Slide 1

Root, Shmoot

Mark Aronoff
Stony Brook University
KU Leuven
CRISSP Seminar
13 March 2024

Slide 2

Or

Don't confuse me with the facts!

Slide 3

WARNING

THIS PRESENTATION MAY OFFEND SOME

Slide 4

For, behold, the day cometh, that shall burn as an oven; and all the proud, yea, and all that do wickedly, shall be stubble: and the day that cometh shall burn them up, saith the LORD of hosts, that it shall leave them neither root nor branch.

Malachi 4.1

The Armenian nightingale

- The Armenian nightingale sings only when the moon is full
- And what does it do when the moon isn't full?
- It sings anyway!
- "Roots of course can take on special non-compositional meanings in particular environments" (Marantz 1997)

5

The point of this presentation

- 'Syntactic' theories of morphology employ the notion of meaningful roots that have no lexical category.
  - Call them *acategorial roots*
- There is no empirical need for acategorial roots
- There is no evidence for acategorial roots in natural languages
- Roots do exist but all evidence says that they are morphological and categorial

6

Ockham's razor

- William of Ockham (1285 -1347)
- Ockham's razor:
- *Entia non sunt multiplicanda praeter necessitatem*
- *Numquam ponenda est pluralitas sine necessitate*
- *Frustra fit per plura quod potest fieri per pauciora*
- '[it] becomes futile [to do] with more [things] what can be done with fewer'

7

Less is more   Ludwig Mies van der Rohe

- Ockham's razor has nothing to do with description length, or any other length-based principles of economy
- It rather focuses on the number (*pluralitas, plura*) of (types of) things (*entia*)
- Do not posit unnecessary things or concepts

8

What is to be done

1. Latinate roots in English morphology
2. A brief history of the notion of 'root' in traditional Arabic and Hebrew grammar
3. Acategorial roots
4. Acategorial roots meet William of Ockham
5. What empirical evidence is there for acategoriality?

9

What is not to be done

- This presentation says noting about other aspects of DM or related 'syntactic' approaches to morphology such as Borer's Exoskeletalism
- These approaches do not rise or fall depending on the existence of acategorial roots
- N.B.: The term *root* has now been extended in work on semantics to mean something like 'basic lexical meaning' (e.g., Beavers and Koontz-Garboden 2023). I have nothing to say about such claims

10

Latinate roots in English morphology

- The first serious paper I ever wrote (1972) was titled "Latinate Roots in English Morphology"
- It demonstrated **empirically**
    1. the lack of semantic consistency of individual Latinate roots in English verbs
    2. The highly regular idiosyncratic morphologically conditioned allomorphy of these roots
- These roots thus have non-semantic purely morphological and morphophonological concatenative properties that render them 'linguistically real'

11

Latinate roots in English morphology

- A fixed set of roots
- A fixed set of prefixes
- They occur together as prefix=root combinations only
- Members of neither set occur elsewhere, though there are homophonous prefixes (e.g., *re#*, *pre#*)
- Many roots have conditioned allomorphs before certain suffixes
- One suffix {*ation*, *ion, ition, ution*} has root-conditioned allomorphs

12

Examples of root allomorphy in English

- Receive, reception
- Deceive, deception
- Conceive conception
- Perceive, perception
- Apperceive, apperception

- Deduce, deduction
- Reduce, reduction
- Seduce, seduction
- Induce, induction
- Conduce, conduction
- Produce, production
- Introduce, introduction
- Reproduce, reproduction

13

An aside
Roots or stems? Make up your mind!

- In Chapter 2 of *Word Formation in Generative Grammar* (1976) I called these items *bound stems*, a term taken from *The Sound Pattern of English* (Chomsky and Halle 1968)
- The SPE analysis of Latinate verbs of the form prefix + "stem" (termed 'complex verbs') recognizes that "these stems and prefixes are not, in general, independent words or even separate lexical items." (SPE p. 94)
- In Chapter 5, which provides a detailed analysis of these verbs, I used the term *root* rather than stem, without comment
- I did not notice this discrepancy until the other week, a half-century later!

14

In the beginning   בְּרֵאשִׁית

- The notion of the 'root of a word' was first discussed by early Classical Arabic grammarians
- The first important figure of the Arabic grammatical tradition was Sibawayhi (8th Century)
- Sibawayhi's term for 'root' was ʔaṣl
- ʔaṣl translates variously as 'root', 'trunk' or 'base' (of a tree)

15

The roots in early Arabic grammar are words

- In early Arabic morphological analysis, derivation does not operate directly from the consonantal root, but *de mot à mot* 'from word to word' (Bohas 1984).
- One word-form is said to be the ʔaṣl of another, as follows:
  - The maṣdar is the initial form of the verb
  - The maṣdar is the ʔaṣl of the past form of the verb
  - The past form of the verb is the ʔaṣl of the present
  - The present form is the ʔaṣl of the imperative

16

Which form is the root?

- Word--based derivation makes sense if we assume the influence of the Greek and Syriac grammatical traditions, which were strictly word--and--paradigm and admitted of no morphemes
- The early schools of Basra and Kufa differed on which word-form was the root
- The Kufa school believed that the third-person perfect of the verb was the root of the verb
  - This tradition is reflected in the use of this form to name the binyanim in both the Arabic and Hebrew traditions
- The Basra school believed that the *Masdar* nominal/gerund was the root of the verb

17

More roots

- The term used more commonly in Arabic grammar nowadays is ɟaðr, which is also the term for (square) root in mathematics, another concept from the Arabic scholarly tradition
- It is not clear whether the grammatical use or the mathematical use came first
- The term used in the Hebrew-language grammar tradition is ʃɔrɛʃ 'root'
- Like most Hebrew-language technical terms, this is a direct translation of the Arabic term

18

Hebrew roots

- The first Hebrew grammarian Judah Hayyuj, who lived in Cordoba in the 10th century, adopted Sibawayhi's term in his treatises on Hebrew verbs (written in Arabic)
- Hayyuj's analysis of 'weak' roots with evanescent consonants pioneered abstract morphophonological and phonological analysis. It stands to this day
- Hayyuj's analysis led to the positing of abstract tri-consonantal verbal roots with three (even more) abstract position classes, first (p), middle (ʕ), and last (l), in later Hebrew grammar
- These position classes are defined for the verb (*hapoʕal*), not the noun

19

Medieval Hebrew Grammar written in Hebrew

- True root-based grammatical description first flourishes in early works written in Hebrew
- The term used in the Hebrew-language grammatical tradition is ʃɔrɛʃ 'root'
- Like most Hebrew-language technical terms, this is a direct translation of the Arabic term
- For scholars of Hebrew, the ʃɔrɛʃ is tri-consonantal and abstract (in the original sense), not an actually occurring form. Compare the use of ʔaṣl in the Arabic tradition, which does denote a form
- The most famous of these grammars was written by David Kimħi, who lived in Provence (1160-1235) (Chomsky 1952)
- Kimħi wrote a dictionary entitled *Sefer Hashorashim* 'the book of roots'

20

Roots and categories in Hebrew Grammar

- In traditional descriptions of Hebrew morphology, all roots have categories
- Grammatical description is devoted almost entirely to verbal morphology
- The tri-consonantal root system is based on the word for 'verb': *poʕal*
- The word for 'conjugation', *binyan*, is reserved only for verbs
- Verbs must occur in a binyan
- The word for noun pattern is *miʃqal* 'weight'
- Nouns do not all fall into patterns

21

From ʃɔrɛʃ to root

- ʃɔrɛʃ is the most common term for 'root' in the original Hebrew text of the Bible
- It also means 'heel' or 'sole' of the foot
- The term occurs close to 90 times in the text of the Old Testament
- Western scholars were more familiar with Hebrew than with Arabic, because of the religious importance of the Old Testament in Christianity
- They translated the Hebrew grammatical term ʃɔrɛʃ into Latin (*radix*), English(*root*), French (*radical*), etc.
- The Hebrew tradition of abstract grammatical analysis influenced Renaissance grammarians and led to modern morpheme-based linguistics

22

OEDO definition of root in linguistics (Technical III.16), with early citations

- An ultimate unanalysable element of language; a morpheme, not necessarily surviving as a word in itself, as a base from which words are formed by means of affixation or other modification. Also: a word from which another or others derive; an etymon.

1530 His thre chefe rotes, that is to say, his theme, his preterit participle, and his present infynityve. *J. Palsgrave, Lesclarcissement Introduction 31*

1599 Recourse must be had to the Hebrew, euen to a false roote

23

Root vs. stem (Aronoff 1994)

Although root and stem both designate sound forms of lexemes, the most important difference between them is that a root is defined with respect to a lexeme, while a stem is always defined with respect to a realization rule. . . . A root is what is left when all morphological structure has been wrung out of a form. (p. 40)

24

Definition of stem (Aronoff 1994)

- The stem of any realization rule is the form of the lexeme on/from which the rule/function performs its realization, i.e. the phonological domain of the rule. While a lexeme has arbitrary properties on three dimensions – sound form, syntax, and meaning – as well as the usually arbitrary association among them, I will reserve the term stem for only the sound-form part of this trinity. A stem, in my use of this term, is a sound form. In particular, it is the phonological domain of a realization rule: that sound form to which a given affix is attached or upon which a given non--affixal realization rule operates. (p. 39)

25

Definition of root (Aronoff 1994)

- Defined only informally: A root is what is left when all morphological structure has been wrung out of a form
- This informality is due to the fact that a root plays no large role in the types of phenomena that we are interested in in that book
- Any lexeme-based framework is interested in the functions that form words
- The domain of the functions does not receive much attention
- It could, but it does not

26

Alloroots

- Roots are pure form and have no meaning
- They are morphological in the original sense
- Roots may have morphologically conditioned *alloroots*
- Each alloroot appears in a morphologically conditioned environment
- In the best cases, there are no exceptions to the morphological distribution

27

English roots

- The root {stand, stood}
- The past tense of any verb whose unmarked root is *stand* is *stood*
- *Stand* (in all its senses!!), *withstand, understand*
- Entries with the most separate senses listed in OED: *run* (645 senses), *set, put, take, go, be*
- **NO EXCEPTIONS FOR ANY OF THESE !!!!**
- denominal verbs lose the root (Kiparsky 1982): **grandstood*

28

The lexicalist hypothesis

- Noam Chomsky. 1970. Remarks on nominalization. In Jacobs and Rosenbaum, eds. Readings in English Transformational Grammar. Ginn and Company: Waltham, MA. 184 – 221.
- We might extend the base rules to accommodate the derived nominal directly (I will refer to this as the "lexicalist position . . .(p. 188)

29

The origins of acategoriality

Let us propose then, as a tentative hypothesis, that a great many items appear in the lexicon with fixed selectional and strict subcategorization features, but with a choice as to the features associated with the lexical categories noun, verb, adjective. The lexical entry may specify that semantic features are in part dependent on one or another of these categorial features. This is of course the typical situation within the lexicon; in general, lexical entries involve certain Boolean conditions on features, expressing conditional dependencies of various sorts (cf. Note 11). Insofar as there are regularities, these can be expressed by redundancy rules in the lexicon.[12] (p. 190)

30

Not so fast, my friends!

From footnote 11: The scope of the existing subregularities, I believe, has been considerably exaggerated in work that takes the transformationalist position. . . In general, there are few cases where problems of this sort do not arise. Correspondingly, the transformationalist position is impossible to support, and difficult even to maintain, on semantic grounds.

31

Not so fast, my friend!

Footnote 12: It is immaterial for present purposes whether a lexical entry is regarded as a Boolean function of specified features or is to be replaced by a set of lexical entries, each of which consists of a set of specified features. It is unclear whether these approaches to problems of range of meaning and range of function are terminological variants, or are empirically distinguishable. . .

32

Acategorial roots

Chomsky's "tentative hypothesis, that a great many items appear in the lexicon with fixed selectional and strict subcategorization features, but with a choice as to the features associated with the lexical categories noun, verb, adjective" led to the suggestion by proponents of Distributed Morphology and Exoskeletalism, that such putative items be called 'roots'.

The suggestion was rendered visually by using the root symbol √

These roots bear little resemblance to either Semitic roots or the Latinate roots of Aronoff (1976) et seq.

33

Antoine Fabre D'Olivet

La Langue hébraïque restituée et le véritable sens des mots hébreux rétabli et prouvé par leur analyse radicale, ouvrage dans lequel on trouve réunis : (1) une dissertation sur l'origine de la parole; (2) une grammaire hébraïque; (3) une série de racines hébraïques; (4) un discours préliminaire; (5) une traduction en français des dix premiers chapitres du Sépher, contenant la Cosmogonie de Moyse (1815).

34

Imposter syndrome

- DM has a long history of taking established terminology and 'adapting' it to its own purposes
- Perhaps the most egregious case is the use of the term *morpheme* for the well-established notion of 'morphosyntactic feature'
- Similarly, the use of the term *root* for an acategorial item bears little relation to the traditional notion of a Hebrew root, which is always treated as categorial within the morphological analysis of the language and is a unit of form

35

Two ways of looking at simplex pairs in Root-based theories

√DOG

dog_N dog_V

N > V
Ex: dog_N > dog_V

V > N
Ex: $bite_V$ > $bite_N$
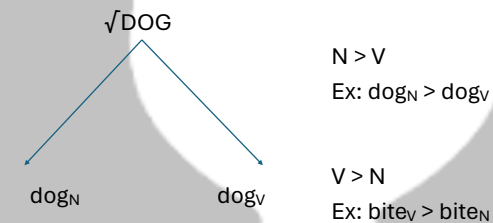
36

DM acategorial roots

The *loci classici* of the use of roots in DM are Pesetsky (1995) and Marantz (1997):

"Roots like √DESTROY and √GROW (to borrow notation from Pesetsky 1995) are category neutral, neutral between N and V."

Pesetsky 1997

"As this typology illustrates, all approaches assume that roots are category-neutral. That is, roots are not stored in the mental lexicon with a category; rather, roots acquire a given category by being merged into a syntactic derivation." Lohndal 2020, ORE

"Furthermore, there is also consensus in roots having some kind of conceptual content. For example, roots like √BOOK and √SING represent concepts, and they become either a noun or a verb depending on which syntactic environments they occur in." (ibid.)

"Roots of course can take on special non-compositional meanings in particular environments" (ibid.)

37

DM roots and meaning (Marantz 1997)

- One of the basic tenets of DM and other 'syntactic' theories is that roots have meaning
- "Things with special meanings are roots."
- "Everyone agrees that there has to be a list of basic sound/meaning connections for the atomic building blocks of language (= the 'morphemes'). There also has to be a list of idiosyncratic properties associated with the building blocks."
- "This List 1 contains the atomic roots of the language"
- "The Encyclopedia lists the special meanings of particular roots, relative to the syntactic context of the roots, within the local domains)"
- Oy vey iz mir!

38

Think before you speak!

- "Nouns like "transmission," "ignition," and "administration" carry the semantic implication of their internal structure, which includes an aspectual pre-verb, a verbal stem, and a nominalizing suffix. (Marantz 1997)
- If these words refer to things, then these things should be for accomplishing something – **and this is indeed the case**." (ibid.)
- *N. B.: ignition* has no "aspectual pre-verb"
- The verb *ignite* is back-formed in English from *ignition*
- What is an "aspectual pre-verb" anyway?

39

NO THEY DON'T!
Administration

- Definitions from Merriam-Webster online (https://www.merriam-webster.com/dictionary/administration)

- 4.c: a governmental agency or board
  - The U.S. Food and Drug *Administration*
- 5: the term of office of an administrative office or body
  - During the first Cinton *administration*

40

NO THEY DON'T!
Ignition

- One frequent sense of *ignition* is the slot where one puts the key to start a vehicle.
- This sense bears no connection to the verb *ignite*
- The phrase "key in the ignition" produces over 3,000,000 hits in a Google search
- In the last couple of decades, this sense has morphed into (car) 'start button' (1.4B Google hits)
- The term "Tesla ignition" yields about 65,000 hits in a Google search, although electric vehicles have no ignition system
- It refers to the start system, which does not ignite anything
- No one ever wants an electric vehicle to ignite!

41

Have your cake and eat it too!

- "Roots of course can take on special non-compositional meanings in particular environments"   Marantz 1997
- "Thus the noun dog is at least bimorphemic, with the structure: [[√DOG] n], where n is a nominalizing morpheme that establishes the category of the word." Bobaljk 2017. ORE
- "At the limit, as mentioned above, even the simplest of words (from the lexical classes) have an internal syntactic structure: dog is structurally [[√DOG]n], where n is a syntactic terminal that established the category feature: Noun." Bobaljk 2017. ORE

42

Ockham's razor redux

- *Entia non sunt multiplicanda praeter necessitatem*
- *Numquam ponenda est pluralitas sine necessitate*
- *Frustra fit per plura quod potest fieri per pauciora*
- Ockham's razor focuses on the number (*pluralitas, plura*) of things (*entia*), not on description length

- Do we ever need acategorial roots or are they *Entia multiplicanda praeter necessitatem?*

43

Arad on Modern Hebrew

- The best-known example of using both acategorial root-based and lexeme-based derivation is Arad (2005)
- "We have shown that the empirical claims in Arad (2003, 2005) are not convincingly supported by the data. In particular, the claimed correlation between idiosyncratic and unpredictable semantics and phonology on the one hand, and root-derived ("de-root") status on the other, does not seem to be borne out, both because of confounds in examples that were given as positive support for the correlation, and because of counterexamples that directly contradict it. We therefore conclude that these arguments, at least, fail to support the existence of category-neutral roots in the syntax."
  Ezer Rasin, Omer Preminger, and David Pesetsky (2021). A re-evaluation of Arad's argument for roots. WCCFL 39

44

The Armenian nightingale

- The Armenian nightingale sings only when the moon is full
- And what does it do when the moon isn't full?
- It sings anyway!
- "Roots of course can take on special non-compositional meanings in particular environments" (Marantz 1997)

45

Armenian ignition

- One might claim within a DM analysis that the two senses of 'ignition', one related to the verb and one not related to the verb are derived in two different ways, á la Arad.
- The related sense is derived from the verb in the standard way: V → N
- The pair in which the verb and noun are unrelated both derive from the acategorial root √IGNITE
- In that case they share no related semantics
- We can do this whenever two etymologically related words diverge semantically
- The strategy tells us nothing

46

Finally, some data and analysis

- English has a large number of simplex nouns and verbs
- Many of them come in pairs
- Traditionally, one member of each pair is considered basic, the other derived
- N > V or V > N
- If the DM claim about acategorial roots is correct, there should be no strong evidence for directionality in these pairs
- There should be a good number of cases of non-directional but clearly related [N, V] homophones

47

Anecdotal evidence (Aronoff 1980)

- Assuming N > V directionality, Clark and Clark (When nouns surface as verbs) provided a catalog of semantic types of zero-derived denominal verbs
- Aronoff (1980) unifies them all under one simple rule N > V, where **V has something to do with N**
- The rest is **ALL** pragmatics and history or chance
- Examples: *Houdini, lynch, boycott, bork, Bobbitt, bogart, Windex*
- The verb must be derived from the noun with no fixed semantics
- Nothing else makes sense
- Compare Kripke's (1980) causal theory of reference for proper names

48

"The meaning of zero nouns and zero verbs" Mitetelu et al. 2023

"We work with a dataset of 4,879 N-V word sense pairs extracted from the Princeton WordNet standoff file containing pairs of morphosemantically related noun and verb senses (Fellbaum 1998; Fellbaum et al. 2009). For these pairs, we use the derivational direction provided by the Oxford English Dictionary."

"Do the individual morphosemantic relations particularly associate with one derivational direction or the other? That is, do we find relations that primarily involve zero nouns or zero verbs?"

"zero nouns are primarily formed to denote events: 70.3% of them appear in this relation"

"[For zero verbs] The meaning of the base only loosely contributes to some pragmatically and contextually determined manner specification on the event of the output zero verb."

49

"The meaning of zero nouns and zero verbs" Mitetelu et al. 2023

In §4.1, we saw that the Event relation is the most frequent one in the dataset and substantially more frequent with zero nouns (63.5%) than with zero verbs (36.5%), if we keep in mind that the overall dataset is biased towards zero verbs (60%). The relation Event between nouns and verbs is expected to be frequent, given the primarily eventive meaning of verbs (Koontz-Garboden 2005; Baker & Croft 2017). Its higher frequency with zero nouns, however, also confirms the intuition that, to create an Event relation, the input category must be that of the verb, whose primary ontological meaning is eventive. The noun is the output category, as events are not typical denotations in the ontology of nouns (pace some exceptions like trip, movie).

50

THANK YOU

Special thanks to

Robert Hoberman

Jose Martinez Delgado

51