Natural language without semiosis

Omer Preminger
omer@lingsite.org
Introduction

At the center of much of the discussion of language in the first half of the 20th century —

the SIGN

ENG: /ˈtej.bl/

SPA: /ˈme.sa/

HEB: /ʃul.ˈxan/
Much was made about the often **arbitrary** nature of the relation between the meaning of a SIGN and its form (Saussure 1916, Hjelmslev 1943)

*as contrasted with: “onomatopoeia”, “iconicity”, etc.*
Of course, anyone who has thought about this carefully has noted that this is wrong — at least for human language.

*Cf. Seyfarth & Cheney (1980), and subsequent work, on vervet monkey alarm calls.*
considerably more abstract!

\[ \lambda x. \text{blah}(x) \]

\[ /\text{bla}:/ \]
(1) This book is old and crumbling, but will captivate you like no other.

(2) This window is double-glazed and has a magnificent view.

cf.: (3) #This bug can record 3 hours of conversation and its bite will cause a rash.
considerably more abstract!

also very abstract!

CENTRAL QUESTION:
Can adequately abstract notions of "FORM" and of "MEANING" salvage a semiotic view of linguistic atoms as <FORM, (SYNTAX,) MEANING> mappings?

CLAIM:
The answer is no.
WHAT I WILL ARGUE FOR:

- syntactic terminals don't "have forms" and they don't "have meanings"

- they are, instead, fully abstract

- they come to be associated with FORM via many-to-one rules from syntactic terminals to exponents

- they come to be associated with MEANING via many-to-one rules from syntactic terminals to listed meanings

(see also Pesetsky 1985)
NB: contiguity
In other words, the proposed architecture of *listed* (a.k.a. "lexical") knowledge:

(A) fully abstract syntactic atoms (e.g. DOG, PAST, RUN, IN, etc.)

(B) many-to-one rules from sets of nodes in (A) to units of FORM

(C) many-to-one rules from sets of nodes in (A) to units of MEANING

NB: contiguity
IF THIS PROVES TO BE CORRECT, THEN:

“What does the word/morpheme $w$ mean?”

“How do speakers (of this language) pronounce the meaning $m$?”

NOT, STRICTLY SPEAKING, COHERENT QUESTIONS!
E.g.: suppose, hypothetically, that we found that sentences with "again" in them are interpreted in two different ways, depending on their syntactic structure.

Conventional responses:

(1) Posit two, homophonous "again"s (cf. bug-bug), each with a restricted syntactic distribution.

(2) Try to find a single, "flexible" semantics for "again", which will give the right reading in each syntactic environment.

 Typically treated as "uninteresting"...
**E.g.:** suppose, hypothetically, that we found that sentences with "again" in them are interpreted in two different ways, depending on their syntactic structure.

But these responses both assume that there is a bona fide linguistic object "again", which is submitted to interpretation.

If the arguments I will present today hold, this is a false assumption.
Notice that there is NO APPEAL TO "HOMOPHONY" here – any more than there's an appeal to "homophony" in:

(3) Kim picked up the hammer and hammered the metal into a blade.
Kim picked up the **hammer** and **hammered** the metal into a blade.

**importantly, the verb "hammer" is indeed non-compositional**
Q: Okay, but is this mode of explanation "interesting"?

A: That's the wrong question to ask. We're in the business of neither mathematics nor aesthetics, but of cognitive science.

If these kinds of representations are available to the child, those who would claim that she doesn't use them are on the hook to explain why.
A Methodological Note:

The discussion of MEANING in this talk will mostly center on open-class items.

Whereas most formal semantics these days is about closed-class items.

⇒ Problem…?
The focus on **closed-class** items in formal semantics is merely a *heuristic* choice.

**CENTRAL IDEA:**
**Open-class** items *(dog, beauty)* will involve the same principles & mechanisms as **closed-class** items *(every, the)*. But we have a better guess for what the latter mean…
Thus, by parity of reasoning:

*If* we're able to learn something about interpretation & meaning from **open-class** items —

It should be taken to be general, as well, and apply to **closed-class** items too.
PRELIMINARIES:

(I) The term "word" is not useful in the context of FORM-MEANING relations. (Marantz 2001, i.a.)
PRELIMINARIES:

(I) The term "word" is not useful in the context of FORM-MEANING relations. (Marantz 2001, i.a.)

(a) "phonological word" is not suited to serve as the relevant notion of "word"
There's (probably) such a thing as "phonological words" —

but phonological words can correspond to *composed* meanings:

[ðə.'dɒɡ]
“the dog”

they need not even be *constituents*:

[ðeɪd.bɪj.hɪ.r]  
“They’d be here.”
PRELIMINARIES:

(I) The term "word" is not useful in the context of FORM-MEANING relations. (Marantz 2001, i.a.)

✓ (a) "phonological word" is not suited to serve as the relevant notion of "word"

(b) and neither is "orthographic word"
There are (sometimes) such things as orthographic words…
BUT:

- speaks English
- doesn't know how to read/write

⇒ doesn't know "words"?

⇒ doesn't know units of FORM-MEANING correspondence?
Many writing systems (incl. early Latin & Greek) lacked spaces altogether

⇒ no "words"...?

"scriptio continua"
Furthermore:

- The writing system for modern-day Vietnamese, for example, has spaces – but they individuate ~syllable-sized units smaller than anything that could realistically be called "word" in the language (Noyer 1998)

- and, of course, not every natural language even has a writing system
PRELIMINARIES:

(I) The term "word" is not useful in the context of FORM-MEANING relations. (Marantz 2001, i.a.)

✓ (a) "phonological word" is not suited to serve as the relevant notion of "word"

✓ (b) and neither is "orthographic word"
At this juncture, one typically launches a final attack on any remaining, "intuitive" notion of word (see, e.g., Marantz 2001).

MEANING:  "chew the fat" (cf. chew, the, fat)  
"believable" (cf. believe, -able)  
"terrific" (cf. terrify, -ic) 

FORM:  дʒəwənə ɓɛt (cf. *dʒəjwənə ɓɛt)  
"ownership" (cf. owner, -ship)  
"cat" (cf. cap, hat, …)
But I've come to believe that this is completely unnecessary —

_In science, we do not need to refute intuitive, nebulous "proto-theories" based on folk-scientific notions._

Unless & until someone presents an explicit, non-phonological non-orthographic definition of "word" that is not post-hoc...
PRELIMINARIES:

✓ (I) The term "word" is not useful in the context of FORM-MEANING relations. (Marantz 2001, i.a.)

✓ (a) "phonological word" is not suited to serve as the relevant notion of "word"

✓ (b) and neither is "orthographic word"
(I) The term "word" is not useful in the context of FORM-MEANING relations.  
(Marantz 2001, i.a.)

(II) Morphological exponents cannot serve as units of FORM-MEANING mapping, either.  
(Aronoff 1976, i.a.)
(a) Just like "chew the fat" requires X-MEANING mapping where $X \succ \"word\"$...

   it also requires X-MEANING mapping where $X \succ$ morphological exponent.

(b) And so does "terrific" (cf. *terrify*, -ic).
(c) **suppletion:**

You provided `go – went` to ask what's the FORM side of the FORM-MEANING mapping, here?

---

Anishinaabemowin (Algonquian);
Sigwan Thivierge, p.c.:

- miskomin-ag
  - raspberry_ANIM-ANIM.PL
  - 1-PST-eat.TA-DIR-ANIM.PL
  - ‘I ate raspberries.’
- miin-an
  - blueberry.INAN-INAN.PL
  - 1-PST-eat.TI-TI3-INAN.PL
  - ‘I ate blueberries.’
(d) forms without meaning:

complete ~ completion
compete ~ *competition (cf. competition)

What is this "extra" -ti/-it?
In particular: what does it MEAN?

Just “morphotactics”…? Not quite…

(Harley 2006)
(d’) in cahoots
short shrift
spick and span

(cf. competition)

PRELIMINARIES:

✓ (I) The term "word" is not useful in the context of FORM-MEANING relations.  

✓ (II) Morphological exponents cannot serve as units of FORM-MEANING mapping, either.

(Marantz 2001, i.a.)

(Aronoff 1976, i.a.)
TOWARDS A THEORY: the role of *late-insertion* at PF

Let's take *go-went* as a representative case:

- $T[^{past}]$ and $\sqrt{GO}$ are separate syntactic terminals
  
  - *remember*: to argue otherwise, one would have to provide a non post-hoc definition of "word"

- Therefore: the choice between *go* and *went* depends on derived syntactic structure
CONCLUSION:

To the extent that knowledge of English includes something like `<FORM, ⬇️>` —

– the "FORM" part of that data structure needs to be a conditional that is informed by – and operates on the output of – the completed syntactic derivation.
Notice that it's not enough for T[PAST] to be "next to" √GO — the two have to stand in a particular structural relationship:

(1) Sprocket **went** home.

(2) ... and once they did, **going/went**ing home was no longer possible.

So, for example, in Distributed Morphology (DM; Halle & Marantz 1993, 1994):

Adjacency in this notation is not "innocent"! It stands for what is, in reality, a structural relationship.

√GO ↔ went / ___ [FINITE, PAST]
√GO ↔ gon / ___ [PTPL]
√GO ↔ go: / elsewhere
There is no FORM, strictly speaking, that can be associated with $\sqrt{GO}$.

At most, what can be associated with $\sqrt{GO}$ (on the FORM side) is a tiny little "syntax engine" —

one which can have various different FORM outputs, depending on which treelet you feed it.
Arguably more perspicuous to just say...:
We'll have more to say about the choice between these two formalizations —

the terminal-centric (DM) one, and the many-to-one mapping one

— in what follows.

But first...
TOWARDS A THEORY: the role of *late-insertion* at LF

What's less often remarked upon: the very same thing we just saw for **FORM** holds for **MEANING**, as well.
Let's take *terrify-terrific* as a representative case:

- *-ic* is the FORM associated with a syntactic terminal (or multiple terminals)
  - remember: to argue otherwise, one would have to provide a non post-hoc definition of "word"

- Therefore: the choice between what *terrify* means on its own and what *terrify*- means when it occurs in the relevant configuration with *-ic* depends on derived syntactic structure
CONCLUSION:

To the extent that knowledge of English includes something like \(<\text{terr(if(y))}-, \text{MEANING}\>\) —

– the "MEANING" part of that data structure needs to be a conditional that is informed by – and operates on the output of – the completed syntactic derivation.
As before, we could encode this via a DM(-like), terminal-centric "rule block":

And as before, \textit{adjacency} in the notation, here, must stand for what is, in reality, a \textit{structural} relationship.

\[
\sqrt{\text{TERR}} \leftrightarrow \begin{array}{c}
\text{thumbs up} \\
/ \quad \text{IC}
\end{array}
\]

\[
\sqrt{\text{TERR}} \leftrightarrow \begin{array}{c}
\text{person sitting} \\
/ \quad \text{elsewhere}
\end{array}
\]
There is **no** MEANING, strictly speaking, that can be associated with $\sqrt{TERR}$.

At most, what can be associated with $\sqrt{TERR}$ (on the MEANING side) is a tiny little "syntax engine" — one which can have various different MEANING outputs, depending on which treelet you feed it.

And as before, **adjacency** in the notation, here, must stand for what is, in reality, a **structural** relationship.
As before, arguably more perspicuous to just say...:
SOME HOUSEKEEPING:
what these facts are not about

These are all (*went, terrific, etc.*) **syntactically complex**
expressions with **FORMS** or **MEANINGS** that are
nevertheless **non-compositional**.

This is not about whether the compositional form does or
does not exist alongside the non-compositional one —

**cf.** *dreamt ~ dreamed*

*transmission* (opaque thing in my car) ~

*transmission* (result or event of transmitting)
This is also (in case there was any lingering doubt...) about sub-"word" vs. super-"word" compositionality —

**cf.**: *terrific*

(blocks access to compositional meaning)

*transmission*  
(doesn't block access to compositional meaning)

*fat chance*  
(blocks access to compositional meaning; would be an antonym of *slim chance* – but is unavailable)

*kick the bucket*  
(doesn't block access to compositional meaning)
An aside: constituency does not constrain many-to-one mappings at LF

(1) a. I read the shit out of that book.
   b. I chilled the shit out of these beers.

(2) a. I read [the shit] [out [of [that book]]].
   b. I chilled [the shit] [out [of [these beers]]].

**cf.:**

(3) I drank [a [quick [cup of coffee]]].
(4) This book had [the shit] read [out of] it.
INTERIM SUMMARY:

What we've seen so far:

– Syntactic terminals don't "have forms".
– Syntactic terminals don't "have meanings".

– At best, syntactic terminals are associated with:
  – a context-sensitive spellout mechanism that determines their contribution to form; and
  – a context-sensitive interpretation mechanism that determines their contribution to meaning
As a result, we can now be quite certain that things like the following are not legitimate parts of a theory of grammar:

(1) *Terminal Nodes* (TN)
    If $\alpha$ is a terminal node, $[\alpha]$ is specified in the lexicon.  

*(Heim & Kratzer 1998: 43)*

And things like the following carve out a particular subset (in DM's terms: *elsewhere* rules; in our terms here, many-to-one mapping rules where 'many' happens to equal 1), *for reasons that are never justified*...

(3) *Terminal Nodes* (TN)
    If $\alpha$ is a terminal node, then $\alpha$ is in the domain of $[\ ]$ if $[\alpha]$ is specified in the lexicon.  

*(Heim & Kratzer 1998: 48)*

... and more importantly, without providing the complementary mechanism for the (as we will see) many, many cases that do **not** fall under (3).
THE QUESTION THAT REMAINS IS:
Which of these two models is better?

**Adjacency** in this notation is not "innocent"! It stands for what is, in reality, a **structural** relationship.

\[
\sqrt{\text{GO}} \leftrightarrow \text{went} / ____ \text{[FINITE, PAST]}
\]

\[
\sqrt{\text{GO}} \leftrightarrow \text{gon} / ____ \text{[PTPL]}
\]

\[
\sqrt{\text{GO}} \leftrightarrow \text{go:} / \text{elsewhere}
\]

And as before, **adjacency** in the notation, here, must stand for what is, in reality, a **structural** relationship.

\[
\sqrt{TERR} \leftrightarrow \text{elsewhere} / ____ \text{IC}
\]

\[
\sqrt{TERR} \leftrightarrow \text{elsewhere}
\]
First, some more data...:

"go off" ~ explode, be triggered

"go" \text{NONPAST} ~ "went" \text{PAST}

"went off" ~ exploded, was triggered
syntactic elements – at minimum:

**PAST** ~ T or Infl or … bearing [+PAST] features

**GO** ~ whatever it is that distinguishes the verb "go" from "run", "dance", etc.

**OFF** ~ whatever it is that distinguishes the preposition/particle "off" from "on", "up", "in", etc.
mappings from syntax to FORM and to MEANING:

/wənt/

PAST
"reference-time is before utterance-time"

/go/

GO
“explode, be triggered”

/ɔf/

OFF
Polish (Slavic);
Asia Pietraszko, p.c.:

a. Bierz się w garść!
take.IMPF.IMP.2SG REFL in handful
‘Pull yourself together (imperfective)!’

b. Weź się w garść!
take.PRFV.IMP.2SG REFL in handful
‘Pull yourself together (perfective)!’
German (Germanic);
Hagen Blix, p.c.:

a. mit jemand-em auf gut-em Fuß stehen with someone-DAT on good-DAT foot stand ‘to get along well with someone’

b. Wir standen damals auf bess-er-em Fuß als heute. we stand.PAST.3PL back.then on good.CMPR-CPMR-DAT foot than today ‘Back then, we got along better than today’
Are these “just” bracketing paradoxes? (No.)

The term *bracketing paradox* describes cases where there’s an apparent mismatch between:
- the constituency needed for morpho-phonology
- the constituency needed for semantics (… and c-selection)

![Diagram](image)

**FORM**

un

happy

er

**MEANING**

er

un

happy
un
  ruly
er

un
  ruly
er

FORM

MEANING

64
This hinges on the question of how FORM representations are related to MEANING representations in the system.
Pesetsky (1985): the answer, at least for cases like *unhappier* and *unrulier*, is Quantifier Raising.
But, crucially for Pesetsky (1985), this movement relation disrupts idiomatic (= semantically noncompositional) readings:

\textbf{compare:} a rarity vs. *a severity
But idiomaticity (=semantic noncompositionality) is precisely what’s going on in GO OFF.
⇒ Instead:

NB: contiguity
Many-to-one mappings: rare?

At this juncture, a potential concern: are we reducing-to-the-worst-case based on a handful of "unusual" examples?
While the “variety” of meanings in (1) is indeed striking —

an often overlooked fact is that the individual meanings
in, e.g., (2) are no more predictable than their
counterparts in (1).
(1) a. /k-b-f/ + CaCuC kvufim ‘pickles’ (Hebrew)
  b. /k-b-f/ + CCiC kvif ‘road’
  c. /k-b-f/ + Ci(C)CuC kibuif ‘conquest’ Aronoff 2007

(2) a. /x-f-b/ + CaCaC xafa’v ‘think’
  b. /x-f-b/ + CiC(C)eC xif’ev ‘calculate’
  c. /x-f-b/ + hiCCiC hixif’iv ‘consider’

NB1: Every instance of composition that is not exclusively phonological or exclusively semantic is syntactic.

NB2: NB1 is not an "assumption" — it's the only game in town (unless & until someone comes up with a working, cross-linguistic definition of "word"… don't hold your breath!)
Pretty much every open-class item in Semitic involves a *joint* mapping from at least two syntactic terminals – the /CCC root, and the n/v/etc. associated with the template – to a meaning.

**AND REMEMBER:**

*If we're* able to learn something about interpretation & meaning from open-class items —

It should be taken to be general, as well, and apply to closed-class items too.
Gaps, gaps, gaps

in cahoots
newfangled
short shrift
huckleberry
spick and span

cf.: * s-cahoot in
* shrift short
* spick span and

(55) a. John left me in the *lurch*. (*left me abandoned and vulnerable to attack*)
b. He protects his *kith* and kin. (*close relatives*)
c. Mary played *hood*. (*failed to attend school*)
d. We *loused* it up. (*made a mess of it*)
e. You *grossed* me out. (*thoroughly disgusted me*)

In a many-to-one model:

in \textcolor{red}{\text{cahoots}}:

\[
\{ \sqrt{\text{CAHOOT}} \} \rightarrow \times
\]

\[
\{ n, \sqrt{\text{CAHOOT}} \} \rightarrow \times
\]

\[
\{ \text{IN, D[-def], Num[pl], } n, \sqrt{\text{CAHOOT}} \} \rightarrow \text{“engaged in a conspiracy”}
\]
In a terminal-centric (e.g. DM’ian) model:

in cahoots:

\[ \sqrt{\text{CAHOOT}} \rightarrow \text{“engaged in a conspiracy”} / \quad \{\text{IN, D[-def], Num[pl], n}\} \checkmark \]

\[ \sqrt{\text{CAHOOT}} \rightarrow \text{X} / \quad \{\text{n}\} \]

\[ \sqrt{\text{CAHOOT}} \rightarrow \text{X} / \quad \text{elsewhere} \]
Claims:

In a terminal-centric (e.g. DM’ian) model:

A language that lacked any "cahoot"s/"fangle"s/etc. would be a fairly unremarkable computational object.

It would simply be a language for which every syntactic terminal had a "complete" rule block (i.e., one that included an elsewhere rule).

In fact, this would arguably be the most computationally natural state of affairs.

Given that DM, at least, is explicitly built around a Pāṇinian logic.
Whereas in a many-to-one model...

(A) fully abstract syntactic atoms
   (e.g. DOG, PAST, RUN, IN, etc.)

(B) many-to-one rules from sets of
    nodes in (A) to units of FORM

(C) many-to-one rules from sets of
    nodes in (A) to units of MEANING

... absolutely nothing guarantees that for every \( x \) in list (A),
there will happen to be a member of list (C) whose input
is the set \( \{x\} \).

In fact, that would be an extreme edge-case.
Which brings me to ...
What is "lexical acquisition" on this type of model?

traditionally: the child learns a "word" — its form(s), its meaning(s), and its syntactic properties
What is "lexical acquisition" on this type of model?

*traditionally:* the child learns a "word" — its form(s), its meaning(s), and its syntactic properties

*That's not a thing...*
what does “learning /'tej.bl/”
or “learning ”
or ...

amount to, in the proposed architecture?
Let's make the simplifying assumption that the child has successfully done "morphological segmentation"—

i.e., division of the incoming speech stream into morphological exponents.
This means the child has successfully identified the sequence of exponents (e.g.: "F1 F2 F3").

But this still radically under-determines the *structure* that could have spelled out —

and even more so the *meanings* that this structure could have been associated with.
The role of exposure to isolated words in early vocabulary development

M R Brent, J M Siskind

Abstract

Fluent speech contains no known acoustic analog of the blank spaces between printed words. Early research presumed that word learning is driven primarily by exposure to isolated words. In the last decade there has been a shift to the view that exposure to isolated words is unreliable and plays little if any role in early word learning. This study revisits the role of isolated words. The results show (a) that isolated words are a reliable feature of speech to infants, (b) that they include a variety of word types, many of which are repeated in close temporal proximity, (c) that a substantial fraction of the words infants produce are words that mothers speak in isolation, and (d) that the frequency with which a child hears a word in isolation predicts whether that word will be learned better than the child's total frequency of exposure to that word. Thus, exposure to isolated words may significantly facilitate vocabulary development at its earliest stages.
a variety of word types, many of which are repeated in close temporal proximity, (c) that a substantial fraction of the words infants produce are words that mothers speak in isolation, and (d) that the frequency with which a child hears a word in isolation predicts whether that word will be learned better than the child's total frequency of exposure to that word. Thus, exposure to isolated words may significantly facilitate vocabulary development at its earliest stages.

Why?
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>$M_2$</td>
<td>$S_1$</td>
<td>$S_2$</td>
<td>$S_3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_1$</td>
<td></td>
<td>$S_1$</td>
<td>$S_2$</td>
<td>F</td>
</tr>
<tr>
<td>$M_1$</td>
<td></td>
<td>$S_1$</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>
\{M_1, M_2\}\{S_1, S_2, S_3\} \quad F

\{M_1\}\{S_1, S_2\} \quad F

\{M_1\}\{S_1\} \quad F
Quantitative Linguistic Predictors of Infants' Learning of Specific English Words

Daniel Swingley ¹, Colman Humphrey ¹

Affiliations + expand

PMID: 28146333  PMCID: PMC5538897  DOI: 10.1111/cdev.12731

Table 5
Regression Coefficients and Descriptive Statistics of Significant Predictors in the Word-Saying Analysis

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>Exp(coef)</th>
<th>IQR</th>
<th>90–10 R</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total frequency.c</td>
<td>0.2754</td>
<td>1.3171</td>
<td>2.77</td>
<td>4.01</td>
<td>.0005</td>
</tr>
<tr>
<td>Isolated freq.c</td>
<td>0.5197</td>
<td>1.6815</td>
<td>0.00</td>
<td>1.10</td>
<td><strong>.0005</strong></td>
</tr>
<tr>
<td>MLU.c</td>
<td>−0.1147</td>
<td>0.8915</td>
<td>1.00</td>
<td>3.00</td>
<td>.0527</td>
</tr>
<tr>
<td>Duration ratio.c</td>
<td>0.2922</td>
<td>1.3393</td>
<td>0.53</td>
<td>2.42</td>
<td>.1233</td>
</tr>
<tr>
<td>Class(closed)</td>
<td>−0.7239</td>
<td>0.4848</td>
<td>na</td>
<td>na</td>
<td>.1709</td>
</tr>
<tr>
<td>Class(pred.)</td>
<td>−1.5665</td>
<td>0.2088</td>
<td>na</td>
<td>na</td>
<td>.0031</td>
</tr>
</tbody>
</table>

Note. Coef refers to the estimated beta coefficient. Exp(coef) provides the number by which the odds of saying a word should be multiplied given an increase of 1 in the predictor's value. IQR (interquartile range) is the difference in value between the 75th and 25th percentiles for values of the numerical predictors. 90–10R is like the IQR but uses the 90th and 10th percentiles. MLU = mean length of utterance.
Learners attempt to "penetrate" this massive many-to-many-to-many mapping problem by establishing single-exponent (or low-number-of-exponent) foot-holds.

As evinced by their over-reliance on fragmentary ("one-word") utterances.
SUMMARY

I have sketched a grammatical architecture in which *listed knowledge* consists in:

(A) fully abstract syntactic atoms (e.g. DOG, PAST, RUN, IN, etc.)

(B) many-to-one rules from sets of nodes in (A) to units of FORM

(C) many-to-one rules from sets of nodes in (A) to units of MEANING

NB: contiguity
We also looked at data that seemed more natural to account for in a many-to-one architecture of the kind proposed here — as opposed to a terminal-centric architecture (e.g. Distributed Morphology).

Including:

- cases of "proper partial overlap"
- the "gaps, gaps, gaps" data
Lastly, we saw some data from so-called "lexical acquisition" that showed over-reliance on isolated forms in vocabulary development —

data which, I argued, made perfect sense from the perspective of the architecture proposed here.
METHODOLOGICAL CONSEQUENCES:

Things like...

“What does the word/morpheme w mean?”

“How do speakers (of this language) pronounce the meaning m?”

... belong in the folk-science dustbin alongside “nouns are things” and “verbs are actions.”
The following, in contrast, are ontologically sound questions:

“How is syntactic structure $S$ pronounced?”

“How is syntactic structure $S’$ interpreted?”

But it is incorrect to presume that the chunks of syntax that are mapped onto discrete units of meanings are the same chunks that are mapped onto discrete units of form.
Comparison with other non-terminal-centric architectures

– Nanosyntax (~phrasal spellout)
  Starke (2009), Caha (2019 / to appear)


These are frameworks which —

like the architecture proposed here and unlike DM —

take the operand of syntax-interface mappings to be structured (rather than atomic).
But to the extent that they pair this structured object with both a **FORM** and a **MEANING** —

they are still fundamentally semiotic in their approach:

```
< FORM , , MEANING >
```

We have seen, however, that natural language simply isn't semiotic in this fashion:

/ˈwɛnt/ /əf/

```
PAST GO OFF
"reference-time is before utterance-time" “explode, be triggered”
```
If, however, we consider a modification of these frameworks that eschews the \(<\text{FORM, STRUCTURE, MEANING}>\) triad — in favor of \(<\text{FORM, STRUCTURE}>\) pairs and, separately, \(<\text{STRUCTURE, MEANING}>\) pairs — then we are looking at something much closer to what is being proposed here.

Indeed, the current proposal can be informally characterized as "spanning with dissociated PF- and LF-spans."
Thank you!