the feminine alternants of Italian irregular plurals must represent distinct pairings of root and [n]; that is, technically, distinct nouns; more precisely, distinct lexemes. As detailed in Acquaviva (2008), this hypothesis accounts for the empirical differences between these Italian nouns and the similar-looking Romanian neuter nouns, which likewise have a masculine singular and a feminine plural. Apart from several other important differences, above all in the numerical size of the class, the Romanian neuters deterministically take on feminine gender when plural as a rule of inflection, without giving rise to doublets and without any accompanying effect on the nouns' lexical semantics. Italian feminine plurals like *ossa* 'bones', by contrast, are lexemes in their own right and therefore do not stand in competition with the regular masculine plurals, which may or may not be available (usage varies in this respect) but are in principle allowed by the grammar. This instance of pluralization, accompanied by feminine gender, has the function of nominal derivation. This immediately accounts in a unified fashion for the semantic difference with respect to the regular alternant, where the latter exists. What is peculiar, however, is that the feminine alternants are not just gender doublets like 'cousin' in (16): they are *plural*, and in some cases they are the only plural forms for the respective masculine singular. For nouns like *uovo* ~ *uova* 'egg' (masc.) ~ 'eggs' (fem.), where the feminine is the only possible plural, inflection and derivation mix in a curious way: the number opposition is unambiguously inflectional, but it is realized by derivational means, in that the plural form arises by deriving a distinct noun (defined by gender).

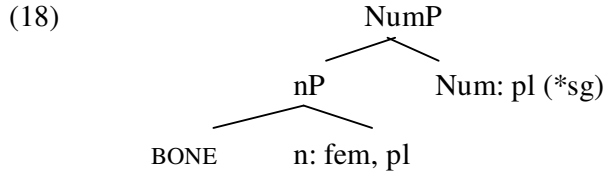
To account for this in precise structural terms, we need to remind ourselves that in Italian, unlike for instance in Spanish, gender and number always have a fused exponence. There is no discrete plural morpheme comparable to the Spanish (or English) *-s*, which expresses number independently of the noun's gender. Instead, the end vowels of Italian nouns (and adjectives, and some pronouns and determiners) always act as portmanteau affixes, fusing gender and number. If the two features are hosted on different heads, [n] and the immediately higher [Number], this means that the two heads must undergo Fusion as a property of Italian morphology:<sup>1</sup>

<sup>1</sup> A reviewer flags an important potential problem: does the proposed analysis entail that all such gender doublets ultimately go back to the Fusion of gender and number? If so, (modern) Greek repre-

(17) Italian: Fusion of gender and number

$$[n: X] [\text{Num}: Y] \rightarrow [n: X, \text{Num}: Y]$$

The resulting node is then spelled out as a single terminal. Since gender and number must end up fused in the same node, nothing bars [n] from carrying number information from the start, that is, in the syntax. When this happens, the number value on the higher head Number must agree (in the informal sense of the word): if [n] is specified as plural, the head [Number] must also be plural. The initial structure of the irregular plurals, before Fusion, is therefore the following:



This analysis makes precise the intuition that the feminine plural alternants (which are irregular also in their ending) are not feminine nouns in a plural context, nor masculines that change gender in the plural (like in Romanian), but inherently plural feminine nouns. They are, then, root nominalizations, on a par with their respective singular (which, incidentally, not always exists). In the case of *osso* 'bone', we have three forms: the regular singular and plural masculine *osso* ~ *ossi*, and the irregular feminine plural *ossa*:

- |         |                                       |                      |
|---------|---------------------------------------|----------------------|
| (19) a. | <i>osso</i> 'bone'                    | [[BONE ] n: masc]    |
| b.      | <i>ossa</i> 'bones (connected parts)' | [[BONE ] n: fem, pl] |

One crucial property of these feminine plurals is that they often, but not always, induce a distinctive interpretation. If a masculine alternant is present, the two are always differentiated, no matter how subtly. This follows from the hypothesis that they are distinct root-[n] pairings. It is well known that the association of the same root with distinct gender values (with distinct contents for [n], in this model) can, but does not have to, bring about variously differentiated interpretations. Italian examples range from the subtle nuance of *buco* (masc.) ~ *buca* (fem.) 'hole, gap (bidimensional, as in a sheet of paper or in a blanket) ~ 'hole' (tridimensional, as in a road or a snooker table), through the standard sex-differentiation of *cugino* ~ *cugina* seen above in (19), to the equally systematic distinction between *pera* 'pear' ~ *pero* 'pear-tree', to the arbitrary lexical differentiation of *cassetto* 'drawer' ~ *cassetta* 'cassette'. Likewise, the irregular feminine plurals may comport no special meaning at all (like *uova* 'eggs'), a more or less subtle nuance (like *osso* 'bone' ~ *ossa* 'bones as connected parts, not self-standing pieces'), or a clear semantic distinction (like *membro* 'member', e.g. of a committee, ~ *membra* 'limbs'). It would be wrong to split what is ostensibly one and the same root into two homophones, one of which characterized by feminine gender, plural number, and occasionally a special interpretation. The root is the same across masculine and feminine, singular and plu-

---

sents a counterexample, because noun endings there fuse number and case, but not gender. This is not because gender has a separate morpheme, which it hasn't, but because the form of the endings reference case, number and inflectional class, but not directly gender (cf. Alexiadou, Haegeman, and Stavrou 2007:268-269). Whether this is a real or only apparent problem depends on the precise analysis of Greek doublets, which I cannot offer here. The classical Greek plural doublets discussed in Acquaviva (2008:40) parallel Italian (and Latin), in so far as the plural ending *-a* is exclusive to neuter gender.

ral; what may change is the meaning of the whole [root [n]] complex. All we need to say is that, when a root is combined with [n:fem], a certain reading arises. For some roots, this requires [Num:pl].

Importantly, the meaning of these plurals is neither provided by the root alone nor by [n] alone. As can be seen from the simple overview in (20), the lexical meanings of these nouns do not share any common property independent of plurality: we go from uncountable masses to perfectly countable measures (in fact, standards of countability), passing through notions like substances with manifold appearance (brains, intestines), mutually connected parts (walls in a perimeter, body parts, shouts, peals of laughter), and 'eggs':

(20) <i>midolla</i> 'marrow'	elements of a mass (non-count)
<i>mura</i> 'walls as closed perimeter'	elements making up a whole (non-count)
<i>ossa</i> 'bones as connected parts'	elements naturally related (partly count)
<i>uova</i> 'eggs'	elements of uniform appearance (count)
<i>miglia</i> 'miles', <i>braccia</i> 'arm-lengths'	abstract units of measurement (count)

To find out what ties together these notions, we must consider both lexical semantics and plurality. The common trait is that, when plural, these nouns describe pluralities of weakly individual entities. This encompasses measures, which are indistinguishable because they are abstract equivalence classes, 'eggs', conceptualized as empirically indistinguishable, discrete parts of a whole, seen as tokens whose distinctive identity is secondary with respect to a larger entity they are parts of, and substances, whose elements may be perceptually salient but cannot be uniquely segmented and identified. The lack of individuality presupposes plurality, and so cannot be a property of the root alone; but it also presupposes the lexical meaning of each noun, and so cannot be a property of number (or gender) alone. In particular, there is no 'collective' plural suffix: measures and eggs are not collective in any sense.

These plurals, then, are a clear illustration of the way 'lexical' meaning arises in a syntactic construction. Their meaning, as well as the treatment of gender they justify, is direct evidence for a constructional approach.

### 5. Roots do not coincide with L-nodes

The previous section has shown how root-governed phenomena may be analyzed without direct marking of roots. The next step is a precise specification of the definitional properties of roots in this framework. To achieve that, I distinguish first of all the terminal node occupied by a root from the Vocabulary item that spells it out. This presupposes that roots, like grammatical morphemes, are spelled out after syntax, contrary to what recently claimed by Embick and Noyer (2007). In the absence of evidence to the contrary, there seems to be no reason for exempting roots from Late Insertion.

Following Harley and Noyer (2000), I will call L-node the node realized by a root. For Harley and Noyer, the definitional property of L-nodes is that their spellout is non-deterministic: while the grammar uniquely determines a realization for the feature bundle defining a grammatical morpheme, the insertion of one root or another (say, DOG or CAT) depends on the speaker's choice. The only grammatical constraint is that roots must appear in a compatible licensing context.

The hypothesis that roots are by definition category-free would seem to be a natural additional property of L-nodes. But a one-to-one relation between L-nodes and category-free syntactic terminals, while possible, is not conceptually necessary. In fact, there are reasons for decomposing the notion of root into three distinct concepts, based on properties that do not necessarily overlap:

(21) Root-as-node ≠ Root-as-exponent ≠ Root-as-category-free domain:

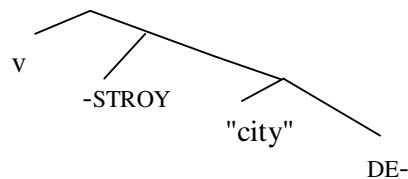
Root =<sub>def</sub> Vocabulary item inserted in an L-node

L-node =<sub>def</sub> terminal allowing non-deterministic Vocabulary insertion

Root domain =<sub>def</sub> syntactic domain inside the first category-assigning head

The suggestion that the lack of category may characterize domains larger than a single L-node is not new. In particular, Marantz (2001) already conjectured that a transitive verb like *destroy* is based not on a simplex root, but on a complex made up of a manner-naming root STROY, a transitive particle DE-, and a [v]: 'Thus, STROY is a manner root that incorporates a particle, spelled out DE, that takes an "inner subject" as the direct object of the syntactically derived verb "destroy"':

(22) Complex root domains: Marantz (2001):



I will not follow Marantz' assumptions that roots like STROY "name" a manner (a notion I find unclear), and that roots and root-like particles head syntactic projections that host arguments in their specifiers. But I accept what I take to be the main point of his analysis, namely that category is assigned not just to simple, atomic category-free heads, but possibly to complex structures. In (22) this consists of the root STROY and the particle DE. The latter is not an open-class element, and so differs from the genuinely open-class STROY; but it is not a preposition or a verb either. It is a root extension, a modifying element that attaches to a root to derive a complex root. It is the whole complex root that is assigned a category; but then what is assigned a category can be a syntactic domain as well as a single head. The atomic elements of this syntactic domain are not necessarily all L-nodes, although this option is open when a language allow root compounds made up of two open-class roots, as argued by Zhang (2007) for Chinese. Viewing the lack of category as a property not of roots per se, but of a potentially non-atomic syntactic domain, opens up the possibility to analyze the internal structure of words like *re-ceive* / *re-ceipt* or *con-ceive* / *con-cept* without having to claim an unlikely verbal or nominal status for *re-* or *con-*, or an invariant semantic content that, as Aronoff (1976) noted, these element simply do not have.

On the other hand, the clear distinction between L-nodes and roots as exponents (in fact, Vocabulary items, which index exponents to morphosyntactic information) has another welcome consequence. It allows us to envisage cases where the same Vocabulary item is inserted in both L- and F-nodes. This seems the correct characterization for what are often called "semi-lexical" categories, which arise when a "lexical" open-class morpheme has an additional use as a grammatical morpheme. A few examples are given in (23); see also Broschart (2000) for Tongan preverbials, and Contini-Morava and Tobin (1999) for a broad-ranging overview:

(23) a. nouns used as classifiers:

hai cái <u>bao</u>	hai <u>bao</u> cam	
two thing bag	two bag orange	
'two bags'	'two bagfuls of oranges'	(Vietnamese: Löbel 2000)

b. lexical verbs used as auxiliaries:

German *werden* 'to become' ~ passive auxiliary

Marie wurde zu einer feiner Dame	'Marie became a fine lady'
Marie wurde von allen bewundert	'Marie was admired by all'

Irish *caith* 'to throw, to consume' ~ modal auxiliary

caitfidh mé an bhliain seo sa Ghearmáin	'I will spend this year in Germany'
caitfidh mé dul go dtí an Ghearmáin	'I must go to Germany'

c. nouns used as prepositions:

German *laut* 'sound' ~ 'according to', *kraft* 'force' ~ 'by the authority of'

die Musik ist zu laut	'the music is too loud'
laut meines Vertrages	'according to my contract'
mit voller Kraft	'with full force'
kraft ihres Amtes	'by the authority of her office'

Coupled with the hypothesis of complex root domains, the idea of open-class roots playing a grammatical role also accounts very naturally for complex adverbials / prepositions like the English *a-long*, *a-part*, or *a-side* (a reviewer also suggests *with-out*, which however does not include a noun; the two cases are related but not identical). Like *de-* is a closed-class particle in the complex root DE-STROY (which underlies *to destroy*, *destructive* and *destruction*; cf. *to instruct*, *instructive* and *instruction*), *a-* is a particle in the complex A-SIDE, where the open-class root *side* forms a preposition and not, as is customary, a noun.

Most importantly, the clear distinction between L-nodes and root Vocabulary items allows us to view class diacritics as a property of the latter, but not of the former. Here I have in mind specifically those diacritics that play no role in syntax, like declension or conjugation class. The arguments in sections 2–4 above militate against associating information with the root as the content of a particular L-node. What must be category-free, according to the strict version of the lexical decomposition hypothesis I adopt here, is a root as the abstract content of a syntactic head, not a root as an exponent. Once Vocabulary insertion takes place, there is no reason why the form that spells out an L-node should not carry morphological diacritics. If these presuppose a category, then that Vocabulary item is restricted to contexts that define that category. In fact, class diacritics are basically insensitive to the distinction between lexical and grammatical elements. In Italian, for example, the nominal classes defined by the endings *-o / -i* (masc. sg. / pl.) and *-a / -e* (fem. sg. / pl.) apply not only to nouns and adjectives, but also to some demonstrative determiners and articles, which are grammatical morphemes that involve no L-node and no category-assigning head, [n] or [adj].<sup>3</sup> The property of falling into one of these classes obviously cannot depend on the idiosyncratic marking (or selectional restrictions) of a lexical category. This is expected if purely morphological diacritics are not the property of the contents of nodes, but of the forms realizing them. This separation, in the

sense of Beard (1995), is a basic feature of Distributed Morphology (as Halle and Marantz 1993 explicitly noted); distinguishing L-nodes from root Vocabulary items simply follows from this separationist stance.

### 6. *L-nodes are indices*

The conception of roots here outlined differs from the received view in two main respects. First, the substantive role of roots is examined here not only in connection with verbs and verb-based nominalizations, as is commonplace, but for all lexical categories. Only by going well beyond verbal / eventive "core" meanings can we appreciate how problematic it is to posit a semantic core shared between entity- and non-entity-denoting linguistic objects. Secondly, I have emphasized the difference between roots as abstract elements in a morphosyntactic representation and roots as unanalyzable forms. These two lines of the argument converge on the same question: what is the content of roots as L-nodes. Having gone as far as claiming that roots have no self-standing meaning, I must say something about what content they have at all, once we abstract away from their form.

To begin with, if one assumes that roots lack not just syntactic category, but all syntactic features, then they are invisible for syntactic operations.<sup>4</sup> Lacking syntactically legible information, they cannot project: there can be, then, no 'RootP', and no argument may therefore appear in the specifier or complement position of a root. Only functional heads take complements — presumably including root-extending particles, which are category-free but do not correspond to L-nodes. Once again, taking this hypothesis seriously leads to a position that is in contrast with most current work in Distributed Morphology. But the logic is inescapable, and the conclusion follows at once if (1) the content of L-nodes has no syntactic features, and (2) syntactic projection arises by merging elements and assigning to the resulting complex the features of one of them. Nor is the argument a novel one: Zhang (2007) has shown that several empirical properties of Chinese root compounds follow from the simple idea that roots are syntactically inert.

One may take this to mean that syntax does not "see" roots at all. For syntactic purposes, every root would then be like every other, and all would count as the one and the same place-holder (cf. Halle 1990 for an early approach along these lines). This is not the position I take, however. The main reason has to do with the architecture of grammar. If Vocabulary insertion takes place on the way to PF, as a preliminary to phonological interpretation, it does not interact with the rest of the derivation, which delivers a complete structure (or a piece corresponding to a Phase) to interpretation at the conceptual-intentional interface. The way a root is morphologically spelled out, then, should not matter for meaning. But distinct roots routinely define distinct meanings within exactly the same syntactic structure, like *cat* versus *dog*, or *to paint* and *to draw*. If syntax sees both of these as the same empty place-holder, and morphological spellout takes place in a separate branch of the derivation from that which leads to semantic interpretation, there can be no way of telling apart syntactic structures that differ only by the choice of a root. But these structures do differ semantically. Therefore, L-nodes are distinct in the syntax.

Other reasons for taking L-nodes to be differentiated before Vocabulary insertion are more theory-internal, and weaker. For example, assuming that certain roots are extended by a category-free particle, and that the node hosting this particle is created in the syntax, then syntax must distinguish between one L-node and another, since not all roots have this type of extension. That is, the node labelled -STROY in (22) must be identifiable, because not all roots fit in that structure. Another theory-internal reason is that certain morphological rules that apply before spellout are sensitive to the choice of one root instead of another, before this is inter-



preted by a Vocabulary item. In (15) in section 3 above, [n] and [Num] must undergo Fusion in the context of certain roots only (cf. Embick 2003 for a similar treatment of root-conditioned allomorphy).

This is not the kind of evidence that would make anyone change his or her mind, however. Therefore, I will just state the claim that L-nodes are different before Vocabulary insertion (not an assumption, because it is motivated by architecture), and ask how this can be reconciled with the previous claim that roots have no autonomous meaning. How can they be different in abstract content, if they have no content?

This fundamental question brings us back to the issue of root-related formations like *to hammer* and *a hammer*. The problem, it will be recalled, is that neither the verb's lexical semantics entails reference to the noun, nor conversely. The shared root ensures a relation, but this is semantically so tenuous that it cannot be identified without reference to an entity or an action, which correspond to the noun and the verb respectively. We see a relation between the two meanings, but we cannot pin it down without falling back onto one of the two meanings.

There are very good reasons for this state of affairs, as already indicated. Any kind of interpretation requires semantic typing, that is, a determination of the category of semantic primitives that an expression refers to; traditionally, the fundamental types for formal semantics are entity and truth-value (and possibly situation). It is certainly possible to assume that roots, as minimal interpretive units, uniformly denote entities, and derive all other interpretations through syntactic composition with other heads. I would like instead to explore another avenue, and take at face value the time-honoured intuition that semantic type correlates with grammatical type. If roots lack the latter, they should not have the former either. The minimal units of interpretation are those that define a semantic type, and these are *not* roots, but core nouns and verbs. Roots are smaller; in this sense, they have no meaning by themselves but co-occur with category-assigning heads to form interpretable typed grammatical entities. But how does DOG differ from CAT, then, if both have the structure [root + n]?

My answer is that the root DOG acts as an index that makes the noun *dog* different from nouns based on other roots. In the abstract syntactic representation before Vocabulary insertion, roots do not mean anything by themselves, but act as name-tags which define identity and difference. Their function is differential, not substantive, like that of the indices 1 and 2 in *he<sub>1</sub> likes broccoli, but he<sub>2</sub> doesn't*. What is crucial before spellout, as we have seen, is that the root nodes DOG and CAT should be distinct; they don't have to mean 'dog' and 'cat', because for that, they would have to be categorized as nouns. Rather, we notate as DOG the abstract element which fills the root node in the representation underlying the noun *dog*, as well as the verb *to dog*, the adjective *dogged* and the adverb *doggedly* (assuming for the sake of illustration that they are lexically related); and with CAT the element filling the root node in the noun *cat*. Identity of root signals lexical relatedness, but not a shared meaning that might be isolated from its context. An interpretation only arises when constructs are assembled and become interpretable as kinds of entities, predicates, states, or activities. A noun and a verb that share the same root, but only the root, will share as it were the same name-tag while denoting distinct types of referents, as schematized in (24a–b); a pair like *to tape* and *tape*, on the other hand, shares a larger domain consisting of root and [n], which derives the transparently denominal semantic relationship illustrated in (24c–d):

(24) Roots as indices for the minimal units of semantic interpretation:

- |    |                     |                       |   |
|----|---------------------|-----------------------|---|
| a. | [[ HAMMER ] v ]:    | 'to hammer-do'        | (action labelled <i>hammer</i> )            |
| b. | [[ HAMMER ] n ]:    | 'a hammer-thing'      | (kind of entity labelled <i>hammer</i> )    |
| c. | [[[ TAPE ] n ] v ]: | 'to use a tape-thing' | (action on an entity labelled <i>tape</i> ) |
| d. | [[ TAPE ] n ]:      | 'a tape-thing'        | (kind of entity labelled <i>tape</i> )      |

Roots are the names that, attached to elements which specify a type, identify one particular concept belonging to that type. The template [[ROOT] n] specifies entity-referring concepts (including nominalized properties like *red* in *red is a colour*); substituting DOG or CAT for the root defines two different entity-referring concepts, which underlie the full-fledged nouns *dog* and *cat*. As noted by Aronoff (1976), we do not need to contrive meanings for RE-, CON-, PER- and -CEIVE; these are just morphological labels, including the open-class -CEIVE, which serve to name the different concepts that surface as the verbs *receive*, *conceive*, and *perceive*. The approach to roots I have outlined enables us to express this important insight of word-based morphology within an item-and-arrangement, piece-based approach like Distributed Morphology.

## 7. Conclusion

This sketch is far from being a theory of roots, in Distributed Morphology or otherwise. Its purpose has been rather to follow the theoretical implications of the lexical decomposition hypothesis, under a strict reading of the idea that roots are radically underspecified. In this perspective, I have argued that roots are no exception to Late Insertion, and that therefore we must distinguish between the morphosyntactic terminals in which they appear and their morphological exponence. The radical underspecification applies specifically to the abstract content of L-nodes, not to the Vocabulary items realizing them. In addition, it does not apply necessarily to one node, but possibly to a complex domain, bringing about category-free extended roots. As we have seen, root-dependent phenomena do not stand in the way of the idea that abstract roots have no diacritics, because alternative and arguably superior analyses are available. Finally and most importantly, I have argued that abstract roots are mutually distinct but carry no identifiable semantic interpretation. On their own, they are just differential indices that fix the identity of larger constructs, which constitute the minimal units for semantic interpretation.

If this is on the right track, we can understand better why roots need to be assigned a category in the first place (the "Categorization Assumption" of Embick and Marantz 2008). In part, this is because they could not otherwise interact with the generative grammatical system (assuming that this is driven by features, and that roots lack features). But a still deeper reason would be that roots must form linguistic categories in order to form conceptual categories. In this sense, the present proposal is radically constructional: not only words, but even the minimal units of meaning arise through linguistic construction.

Many questions inevitably arise; to conclude, I would like to point to a few distinctly problematic aspects of this proposal. To begin with, expressing the gender of a noun as a statement of co-occurrence between a root and a certain content for [n] is not particularly enlightening or predictive. The approach I have outlined is designed to allow for multiple categorization of the same root, but it does not provide a formal expression of the stable association of most roots with gender, class, or other diacritics. This seems to be a weakness, in so far as a noun's idiosyncratic marking should have a formal grammatical encoding, as distinct from a mere statement. Secondly, every analysis where gender features feed morphological realization faces the problem posed by cases where gender depends on the phonological shape of an exponent. The problem is particularly acute in a framework like Distributed Morphology, because abstract features cannot be licensed or copied after Vocabulary insertion. Finally, there is at least one theoretical difficulty specific to Distributed Morphology, and that concerns Impoverishment (the deletion of features from the input to morphological realization, to be replaced by the corresponding unmarked values). In Noyer's (2005) analysis of interclass syncretism in nominal paradigms, class diacritics may be deleted by Impoverishment. This

means that they must be there *before* Vocabulary insertion, contrary to my claim that what has class diacritics is a Vocabulary item, and not an abstract terminal. My discussion has centred on roots, however, and it remains possible that diacritics appear as abstract information on non-L-nodes (typically [n] for nouns). Perhaps, as a reviewer suggests, the solution lies in the cyclic nature of Vocabulary insertion, which takes place from the root outwards: a root node is spelled out by an exponent; this exponent, suppose, is only compatible with a certain choice of diacritics on the adjacent [n], and these diacritics undergo Impoverishment (under specified circumstances) before [n] is spelled out. Like many other questions concerning the relation between class features and morphological exponence, this, clearly, requires much "future research".

### *Acknowledgements*

I would like to thank the organizers and the audience of the Fifth York–Essex Morphology Meeting, as well as two anonymous reviewers. This research was supported by a Research Fellowship of the Alexander von Humboldt Stiftung, which I gratefully acknowledge.

### *Notes*

1. Alexiadou (2001:58) explicitly treats roots as having the interpretation of states. This is more restrictive than treating them as 'concepts' that hardly differ from verbs, but still presupposes a semantic categorization that is clearly more verbal than nominal, more predicate- than entity-oriented. The root of a simple noun like *birch* might well have the semantic value of a predicate, but it's hard to see in what sense it is a state, unless 'state' means just 'predicate' (in which case ' $\lambda x$  [ happen (x) ]' is also a state).
2. In fact, it is not completely correct that regularized alternants of strong verbs are always based on a noun: cf. *she hanged herself / was hanged in her cell* (instead of *hung*), where there is no noun *\*a hang* meaning 'execution by hanging'. A reviewer also notes *to shine / shone* (intransitive), but *to shine / shined* (transitive causative, as in *she shined her shoes*).
3. Contrast Halle and Marantz (1994:282): 'In the morphology, as a consequence of a morphological well-formedness condition [in Spanish], a Theme is adjoined to the DET node. In fact, a Theme node is adjoined to every lexical category node in Spanish.' In so far as there is a distinction between lexical and non-lexical categories, nouns fall into the former and articles into the latter. But introducing diacritics by the means of an added Theme head forces one to stipulate this additional head precisely where it is needed.
4. Thanks to Natascha Pomino for discussion on this point.

### *References*

- Acquaviva, P. (2002). Il plurale in *-a* come derivazione lessicale. *Lingue e linguaggio* 2, 295–326.
- Acquaviva, P. (2008). *Lexical Plurals*. Oxford: Oxford University Press.
- Alexiadou, A. (2001). *Functional Structure in Nominals*. Amsterdam – Philadelphia: Benjamins.

- Alexiadou, A. (2004). Inflection Class, Gender and DP Internal Structure. In G. Müller, G. Lutz, and G. Zifonum (Eds.), *Explorations in Nominal Syntax*, pp. 21-50. Berlin – New York: Mouton de Gruyter.
- Alexiadou, A. and G. Müller. (2008). Class Features as Probes. In A. Bachrach and A. Nevins (Eds.), *Inflectional Identity*, pp. 101–155. Oxford: Oxford University Press.
- Alexiadou, A., L. Haegeman, and M. Stavrou. (2007). *Noun Phrase in the Generative Perspective*. Berlin – New York: Mouton de Gruyter.
- Arad, M. (2003). Locality Constraints on the Interpretation of roots. *Natural Language and Linguistic Theory* 21, 736–778.
- Aronoff, M. (1976). *Morphology in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Beard, R. (1995). *Lexeme-Morpheme Base Morphology*. New York: SUNY Press.
- Broschart, J. (2000). The Tongan Category of Preverbials. In P. Vogel and B. Comrie (Eds.), *Approaches to the Typology of Word Classes*, pp. 351–370. Berlin –New York: Mouton de Gruyter,.
- Contini-Morava, E. (1999). Noun Class and Number in Swahili. In E. Contini-Morava and Y. Tobin (Eds.), *Between Grammar and the Lexicon*, pp. 3-29. Amsterdam – Philadelphia: Benjamins.
- Contini-Morava, E. and Y. Tobin (eds.). (1999). *Between Grammar and the Lexicon*. Amsterdam – Philadelphia: Benjamins.
- Don, J. (2004). Categories in the Lexicon. *Linguistics* 42(5), 931–956.
- Embick, D. (2003). Locality, Listedness, and Morphological Information. *Studia Linguistica* 57(3), 43–169.
- Embick, D. and M. Halle. (2005). On the Status of *Stems* in Morphological Theory. In T. Geerts, I. van Ginneken and H. Jacobs (Eds.), *Romance Languages and Linguistic Theory 2003*, pp. 37–62. Amsterdam – Philadelphia: Benjamins.
- Embick, D. and A. Marantz. (2008). Architecture and Blocking. *Linguistic Inquiry* 39(1), 1-53.
- Embick, D. and R. Noyer. (2007). Distributed Morphology and the Syntax-Morphology Interface. In . G. Ramchand and C. Reis (Eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 289–324. Oxford: Oxford University Press.
- Galani, A. (2004). So, Word Formation: Syntax or Morphology? Formation of Non-Suppletive Stems in Modern Greek. *Proceedings of the III International Conference of Greek Linguistics*, University of Crete.
- Hale, K. and S. J. Keyser. (1992). 'The Syntactic Character of Thematic Structure. In I. Roca (Ed.), *Thematic Structure: Its Role in Grammar*, pp. 107-143. Berlin: Foris.
- Halle, M. (1990). An Approach to Morphology. In *NELS 20*, pp. 150-184. University of Massachusetts, Amherst.
- Halle, M. and A. Marantz. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale and S. J. Keyser (Eds.), *The View from Building 20*, pp. 111–176. Cambridge, Massachusetts: MIT Press.
- Halle, M. and A. Marantz. (1994). Some Key Features of Distributed Morphology. In A. Carnie, H. Harley, and T. Bures (Eds.), *MIT Working Papers in Linguistics 21: Papers on Phonology and Morphology*, pp. 275-288.
- Harley, H. and R. Noyer. (2000). Formal versus Encyclopedic Properties of Vocabulary: Evidence from Nominalisations. In B. Peeters (Ed.), *The Lexicon-Encyclopedia Interface*, pp. 349–374. Amsterdam: Elsevier.
- Harris, J. (1996). The Syntax and Morphology of Class Marker Suppression in Spanish. In K. Zagana (Ed.) *Grammatical Theory and Romance Languages*, pp. 99–122. Amsterdam – Philadelphia: John Benjamins.

- Harris, J. (1999). Nasal Depalatalization *no*, Morphological Wellformedness *sí*; the Structure of Spanish Word Classes. *MIT Working Papers in Linguistics 33: Papers on Syntax and Morphology, Cycle One*, pp. 47–82.
- Josefsson, G. (2001). The Meaning of Lexical Classes. *Nordic Journal of Linguistics 24*, 218–231.
- Kihm, A. (2005). Noun Classes, Gender, and the Lexicon-Syntax-Morphology Interfaces: A Comparative Study of Niger-Congo and Romance Languages. In G. Cinque and R. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax*, pp. 459–512. Oxford: Oxford University Press.
- Kiparsky, P. (1997). Remarks on Denominal Verbs. In A. Alsina, J. Bresnan and P. Sells (Eds.), *Argument Structure*, pp. 473–499. Stanford: CLSI.
- Lecarme, J. (2002). Gender "Polarity": Theoretical Aspects of Somali Nominal Morphology. In P. Boucher and M. Plénat (Eds.), *Many Morphologies*, pp. 109–141. Somerville, Massachusetts: Cascadilla Press.
- Löbel, E. (2000). Classifiers versus genders and noun classes: A case study in Vietnamese. In B. Unterbeck, M. Rissanen, T. Nevalainen and M. Saari (Eds.), *Gender in Grammar and Cognition*, pp. 259–319. Berlin – New York: Mouton de Gruyter.
- Lowenstamm, J. (2006). On *Little n*, and Types of Nouns. Ms, Université Paris 7.
- Marantz, A. (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In A. Dimitriadis, L. Siegel, C. Surey-Clark, and A. Williams (Eds.), *Proceedings of the 21st Annual Penn Linguistics Colloquium. University of Pennsylvania Working Papers in Linguistics 4(2)*, pp. 201–225.
- Marantz, A. (2001). Words and Things. Ms., MIT.
- Noyer, R. (1997). *Features, Position and Affixes in Autonomous Morphological Structure*. New York: Garland.
- Noyer, R. (2005). A Constraint on Interclass Syncretism. *Yearbook of Morphology 2004*, pp. 273–315. Springer: Dordrecht.
- Oltra-Massuet, I. (1999). *On the Notion of Theme Vowel: A New Approach to Catalan Verbal Morphology*, MIT occasional papers in linguistics 19.
- Pinker, S. (1999). *Words and Rules*. London: Weidenfeld and Nicolson.
- Trépos P. (1957). *Le pluriel breton*. Rennes: Imprimeries Réunies.
- Zhang, N. (2007). Root Merger in Chinese Compounds. *Studia Linguistica 61(2)*, 170–184.

*School of Languages and Literatures*  
*Newman Building*  
*University College Dublin*  
*Belfield, Dublin 4*  
*email: [paolo.acquaviva@ucd.ie](mailto:paolo.acquaviva@ucd.ie)*  
*Paolo Acquaviva*