1. INTRODUCTION: THE ISSUE OF ACATEGORIAL ROOTS?

Mental lexicon: core component of natural language & essential part of any linguistic theory, be it a traditional one of packaged parcels of phonological, semantic, and formal information or decomposed into a series of lists as in DM. ‘Root’, in turn, is an indispensable part of the mental lexicon; again, of a traditional and (even more so) a DM one. Under either view, the notion of morphological root is underlined by the intuition that words may share a minimal - bare ‘core’; that which remains invariant when all identifiable morphological formatives have been taken away. Although the term 'root' has traditionally designated a descriptive morphological category, in Distributed Morphology (as in many morphological theories), the term names a theoretical construct which plays an important role in the framework.

Though we will discuss the nature of roots in general, our primary focus will be on empirical domain of cranberry items. In other words, the conclusions we draw about them are merely suggestive for the broader domain of roots.


Lexical roots become nouns, verbs or adjectives in syntax as a result of merger with n, v or a functional heads, which can be phonologically null and whose grammatical content defines a nominal, verbal or adjectival domain:

\[
\text{CAT} \begin{array}{c}
\text{n} \\
\text{Ø}
\end{array}
\]

Thus in DM roots have
- no morphosyntactic category
- no gender or declension/conjugation class (Acquaviva 2009, following Harris 1996, though with caveats, for Harris roots have category)
- no phonology
- for some researchers (e.g., Pfau 2000, 2009, Acquaviva 2009, Harley 2014), no meaning

“On their own, roots are unpronounceable. It is “words”—roots combined with nominal, adjectival, or verbal features—that we pronounce. Roots also lack a fixed or precise semantic interpretation. It is only in the specific environment of certain morphemes that they acquire an actual interpretation as nouns or verbs. The root \sqrt{}hammer, for example, is assigned an interpretation of a manner verb when embedded in a verbal environment, and an interpretation of an instrument used for hammering when embedded in a nominal environment.” (Arad 2003:10)

2. CRANBERRY MORPHEMES AS A CLUSTER OF PHENOMENA

“…because cran is completely isolated from the syntax by its occurrence inside only the one word, there is no way in which it can have syntactic properties of its own, and hence semantic properties” (Aronoff 1974:38)

Lexicon of Linguistics: “a type of bound morpheme that cannot be assigned a meaning nor a grammatical function, but nonetheless serves to distinguish one word from the other.

EXAMPLE: the English word cranberry seems morphologically complex, since it must be distinguished from words such as raspberry, blackberry, and gooseberry. Still, cran has no meaning and does not function as an independent word: cranberry is the only word in which cran appears.”
Usual examples of cranberry morphemes:

(2)  
a.  

\textit{cran-} (cranberry)  

b. \textit{-mit} (permit, transmit, commit, remit, submit, resubmit, recommit, intromit, intermit, remit, but also admit, demit, emit, omit)  

c. \textit{-ceive} (receive, perceive, deceive, conceive, apperceive, misconceive, misperceive, preconceive…)  

d. \textit{cob-} (cobweb)  

e. \textit{were-} (werewolf, werebear, wereboar, wererat, weretiger)  

f. \textit{-kemp(t)} (unkempt)  

g. \textit{ruth-} (ruthless)  

h. \textit{reck-} (reckless, from OE reccelēas)

What these morphemes have in common: they are bound. That’s it.

2.1. Lack of meaning

Not necessary: \textit{ed-} (edible), \textit{cob-} and \textit{-ept} are quite transparent

2.2. Hapax morphemes and lexical category

Cranberry morphemes in (2b-c, e) are not restricted to just one environment. Even \textit{cran} isn’t: there is \textit{crantini} (a martini-like drink with cranberry juice) and some other coinages

When there are several environments of use, the lexical category can still remain unclear:

(3)  
a. \textit{nutrient}, \textit{nutricious}, \textit{nutrition}  

b. \textit{terror}, \textit{terrible}, \textit{terrific}, \textit{terrify}  

c. \textit{horror}, \textit{horrible}, \textit{horrific}, \textit{horrify}, \textit{horrid}, \textit{horrendous}  

d. hind, fore

NB: (3b,c) are extremely close semantically, yet their distribution is limited differently, unlike in (3d)

Conversely, a morpheme can be single-use (hapax) and still clearly have a lexical category:

(4)  
a. \textit{ruth-}, \textit{reck-}, gorm—: \textit{-less} adjoins only to nouns  

b. \textit{(il)leg-}, \textit{plaus-}, \textit{ed—}: \textit{-ible} is an allomorph of \textit{-able} and only adjoins to verbs (cf. \textit{permissible}, \textit{admissible}, \textit{accessible}…)  

c. \textit{-ept}: the negative \textit{in-} is an adjectival prefix

Hence there is no direct connection between the underdetermination of a lexical category and the hapax status

When the distribution is not limited, is the category clear from it?

2.3. Compounding vs. derivation

The first member of an English nominal compound can belong to any lexical category (from Andrew McIntyre’s handout):

(5)  
a. [N N]: chess table, strawberry jam, diesel motor, bookshelf  

b. [V N]: crybaby, scrubwoman, bakehouse  

c. [Participle N]: filing cabinet, reading class, writing table, drinking water  

d. [A N]: blackbird, drydock, redbrick, wetsuit  

e. [Particle/Preposition N]: outhouse, outgrowth, undergrowth, offprint  

f. [N A]: bloodthirsty, pain-free, theory-neutral, colourblind, class-specific

So a cranberry morpheme as the first member of a compound can be of any lexical category... in principle
A lexical-semantic class of cranberry morphemes: *Monday, Tuesday*…

These are nouns: the only example of a *day*-compound that has a non-nominal first member is *holiday* and it is also the only one that is not a day

The fine line between compounding and derivation with cranberry morphemes is not clear:

(6)  
   a.  behind, hindmost, hindsight  
   b.  before, foremost, foreleg, foresight…

Probably, nouns

*On cranberry* itself see the appendix

Suppose we find a better example. What then?

2.4. Root vs. affix

So far we have been looking at limited-distribution roots

But there are limited-distribution and even **hapax affixes**:

**NB:** For our purposes, the fact that hapax-affixed constituents can serve as input for subsequent derivation does not affect their status: after all, such morphemes still only occur in the context of a single root

(7)  
   a.  -uz-: only used in *franc-uz* ‘a Frenchman’, cf. *Francija* ‘France’  
   b.  -ës-: only used in *bel-ës-yj* ‘offwhite, whitish’, cf. *bel-yj* ‘white’  
   c.  -s-: only used in *plak-s-a* ‘crybaby’ from *plak-at* ‘to cry’; marginally there’s also *xnyk-s-a* from *xnyk-at* ‘to snivel’ and *krik-s-a* from *kryč-at* ‘to yell’, with the same semantics of habituality

(8)  
   a.  diev-egge ‘female thief’, cf. *dief* ‘thief’  

All of these do have a lexical category, so once again, being hapax does not mean acategorial

2.5. Free cranberry morphemes

Single-use morphemes need not be bound (the so-called **fossil words**, or cranberry words (see Aronoff 1974))

(9)  
   a.  **kith** and kin  
   b.  **ekte** (only in *ekte out*)  
   c.  **spick** and span

Fossil words can be functional (e.g., *to and fro, hither and yon*)

And be restricted in their use as well:

(10)  
   a.  **hark** (*hark back to or hark at you*)  
   b.  **hither** (*come hither, hither and thither, hither and yon*)  
   c.  **mettle** (*in high mettle, test/prove/show (one’s) mettle, on (one’s) mettle*)

And, very rarely, their category is underdetermined:

(11)  
    **days of yore** (cf. *Days of Future Past vs. days of old*)

There might be some fossil words underdetermined for category. So what? **They must have a category, this is syntax!**

2.6. The issue of phonaesthemes

Firth 1930, Bolinger 1950, Bergen 2004: frequent sound-meaning pairings that are clearly not morphemes (though see Rhodes and Lawler 1981):
(12) a. *gl-* ‘light, vision’: glimmer, glisten, glitter, gleam, glow, glint, etc.
    b. *sn-* ‘nose, mouth’: snore, snack, snout, snarl, snort, sniff, sneeze, etc.

For (12a): these are all verbs. So:

- shared sound
- shared meaning (the lexical-semantic class of the totality)
- shared category
- NOT a morpheme

Why?

- phonaesthemes are not syllabic (but then neither is the past tense -d)
- English prefixes are not category-determining heads (with the possible exception of *en-* in *envisage*, *enliven*, etc., though probably not *enjoy*; see Hammond 1993)

Making *gl-* a morpheme would lead to an impossible syntactic structure

Hypothesis: it is having lexical category that determines whether something is a morpheme

2.7. **Intermediate summary**

The lexical indeterminacy of cranberry morphemes is hugely exaggerated

Limited distribution is not a property that is only relevant for morphology: fossil words are free, but have limited distribution

The fact of single use does not exclude having lexical category and multiple environments of use do not ensure it

The fact that the lexical category of a morpheme cannot be determined does not entail the lack of a lexical category (cf. fossil words)

In fact, assuming that morphemes must have a lexical category explains why phonaesthemes are not morphemes

Even the fact that a morpheme might have a certain meaning in isolation does not entail that this meaning is retained in composition (cf. *handsome*, *understand* – see Aronoff 1974:37)

3. **AGAINST ACATEGORIAL COMPOUNDING**

**Cranberry morphemes do have a lexical category** (though it cannot be always determined unambiguously)

Compounding doesn’t require bare roots as input:

- there are compounds where the lexical category of the members is clear
- in languages with overt morphology are there any others?

3.1. **Overt morphology in compounds**

Argument: category-specific morphology requires the presence of the category

3.1.1. **Imperative compounds in SC**

In all the instances of these compounds, the first element is a verb in the 2PsSg Imperative form. Whatever the analysis of their origin etc., what is important for us is that these cases of compounding clearly do not involve cannot be category-less roots.

The origin, semantic types, and use of these compounds have attracted a great deal of interest in the literature on SC (see Daničić 1876, Stevanović 1956, Klajn 2002, Progovac 2006, among others). Though interesting, these issues are outside of the domain of this paper. What is relevant for us here is that the first element of these
compounds cannot be a bare root – it must be a verb as it appears in the form of 2SG imperative, ending in either ‘i’ or ‘j’, depending on the verb. In SC, -o- and -e- are the linkers, -i- is not

(13)  

a.  

\text{vuci-batina (pull-whip) (vuci.IMP, as opposed to vuci.INF, vuče.3SG.Present, vuK.ROOT, as evident in vuk-ao.PastParticiple)}

b.  

\text{seći-kesa (cut-purse) (seći.IMP, as opposed to seK.ROOT (e.g. trbo-sek (stomach-ripper) sek-aeć (cut-ter), seći. INF seče.3SGPresent)}

c.  

\text{deri-koža (rip-skin) (deri.IMP, as opposed to derati.INF, dere.3SG.Present, der.ROOT, as in oblak-o-der (skyscraper(er)) (Progovac 2006: Ex7)}

d.  

\text{some proper names: Branimir (defend-peace), Želimir (want-peace), Budislav (be-glorious); Branislav (defend-glory)}

There are no compounds that allow verbal roots as a first member and yet show no verbal morphology

3.1.2. Infinitive-containing compounds in Dutch

Though the morphemes \text{en} and s (homophonous or identical to plural markers) are attested in compounds in Dutch, here -s- is unambiguously a linker and -en-, the infinitive marker:

(14) \text{eten-s-tijd (eating time), varen-s-man (sea farer), zien-s-wijze (lit. seeing way, view); uitgaan-s-verbod (lit. outgoing prohibition, curfew)}

For further examples, overview, and discussion of compounding in Dutch, see Booij 1992

3.1.3. Nouns & adjectives as non-heads

If it is plural, it has to be nominal:

(15) \text{English: mice-infested, menbashing, oxenpower (ox-oxen), dice thrower (die – dice)}

Though an intriguing topic in its own right, the issue of the restriction on regular plurals to take part as the first member of the compounds is orthogonal to our discussion here. The reader is referred to Berent and Pinker 2007 for discussion and references. Upshot: “phonological unfamiliarity in general and phonological properties of regular plural in particular, do not taint the acceptability of compounds […] the morphological regularity of a non-head does taint the acceptability of compounds when phonological familiar and plural semantics is controlled” (Berent and Pinker 2007:135-136)

(16) \text{Serbo-Croatian suppletive plurals}

a.  

\text{dec-o-ubica (child-murderer): dete (child) – deca (children)}

b.  

\text{ljud-o-žder (cannibal): čovek (man) – ljudi (people - suppletive)}

c.  

\text{viš-e-grad (toponym, lit. higher-town): visok (high/tall) – viši- (taller)}

Similar facts in Russian

3.1.4. Are Slavic compounds always predictive?

Russian, two core types: with a derived second member and with a bare root one (Vinogradov 1952:273-277 for nominal compounds)

Nominal and adjectival compounds all involve the linker o (allomorph e after palatalized consonants)

(17) \text{exocentric [N-\sqrt{v}] compounds (underived verbal root as the second member)}

a.  

\text{sen-o-kos ‘haying’ ← sen-o ‘hay’ + kos-it’ ‘mow’}

b.  

\text{pyl-e-sos ‘vacuum cleaner’ ← pyl’ ‘dust’ + sos-at’ ‘suck’}

c.  

\text{sneg-o-pad ‘snowfall’ ← sneg ‘snow’ + pad-at’ ‘fall’}
The first member is a noun. However, there are exceptions (probably, a closed class):

d. skor-o-xod ‘footman, foot courier’ ← skor ‘fast’ (adverb) + xod-it ‘walk’

And there are A-N compounds of the same type:

e. čern-o-zēm ‘black earth’ ← čern-yj ‘black’ + zem-lja ‘earth’

(18) derived nominal compounds (a derived deverbal noun as the second member)

a. mjas-o-rub-k-a ‘meatgrinder’ ← mjas ‘meat’ + rub-it ‘chop’ + diminutive

b. vod-o-kač-k-a ‘water-tower’ ← vod-a ‘water’ + kač-at ‘pump’ + diminutive

The first member is a noun. However, there are exceptions (probably, a closed class):

c. tix-o-xod-k-a ‘tardigrade’ ← tix-o ‘quiet’ (adverb) + xod-it ‘walk’

And there are A-N compounds with the same suffix:

d. bos-o-nož-k-a ‘sandal’ ← bos-oj ‘barefoot’ (adjective) + nog-a ‘foot’

it seems impossible to unambiguously determine the category of Russian compound members without looking into the details of semantic composition

Yet the first member of a compound created with the linking vowel -o- cannot be a verb

3.2. Category-specific class restrictions

Empirical cases that require access to both PF & LF & the categorial information (the declensional/gender information, in turn presupposed by the category label – see Stevanović 1964 for examples and overview)

(19) Suffix –ad creating collective nouns has very specific restrictions: Neuter, declension class III (exception dete (child) II) ; e-stem; typically [+younglings]: tele (calf) – telad (collective); jagnje (lamb) – jagnjad; pile – pilad (chick); pač (duck) – pačad etc.

(20) Suffix –ov added to derive descriptive labels, specifically for dogs and horses:
zeljov (a dog with gray-green coat – zeleno (green)), žutov (a dog with a yellow coat (žut – yellow), mrkov (a brown horse – mrk (brown)), šarov (a dog with a multi-colored coat – šaren (multicolored))

4. Conclusion

A careful examination of cranberry morphemes reveals that the use of the term is confused

We maintain that parameters such as non-compositional meaning, limited/hapax distribution, derivational status (compounding or affixation), being free or bound and prosodically well- or ill-formed does not entail the lack of the lexical category

Conversely, the under-determination of the lexical category is not limited to morphology

Such distributional properties of roots as conjugation class or gender are difficult to achieve if roots are acategorial (see Matushansky 2015 for discussion)

A closer examination of compounding suggests that the lack of a lexical category in the first member of a compound may be only apparent: while the grammatical category of cob-, cran- or luke- might not seem as obvious as that of ruth- or -ceive, the very fact that they appear in a compound requires them to have a category modulo other language groups (see Zhang 2007 who argues that Chinese exhibits root-compounding). For inflectional languages – even those with morphologically eroded systems like Dutch or English – we suggest that the aphorism, Absence of evidence is not evidence of absence applies.
5. **APPENDIX: THE CASE OF CRANBERRY**

A random search of compounding involving -berry suggests that the category of the first part of this compound is overwhelmingly a noun:

(21) *strawberry* (also *eordberge* *(earth-berry)* and its counterparts in Modern German and Dutch, a.o.): *dewberry; cowberry; mountain berry; foxberry; auailberry; bearberry; beverberry; courgarberry; partridgeberry; lingonberry (lingon - cow), salmonberry; huckleberry (dialect: hurtle-berry - whortleberry); gooseberry; wolfberry (goji berry); bilberry (billo -ball); acai berry; cloudberry; baneberry; buffalo berry; bunch berry; juneberry; elderberry (ME, elder (alder) - a kind of tree); mistletoe berry; nannyyberry; sugarberry; ivy berry; virginia creeper berry; wineberry; whortle berry (whortle – wyrt, Germ. wurzel, root - small shrub); windberry; mooseberry; marion berry; dogberry; poisonberry; brambleberry; crackerberry; mulberry (a tree of the genus *morus* – Modern German: *maulbeere*); raspberry (15th raspite – a kind of wine (sweet))

So, overwhelmingly, the -berry compounds seem to be of the [N-N] type. Even those cases that seem to be cases of A (*youngeberry*) need to be checked carefully. As it turns out, the first member here is a proper name - the founder’s surname. *Pokeberry* derives its name from the name of American Indian Powhatan people, *bovsenberry* owns its name to Rudolf Bovsen and *loganberry* to James Harvey Logan. Finally, *tayberry* is named after the River Tay in Scotland and *farkleberry* owns its name to a dice game called *Farkle*. Even the relatively newly coined *crantini* (analogy to *appletini*) suggests that cran- is treated as a noun. So, there seems to be enough evidence that cran- in ModE is a noun.

Crucially for our discussion, however, our argument would be unaffected if we were to find that cran- were an A, even though most berry-compounds are noun-noun ones. After all, *blackberry* and *blueberry* are likely [A-N] compounds, whereas *goldenberry* is clearly an [A-N] one. The point is that our argument that compounding takes items with a syntactic category as input would still hold. What we wanted to demonstrate with this small corpus exercise and ‘thought experiment’ is that there is far less unpredictability and that even in such an ‘obscure’ set as the set of the original cranberry morphemes it is not the case that ‘anything goes’ universe holds. After all, noun-noun compounds are the most common types of nominal compounds, followed by the adjective-noun and then verb-noun ones (Carstairs-McCarthy 2002, a.o.)

6. **REFERENCES**


