Marijana Marelj² & Ora Matushansky^{1,2} ¹SFL (CNRS/Université Paris-8/SFL)/²UiL OTS/²Utrecht University

ON CRANBERRIES, CATEGORIES, AND CORE Morphology Days in the Low Countries 2021, April 22-23, 2021

1. INTRODUCTION: THE ISSUE OF ACATEGORIAL ROOTS?

Background assumptions: i) 'root' an indispensable part of the mental lexicon; of a traditional and (even more so) a DM one & ii) the notion '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.

QS: What is the nature of roots?

Within DM, there is broad consensus that <u>roots are acategorial</u> (Josefsson 1995, 1997, 1998, 2001, Marantz 1997, 2001, Arad 2003, Embick and Marantz 2006, Embick and Noyer 2007, Acquaviva 2009, Harley 2014, Haugen and Siddiqi 2013, among many others).



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)

Upshot: Roots have a category.

Disclaimer: Though we discuss the nature of roots in general, our primary focus will be on empirical domain of cranberry items. Hence, the conclusions we draw about them are merely suggestive for the broader domain of roots, but due to the nature of the empirical domain probed, they seem quite indicative.

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."

The usual suspects:

- (2) a. *cran* (cranberry)
 - b. *-mit* (permit, transmit, commit, remit, submit, resubmit, recommit, intromit, intermit, remit, *but also* admit, demit, emit, omit)
 - c. *-ceive* (receive, perceive, deceive, conceive, apperceive, misconceive, misperceive, preconceive...)
 - d. *cob* (cobweb)
 - e. *were-* (werewolf, werebear, wereboar, wererat, weretiger)
 - f. -kemp(t) (unkempt)
 - g. *ruth* (ruthless)
 - h. *reck* (reckless, from OE reccelēas)

What these morphemes have in common: they are bound.

2.1. Lack of meaning

Not necessary: *ed*- (edible), *cob*- (cobweb) and *-ept* (inept) are quite transparent and comfortable for further derivation: *in-ept-ly*, *in-ept-ness* etc.

2.2. Hapax morphemes and lexical category

Cranberry morphemes in (2b-c, e) are not restricted to just one environment. Even *cran* isn't (see Appendix for *crantini* etc.)

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

(3) a. ruth-, reck-, gorm-...: -less adjoins only to nouns
b. (il)leg-, plaus-, ed-...: -ible is an allomorph of -able and only adjoins to verbs (cf. *permissible, admissible, accessible...*)
c. -ept: the negative *in*- is an adjectival prefix

No direct connection between the underdetermination of a lexical category and the hapax status.

2.3. Root vs. affix

So far we have been looking at limited-distribution roots. There are limited-distribution and even **hapax affixes**:

- (4) a. -uz-: only used in *franc-uz* 'a Frenchman', cf. *Francija* 'France' Russian 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

(5) a. diev-egge 'female thief', cf. *dief* 'thief'b. frans-oos 'Frenchman', cf. *Frans* 'French'

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

2.4. Free cranberry morphemes

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

(6)	a.	kith and kin	noun
	b.	eke (only in <i>eke out</i>)	verb
	c.	spick and span	adjective

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

And be restricted in their use as well:

(7) a. hark (hark back to or hark at you) verb
b. hither (come hither, hither and thither, hither and yon) demonstrative
c. mettle (in high mettle, test/prove/show (one's) mettle, on (one's) mettle) noun

And, very rarely, their category is underdetermined:

(8) days of yore

There might be some fossil words underdetermined for category. So what? **They must have** a category, this is definitely syntax: *But Satan now is wiser than of yore, and tempts by making rich, not making poor* (Alexander Pope); *A nation now poorer than of yore.*

2.5. 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):

(9) a. *gl*- 'light, vision': glimmer, glisten, glitter, gleam, glow, glint, etc.
b. *sn*- 'nose, mouth': snore, snack, snout, snarl, snort, sniff, sneeze, etc.

For (9a): these are all verbs. So:

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

Why?

- \blacktriangleright 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.

Dutch

2.6. 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

3.1. Compounding vs. derivation

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

- (10) a. [N N]_N: chess table, strawberry jam, diesel motor, bookshelf
 - b. [V N]_N: crybaby, scrubwoman, bakehouse
 - c. [Participle N]_N: filing cabinet, reading class, writing table, drinking water
 - d. [A N]_N: blackbird, drydock, redbrick, wetsuit
 - e. [Particle/Preposition N]_N: outhouse, outgrowth, undergrowth, offprint
 - f. [N A]_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:

(11) a. behind, hindmost, hindsightb. before, foremost, foreleg, foresight...

Probably, nouns. On cranberry itself see the appendix

Suppose we find a better example. What then?

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

Minimally, 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?

3.2. Overt morphology in compounds

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

3.2.1. 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

- (12) a. *vuci-batina* (pull-whip) (*vuci*.IMP, as opposed to *vući*.INF, *vuče*.3SG.Present, *vuk*.ROOT, as evident in *vuk-ao*.PastParticiple)
 - b. *seci-kesa* (cut-purse) (*seci*.IMP, as opposed to *sek*.ROOT (e.g. *trbo-sek* (*stomach-ripper*) *sek-ač* (cut-ter), *seći*. INF *seče*.3SGPresent)
 - c. *deri-koža* (rip-skin) (*deri*.IMP, as opposed to *derati*.INF, *dere*.3SG.Present, *der*.ROOT, as in *oblak-o-der* (skyscrape(er))

(Progovac 2006: Ex7)

d. 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

3.2.2. Infinitive-containing compounds in Dutch

Though the morphemes *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:

(13) *eten-s-tijd* (eating time), *varen-s-man* (sea farer), *zien-s-wijze* (lit. seeing way, view); *uitgaan-s-verbod* (lit. outgoing prohibiton, curfew)

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

3.2.3. Nouns & adjectives as non-heads

If it is plural, it has to be nominal:

- (14) English: mice-infested, menbashing, oxenpower (ox-oxen), dice thrower (die dice)...
- (15) Serbo-Croatian suppletive plurals and case-marked nonheads:
 - a. **dec**-o-ubica (child-murderer): dete (child) deca (children)
 - b. **ljud**-o-žder (cannibal): čovek (man) ljudi (people suppletive)
 - c. **Viš**-e-grad (toponym, lit. *higher-town*): visok (high/tall) viši-(taller)
 - d. **Svrsi**-shodan (useful) svrha nominative svrsi –dative (purpose)
 - e. **Bogu**-mrzak (God-displeasing) bog-nomintive bogu-dative (god)

Particularly interesting are (d) and (e): no linker – nonheads directly adjacent to the head.

3.2.4. 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)

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

- a. sen-o-kos 'haying' \leftarrow sen-o 'hay' + kos-it' 'mow-INF'
- b. pyl-e-sos 'vacuum cleaner' \leftarrow pyl' 'dust' + sos-at' 'suck-INF'
- c. sneg-o-pad 'snowfall' ← sneg 'snow' + pad-at' 'fall-INF'

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

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

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

e. čern-o-zëm 'black earth' ← čern-yj 'black' + zem-lja 'earth'

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

 - b. vod-o-kač-k-a 'water-tower' ← vod-a 'water' + kač-at' 'pump-INF' + diminutive

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

c. tix-o-xod-k-a 'tardigrade' $\leftarrow tix-o$ 'quiet' (adverb) + xod-it' 'walk-INF'

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

3.3. 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)

- (18) 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); pače (duck) pačad etc.
- (19) 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 deceptive.

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 (recall 2.4 and see also: *I know those Algerine cut-throats of yore; and if they are met bravely, they quickly show the white feather.* "Roger Willoughby" by William H. G. Kingston; *Miners Alley visits a Christmas of yore. http://www.boulderweekly.com/article-4022-swansea-song.html*

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. **References**

Acquaviva, Paolo. 2009. Roots and lexicality in Distributed Morphology. In *York-Essex Morphology Meeting 2*, ed. by Alexandra Galani, Daniel Redinger, and Norman Yeo, pp. 1-21. York: University of York.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21, pp. 737-778.

Aronoff, Mark. 1974. Word-structure, Doctoral dissertation, MIT.

- Berent, Iris and Steven Pinker. 2007. The dislike of regular plurals in compounds. *The Mental Lexicon* 2, pp. 129-181.
- Bergen, Benjamin K. 2004. The psychological reality of phonaesthemes. *Language* 80, pp. 290-311.
- Bolinger, Dwight L. 1950. 'Shivaree' and the phonestheme. American Speech 25, pp. 134-135.
- Booij, Geert. 1992. Compounding in Dutch. Rivista di Linguistica 4, pp. 37-59.
- Carstairs-McCarthy, Andrew. 2002. An Introduction to English Morphology: Words and Their Structure. Edinburgh: Edinburgh University Press.
- Daničić, Đuro. 1876. Osnove srpskoga ili hrvatskoga jezika. Beograd: Državna štamparija.
- Embick, David and Alec Marantz. 2006. Architecture and blocking. Ms., University of Pennsylvania and Massachusetts Institute of Technology.
- Embick, David and Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. In *The Oxford Handbook of Linguistic Interfaces*, ed. by Gillian Ramchand and Charles Reiss, pp. 289-324. Oxford: Oxford University Press.
- Firth, John R. 1930. Speech. London: Benn.
- Hammond, Michael. 1993. On the absence of category-changing prefixes in English. *Linguistic Inquiry* 24, pp. 562-567.

Harley, Heidi. 2014. On the identity of Roots. Theoretical Linguistics 40, pp. 225-276.

- Harris, James W. 1996. The syntax and morphology of class marker suppression in Spanish. In Grammatical Theory and Romance Languages: Selected Papers from the 25th Linguistic Symposium on Romance Languages (LSRL XXV), Seattle, 2-4 March 1995, ed. by Karen Zagona, pp. 99-122. Amsterdam: John Benjamins.
- Haugen, Jason and Daniel Siddiqi. 2013. Roots and the derivation. *Linguistic Inquiry* 44, pp. 493-517.
- Josefsson, Gunlög. 1995. The notion of word class and the internal makeup of words. *Working Papers in Scandinavian Syntax* 56, pp. 1-45.
- Josefsson, Gunlög. 1997. On the Principles of Word Formation in Swedish. Lundastudier i nordisk språkvetenskap A 51. Lund: Lund University Press.
- Josefsson, Gunlög. 1998. Minimal Words in a Minimal Syntax. Amsterdam: John Benjamins.
- Josefsson, Gunlög. 2001. The meaning of lexical classes. Nordic Journal of Linguistics 24, pp. 218-231.
- Klajn, Ivan. 2002. Tvorba reči u savremenom srpskom jeziku: prvi deo: slaganje i prefiksacija. Beograd/Novi Sad: Zavod za udžbenike i nastavna sredstva Institut za srpski jezik SANU Matica srpska.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams. *University of Pennsylvania Working Papers in Linguistics* 4.2, pp. 201-225. Philadelphia: University of Pennsylvania, Penn Linguistics Club.
- Marantz, Alec. 2001. Words. Paper presented at WCCFL 20, University of Southern California, Los Angeles
- Matushansky, Ora. 2015. *n* is for "Not there". Paper presented at *FASL 24: Special Workshop* on Approaches to Slavic Morphology, New York University, May 7, 2015.
- Pfau, Roland. 2000. Features and Categories in Language Production, Doctoral dissertation, University of Frankfurt am Main.
- Pfau, Roland. 2009. Grammar as processor. A Distributed Morphology Account of Spontaneous Speech Errors. Linguistik Aktuell 137. Amsterdam: John Benjamins.
- Progovac, Ljiljana. 2006. Fossilized imperative in compounds and other expressions. In Online Proceedings of the Inaugural Meeting of the Slavic Linguistics Society (SLS), Bloomington, IN, September 2006.
- Rhodes, Richard and John Lawler. 1981. Athematic metaphor. In *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society 1981 (CLS 17)*, ed. by Roberta Hendrick, Carrie Masek, and Mary Frances Miller, pp. 318-342. Chicago: Chicago Linguistic Society.
- Stevanović, Mihailo. 1956. Imperativne složenice. Naš jezik VIII, pp. 6-18.
- Stevanović, Mihailo. 1964. Šavremeni srpskohrvatski jezik: Uvod, fonetika, morfologija. Beograd: Naučno delo.
- Vinogradov, V. V. ed. 1952. Grammatika russkogo jazyka. Moscow: Soviet Academy of Sciences.
- Zhang, Niina Ning. 2007. Root merger in Chinese compounds. Studia Linguistica 61, pp. 170-184.

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:

(20) strawberry (also OE eorðberge (earth-berry) and its counterparts in Modern German and Dutch, a.o.); dewberry; cowberry; mountain berry; foxberry; quailberry; bearberry; beverberry; courgarberry; partridgeberry; lingonberry (lingon - cow), salmonberry; huckleberry (dialect: hurtle-berry - whortleberry); gooseberry; wolfberry (goji berry); bilberry (billo -ball – Scandinavian origin/Danish); acai berry; cloudberry; baneberry; buffalo berry; bunch berry; juneberry; elderberry (ME, elder (alder) - a kind of tree); mistletoe berry; nannyberry; 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 raspise – a kind of wine)

So, overwhelmingly, the *-berry* compounds seem to be of the [N-N] type. Even those cases that seem to be cases of A (*youngberry*) one needs to be very careful. 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, *boysenberry* owns its name to Rudolf Boysen 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 than typically assumed 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.)