On the Identity of Roots
(Harley, 2013)

John Gluckman

October 28, 2014

> "**[S]yntactic roots are individuated as pure units of structural computation, lacking (in the syntax) both semantic content and phonological features**"

- There are three parts to this paper.
  - I. Roots are underspecified for semantic and phonological information.
  - II. Roots are **not** "syntactically deficient"; they take complements and specifiers.
  - III. A discussion of the "domain of interpretation" of a root – that is, what can trigger allosemy.

1 DM and problems with root insertion

- Harley assumes as standard theory of Distributed Morphology, as represented in Figure 1.

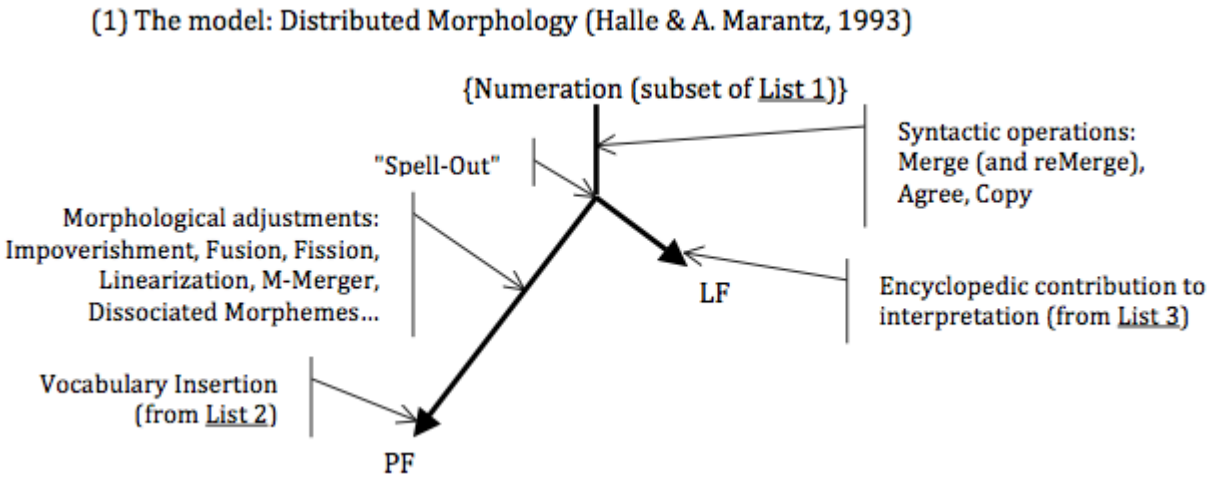


Figure 1: Architecture of the Distributed Morphology model

- In the original model (and subsequent incarnations), roots are underspecified for phonological information, only getting a phonological form at the point of Vocabulary Insertion.
  - The impetus was that syntactic processes could affect phonological form, thus, the phonology had to wait until syntax was finished, before it could insert the correct form.
- Moreover, in the original model, roots were "differentiated late", meaning that there was nothing that differentiated $\sqrt{CAT}$ and $\sqrt{DOG}$ during the syntax.
  - In other words, the syntax only sees that there IS a root, but nothing inherent ABOUT the root.[1]
- The unfortunate consequence was that both PF and LF needed to be read "simultaneously" at the interpretative interface to make sure that [dɔg] was inserted when the speaker meant *dog*, and not, say, *cat*.
  - This lead to FREE-CHOICE insertion of roots, since they were not subject to competition, like other morphemes. Rather, they were inserted "based on the entire morphosyntactic derivation to that point, and [the speaker's] communicative intent."

[1] However, features like [+count], [+animate], etc, which are syntactically active, can be available on the root for syntactic manipulation.

  - This contrasts with non-root Vocabulary Insertion, which has COMPETITION-based spell-out: the familiar "Lists" of DM, where more specified items are inserted before less specified ones, and the Subset Principle and Elsewhere Condition are in effect.
    a. +pl ↔ [ən] / $\sqrt{OX}$
    b. +pl ↔ [z]

- Note that Free-Choice and Competition for roots are mutually exclusive.
  - A root which has a contextual allomorph would be ranked first among all roots, as it is the most specified and roots are not differentiated.
- Thus, if there are cases of true root suppletion, then the Free-Choice Model is unsustainable.[2]
- Unfortunately for the original model, root suppletion is actually quite common in many languages, generally being conditioned by number and tense/aspect within the verbal category.[3]
- An example from Hiaki.

(1) a. *Aapo aman **vuite**-k*
3sg there **run.sg**-PERF
'He ran over there'

b. *Vempo aman **tenne**-k*
3pl there **run.pl**-PERF
'They ran over there'

- Given the suppletion, it is tempting to create spell-out rules which insert different phonological forms of an abstract semantic concept **RUN**, conditioned by a plural feature.
  a. $\sqrt{RUN}$ ↔ /tenne/ / [DP$_{pl}$ [$_{vP}$_____]]
  b. $\sqrt{RUN}$ ↔ /vuite/
- So, it seems clear that roots must be differentiated before spell-out, as they are sensitive to syntactic contexts. Semantic differentiation of roots works for most languages, however it fails completely in others.

[2] And cases of suppletion in English (e.g., *go* → *went*) were explained as allomorphy of a functional category (e.g., *v*), not the root. It was assumed that this suppletion was available for "light" categories.

[3] Harley notes that an survey of suppletion finds many instances where the suppletive verb is not "light" – meaning that it has a highly specified meaning. She proposes that the crucial ingredient is frequency of the verb form.

  - This is famously true of Hebrew roots, which vary drastically in meaning, depending on syntactic context.

    root: $\sqrt{\textbf{kbʃ}}$
    kviʃ, 'paved road, highway'
    mixbaʃa, 'pickling shop'
    hixbiʃ, 'subdue, subjugate'
    kuvuʃim, 'conserves, preserves'

- An analogous situation arises in English with what Harley calls "caboodle" items (also sometimes called "cran-" morphs)
  - *-ceive*
    deceive, receive, conceive, perceive
  - *-pose*
    suppose, depose, compose, repose, propose
- These are clearly identifiable roots, acting similarly with respect to phonology, prosody, and morphological selection, but each root lacks an identifiable semantics, the meaning being dependent on what other affix it combines with.
- Harley argues that the same effects occurs in some idioms (hence the term "caboodle" items)
  - kit and **caboodle**, 'everything'
  - run the **gamut**, 'include a whole range'
  - high **jinks**, 'mischief'
- The point is, "jinks" has no meaning outside of the phrase *high jinks* – that is, its meaning is entirely context dependent.
- Note that "jinks" still is syntactically active, triggering agreement, for example, *These high jinks are killing me*.

> "[Roots] must be individuated, but no single type of independent interface property can be taken to individuate them. They are simply units of morphosyntactic computation – abstract morphemes in the truest sense. We cannot individuate them by their phonological properties, which may depend on the derived morphosyntactic context; neither can we individuate them by their interpretive properties, for the same reason." (p. 14)

- **List 1**
  FORMATIVE LIST: the list of differentiated roots that can be selected by syntax
- **List 2**
  EXPONENT LIST: the list of mappings from root to phonological form
- **List 3**
  ENCYCLOPEDIC LIST: the list of mappings from root to semantic "form"

2 Root identification

- Since roots must be identified, but cannot be identified by semantic or phonological means, Harley assumes that they are indexed.[4]

	PF instructions	LF instructions
a.	$\sqrt{}_{278} \leftrightarrow$ [teɪp]	a. $\sqrt{}_{278} \leftrightarrow$ TAPE

- It's possible that the PF component has a dependent allomorph

	PF instructions	LF instructions
a.	$\sqrt{}_{34} \leftrightarrow$ [wɛn] / [T_{past} [v [_____]]]	
b.	elsewhere [gou]	a. $\sqrt{}_{34} \leftrightarrow$ GO

- Alternately, the phonology might be constant, but the meaning might be contextually determined.

	PF instructions	LF instructions
a.	$\sqrt{}_{67} \leftrightarrow$ [θrou]	a. $\sqrt{}_{67} \leftrightarrow$ VOMIT / [v[[_____][up]]]
		b. elsewhere THROW

- Harley cites *-ceive/-cept* as an example of the last logical possibility, where both PF and LF trigger allomorphy/-semy.
  - I think there may be better examples, e.g., *went crazy*, seems to me to involve both PF and LF changes of the root $\sqrt{GO}$.

[4] Harley assumes that "the various $\sqrt{}$ items may have interpretations as predicates of entities... predicates of properties... or predicates of events.

The four logical possibilities:

	no sem suppletion	sem suppletion
no phon suppletion	tape	throw (in throw up)
phon suppletion	went	-ceive/-cept

- What is crucial for Harley, is that, unlike on the PF side, the LF lists do not always contain an Elsewhere Condition. This must be true for the caboodle items, since they are only contextually defined.
  - To Harley, PF and LF are fundamentally different in this regard, and she explains the difference in the following manner,

> "Model-theoretic interpretations must compose with the interpretations of other elements in their syntactic environment using one of a limited number of composition operations, most commonly function application. . . Even the 'literal' meaning of a root is only well-formed if its type-theoretic restrictions are satisfied by the entities with which it is merged." (p. 17)

  - I think the crucial word here is "entities". What Harley is saying is that ALL (root-)concepts are ill-formed without being "grounded" by something. But what are those "somethings"?
  - Are they entities of type <et>? Are they events? Categorizers? Something else?
- In any event, getting rid of the Elsewhere Condition in the Encyclopedic List is empirically necessary for caboodle-terms, but it is unclear what it means in theory.

Aside #1: Arregi and Nevins (2014)

- A&N give an enthusiastic thumbs up to Harley's approach, and, among some other points, they suggest that PF has cases which lack an Elsewhere form as well.
- Specifically, they discuss paradigm gaps in Spanish.
- Famously, some verbs lack a phonological form only in certain phonological conditions.
  - For instance, *abolir*, 'to abolish', only has a phonological form if the root is directly followed by a high front vowel.
- A&N suggest that this can be accomplished by the following Vocabulary Item.[a]

  (2) [abol]/ $\leftrightarrow \sqrt{}_{385}$ / _____ [-cons, +hi, -bk]
- Crucially, if the conditions in (2) are not met, the form cannot be spelled-out – although of course all the cells in the paradigm are perfectly well-formed concepts.

[a] I've slightly altered how this rule is stated for consistency; it does not affect their arguments.

3 Roots in the syntax

- Given the sparsity information inherent to roots, they have typically been assumed to lack any syntactic information.
  - They cannot take complements, specifiers, or impose selectional restrictions on the structure
  - They must combine with a categorizer first in all cases, in order to project.
- Harley argues that there are indeed arguments in favor of "syntactically active" roots.
- In particular, she argues that roots can, and do, take complements.[5]

[5] I think it's fair to say that Harley does not take any of the following points as a knock-down argument

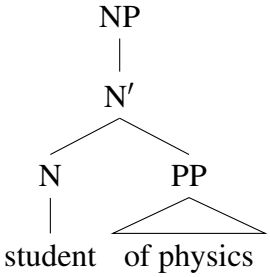
1. *One*-substitution

(3) a. *This [student]$_N$ [of chemistry]$_{PP}$ and that [one]$_N$ [of physics]$_{PP}$ sit together

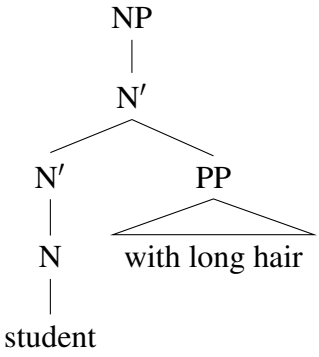
b. This [student]$_N$ [with short hair]$_{PP}$ and that [one]$_N$ [with long hair]$_{PP}$ sit together

- *One*-replacement is normally taught to undergrads as N′ substitution. This cannot be maintained in Bare Phrase Structure, as there is no intermediate non-branching layer equivalent to N′.

(4) a.



b.

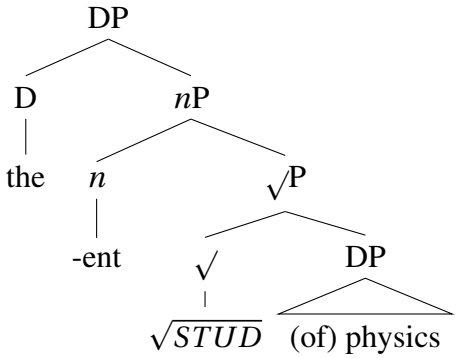


- Under Harley's proposal, *one* substitutes for *n*P, which is the locus of PP adjunction.[6]

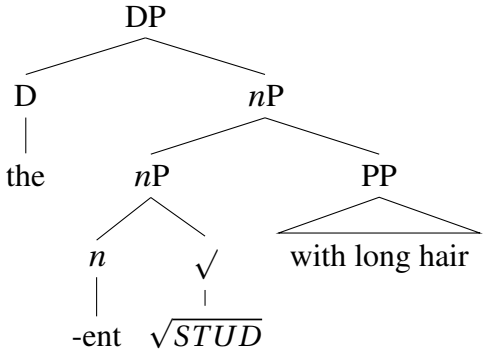
in her favor. Rather she claims that these are consistent, and suggestive, of the approach she is laying out. She definitely agrees that more work will need to be done to confirm her hypothesis.

[6]Here and for the rest of the paper, I'll ignore the index notation for roots, although we can assume that they are all present. This is purely a choice for ease of explication.

(5) a.



b.



2. Verb-object idioms

   - Idiom interpretations are often dependent on the internal argument.

     kill a bug, 'cause the bug to croak'
     kill a conversation, 'cause the conversation to end'
     kill an evening, 'while away the time span of the evening'
     kill an audience, 'entertain the audience to an extreme degree'
     etc

   - Harley takes the idiom meanings to be evidence of the locality of the internal argument to the root.
   - (I'll note though that this is an extremely weak argument, particularly with respect to her later discussion on the conditioning environment for allosemy. To be fair, though, I think Harley is perfectly aware of the weaknesses here.)

3. Hiaki root suppletion

   - Hiaki verbs are conditioned for number (and may supplete) depending on the number of the **intransitive subject**, and the **transitive object**.

(6) Intransitives

a. *Aapo weye*
3sg walk.sg
'He/she/it in walking'

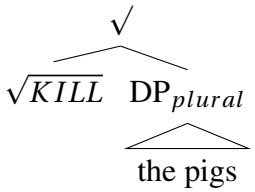
b. *Vempo kaate*
3pl walk.pl
'They are walking'

(7) Transitives

a. *Aapo/Vempo uka koowi-ta mea-k*
3sg/3pl the.sg pig-ACC.sg kill.sg-PERF
'He/They killed a pig'

b. *Aapo/Vempo ume kowi-m sua-k*
3sg/3pl the.pl pig-pl kill.pl-PERF
'He/they killed the pigs'

- If this is true agreement, this contradicts two fairly robust (though not universal) cross-linguistic patterns.
  I. Agreement generally tracks the morphologically unmarked case
  II. Languages which have a nominative-accusative case alignment, do not have an ergative-absolutive agreement alignment.
- Harley proposes that the suppletion here is not agreement *per se*, rather, it is merely allomorphy conditioned by a plural argument local to the root.



- In support, she notes that among the intransitives, only unaccusatives supplete.
- The arguments presuppose that suppletion – or any morphologically triggered phonological change – is subject to strict locality conditions. An intervening categorizer (e.g., *v*) would exclude anything outside of the categorizer from the condition environment.

4 The domain of interpretation

- Finally, given the strict locality conditions on allomorphy, Harley asks the question what the locality constraints on allosemy are.
  - Specifically, are there structural conditions/restrictions on which elements can condition allosemy?
- Here, she tentatively concludes that there is no precise position that can be cross-linguistically identified as the locus beyond which allosemy is precluded, although she hypothesize that the domain is defined by the introduction of an "Agent" or perhaps any specifier.
- Working through a lot of literature on verbal syntax, she eventually adopts a view of syntax which includes a categorizer (v), distinct from the projection that introduces an external argument (Voice).
- She then addresses whether v defines the locality domain for interpretation, as has been argued for allomorphy. Or more broadly, does the **first categorizer** define the domain?
- Harley concludes that this just cannot be the case.

> "Obviously the interpretation assigned at the level of the first categorizing affix will be idiosyncratic, as the root never occurs without such superstructure, and cannot be interpreted in its absence. However, it seems clear that idiosyncratic semantics can also be assigned outside the first categorizer heads, on later cycles of derivation."

a.	edit	edit-or	editor-ial	
			'of or relating to the editor'	*compositional*
			'opinion article'	*idiosyncratic*
b.	nature	natur-al	natural-ized	
			'made natural'	*compositional*
			'became a citizen'	*idiosyncratic*
c.	class	class-ify	classifieds	
			'#things which have been classified'	*#compositional*
			'small newspaper advertisements'	*idiosyncratic*

- Harley hypothesizes (following Marantz), that Voice defines the domain of interpretation – it is the "first real phase-head". Anything before Voice can possibly be a trigger for allosemy. (And everything after cannot.)

Aside #2: Anagnostopoulou (to appear)

- Anagnostopoulou offers some arguments against Voice as defining the domain of interpretation, and in favor of the first categorizer.
- Adjectival participles use v as the domain for interpretation
  - Stative (*open*) versus Resultative (*opened*) adjectives differ according to whether there is an event or not.
  - Greek morphology has been argued reflect this difference everywhere: *-menos* attaches above (eventive) v, while *-tos* attaches directly to the Root.
  - Anagnostopoulou further argues that while the *-tos* forms can be idiosyncratic, the *-menos* forms are purely compositional, i.e., given the root $\sqrt{SPAZ}$, BREAK,

    a. *spas-tos*
    "folding" (umbrella)

    b. *spas-menos*
    "broken"

  - However, she concedes that Voice is a <u>further</u> domain for which allosemy can be determined, except that in this allosemy "the root has exactly the same literal meaning as the root of the corresponding verb, while the root in the *-tos* participles receives a distinct interpretation".

    a. *spas-ta Ellinika*
    'broken Greek' (not very good Greek)

    b. *spas-meni foni*
    'broken voice' (voice that does not sound clear)

5 Conclusion and discussion

- I take the proposals laid out in this paper as a research agenda, to be investigated and fleshed out.

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

1. What is it about Voice – or the projection which instantiates the domain – that defines it as such?
2. What is the "conspiratorial nature" of idioms? That is, in *kick the bucket*, both *kick* and *bucket* fail to get their "normal" meanings.
    - And why doesn't, say, *the* receive a different meaning?
3. What is the psycholinguistic reality of this model? Some work is now being done that seems to suggest that this type of root differentiation does indeed have mental correlate, but findings are controversial and sometimes inconsistent.

References

Anagnostopoulou, E. (to appear). Roots and Domains: Some remarks "On the identity of Roots". *Theoretical Linguistics: Special issue on Roots in Distributed Morphology*.

Arregi, K. and Nevins, A. (2014). A monoradical approach to some cases of disuppletion. http://ling.auf.net/lingbuzz/002049/current.pdf.

Harley, H. (2013). On the Identity of Roots. http://ling.auf.net/lingbuzz/001527.