The Lexicalist Hypothesis

Andrew Murphy
Universität Leipzig
andrew.murphy@uni-leipzig.de

Question
Where does word formation take place?

1 Introduction

What do we mean by word formation?
Two locations for word formation:
• Syntax
• Before Syntax (Lexicon)
What kind of evidence can help to decide between the two approaches?

2 What is word formation?

• Word formation refers to the process of combining a root with either a root or an affix to create a new word.
• Compounding involves combining two lexical roots to form a new word:
  (1) Compounding:
    a. fire truck
    b. birthday cake
    c. coffee cup
  (2) 

• Derivation combines a roots and affixes to form a word of a different category:

3 Architecture of grammar

• A common view of the architecture of grammar:
  (3) Derivation:
    a. dance-r
    b. joy-ful
    c. sad-ness
  (4) 

• Where does word formation happen?
• Syntax is responsible for building words and phrases.
• There is no dedicated module for word formation (‘syntax all the way down’).
• For example, nominalization such as destruction would be built in the syntax:
An alternative view is that word formation is pre-syntactic:

This means that words are formed before syntax. Such approaches are often referred to as lexicalist. Words are formed in a different module of the grammar. In a pre-syntactic approach, the nominalization destruction would be formed prior to syntax:

There is no syntactic representation of a verbal category.

4 The Lexicalist Hypothesis

The logic of modularity is one of encapsulation:

Modularity
Information relevant to one module is not accessible outside of that module.

The assumption of a pre-syntactic word formation module gives rise to what is known as the Lexicalist Hypothesis.

Lexicalist Hypothesis
Words and phrases are built in separate parts of the grammar

How can we test the Lexicalist Hypothesis empirically?

Three predictions of the Lexicalist Hypothesis:

1. Phrases cannot be the input to word formation
2. Rules of syntax cannot refer to word-internal structure (Lexical Integrity)
3. The formation of words is governed by different principles to the formation of phrases
5 Phrases and word formation

- Recall the architecture of a possible lexicalist model:

\[ \text{LEXICON} \xrightarrow{\text{INTERPRETATION}} \text{WORD FORMATION} \xrightarrow{\text{SYNTAX}} \text{PRONUNCIATION} \]

- The word formation component **precedes** the phrasal syntax.

- **Predictions:**
  - Words are formed before phrases.
  - Phrases cannot be the input to word formation.

- Compounding, for example, does not seem to operate on phrases:

  7. a. \( [N \text{ coffee }] \)
     b. \( [N \text{ coffee }] \ [N \text{ shop }] \)
     c. \( [NP \ [AP \text{ freshly ground }] \text{ coffee }] \)
     d. \( *[NP \ [AP \text{ freshly ground }] \text{ coffee }] \text{ shop }] \)

- The same can be seen for derivational processes:

  8. a. \( [A \text{ happy }] \)
     b. \( [N \ [A \text{ happy }] \text{ -ness }] \)
     c. \( [AP \text{ quite happy }] \)
     d. \( *[N \ [AP \text{ quite happy }] \text{ -ness }] \)

- Examples such as these would seem to support such the idea that phrases cannot be the input to word formation.

- However, there are some problematic cases...

5.1 Synthetic compounds

- A potentially problematic case involves **synthetic compounds**: compounds containing a derived category and its argument.

  9. a. truck drive-r

5.2 Nominalization in Tamil

- **Deverbal nominalization in Tamil** (Subramanian 1988):
  - peecu ‘speak’ peecu-tal ‘speaking’
  - vaLai ‘bend’ vaLai-tal ‘curving’
  - tura ‘renounce’ *tura-tal ‘renouncing’
  - mara ‘forget’ *marat-tal ‘forgetting’

- **Nominalization in Tamil** (Subramanian 1988):
  - illaram tura illaram tura-ttal family renounce family renounce-nmlz ‘become a hermit’ ‘becoming a hermit’
  - nanri mara nanri mara-tal gratitude forget gratitude forget-nmlz ‘be ungrateful’ ‘ingratitude
  - nilatt-ai uRu nilatt-ai uRu-tal land-acc plow land-acc plow-nmlz ‘plow the land’ ‘plowing the land’

- Does tal attach to V^0s or VPs?
-tal attaches to VPs, rather than V₀s:

If -tal combined with V₀, the transitivity distinction would be unexpected:

5.3 Verbalization in Indonesian

Indonesian derives verbs with an (active) voice prefix meN- and a transitivizing suffix -kan:

Deadjectival verbs (Stevens & Schmidgall-Tellings 2010):

a. hitam 'black'
   meng-hitam 'to become black'
   meng-hitam-kan 'to blacken sth.'

b. merah 'red'
   meng-merah 'to become red'
   meng-merah-kan 'to redden sth.'

Denominal verbs (Sneddon 1996):

a. cermin 'mirror'
   meN-cermin-kan 'to re/flect sth.'

b. bukti 'proof'
   meN-bukti-kan 'to prove sth.'

c. pasar 'market'
   meN-masar-kan 'to take sth. to market'

d. penjara 'jail'
   meN-menjara-kan 'to send so. to jail'

e. libur 'holiday'
   meN-libur-kan 'to send so. on holiday'

Indonesian also allows prepositional phrases to undergo verbalization:

PPs inside words (Stevens & Schmidgall-Tellings 2010):

ke Jakarta 'to Jakarta'
   meN-ge-jakarta-kan 'to send to Jakarta'

ke rumah 'to home'
   meN-ge-rumah-kan 'to send home'

ke bumi 'to ground'
   meN-ge-bumi-kan 'to lower, take down'

ke pingir 'to the edge'
   meN-ge-pingir-kan 'to move to the edge'

Interrogative PPs can form the input to verbalization:

V₀ to where they ACT-take thing DEM

Where are they taking those things?

V₀ to where-TRANS garbage DEM

Where will father take this garbage?

5.4 Quotations

We find what looks like full clauses inside compounds:

a. He baked her an ['I love you' cake].

b. Uh oh! John has his ['don't talk to me today' face] again.

c. Don't give me the old ['the dog ate my homework' excuse]!
A common approach is to treat these as lexicalized ‘quotations’ (Bresnan & Mchombo 1995; Pafel 2015).

However, such ‘quotations’ are actually quite productive (Bruening 2018):

(23) a. How to end your ['I don't feel like it' syndrome]
    b. I don't need your ['I don't think that's wise' attitude]
    c. Your ['why can't I bait newbies?' tears] are glorious

One could still maintain that, while productive, their internal structure is opaque (Bresnan & Mchombo 1995).

(24) a. John, said "I /*spkr love you"
    b. John, said that I /*spkr love you

(25) John baked her an ['I *spkr love you' cake]

However, this is not completely true (Bruening 2018):

(26) a. Max, has that ['talks to himself, on the bus' look]
    b. Don't give me the old ['the dog, ate my homework' excuse] because I know you don't have one;
    c. He baked me an ['I love you' cake], but I don't think he actually does [vp love me]

Such examples do seem to require clause-like constituents below the word-level.

5.5 Phrases below the word level: Summary

So far, we have seen evidence that there is some evidence suggesting that phrases can form the input to syntax.

(27) ...
7 Movement inside words?

- Can we find movement inside words?

- Possible evidence comes from examples such as unhappier.

(30) unhappier
   a. \[\Lambda\ [\Lambda\ [un-\ happy]\ -er]\]
   b. \[\Lambda\ un-\ [\Lambda\ happy\ -er]\]

- Meaning of unhappier: ‘more [ NOT [ HAPPY ]]; not ‘NOT [ MORE [ HAPPY ]]’

- However, -er does not normally attach to trisyllabic base:

(31) a. *importanter vs. more important
   b. *difficulter vs. more difficult

- We have evidence for both of the structures in (30) (bracketing paradox)

- One way to reconcile these requirements it to posit movement of the affix (Pesetsky 1985):

(32)

(33)

- The affix -er attaches to happy (satisfying the phonological requirements).

- The affix then moves to a higher position where it achieves the correct semantic scope more [ NOT [ HAPPY ]];

- This analysis is not compatible with a strictly lexicalist view.

7.1 Anaphora

- Can anaphoric processes 'look inside' words?

(34) Anaphoric islands (Postal 1969):
   a. Trump\textsubscript{i} was glad that [his\textsubscript{i} followers] were the majority in the room.
   b. Trump-ites were the majority in the room.
   c. *Trump\textsubscript{i} was glad that [him\textsubscript{i}-ites] were the majority in the room.
   d. [Hunters of animals\textsubscript{i}] tend to like them,
   e. *[Animal, hunters] tend to like them,

- On the surface, it seems like words might be opaque for anaphora.

- However, the outbound anaphora is generally possible, but subject to pragmatic constraints
  (Ward et al. 1991):

(35) a. John became a [guitar\textsubscript{i}-ist ] because he thought it\textsubscript{i} was a beautiful instrument.
   b. Next week's [Justin Bieber\textsubscript{i} concert] will be his, last of the year.
   c. Most [Trump\textsubscript{i}-ites] claim that they would vote for him, again.

- Inbound anaphora in English has been argued to be due to the fact that pronouns cannot participate in word formation processes (Sproat 1988):

(36)

- Something like this categorial restriction is required independently:
Other languages do seem to allow anaphoric elements to participate in word formation:

38 **Inbound anaphora in Georgian** (Harris 2006):
Merab-i čamovida tbilisi-ši [ta[w]-is-i-an-eb-tan] ertad
Merb-NOM he.come Tbilisi SELF-GEN-DERIV-PL-with together
'Merab, arrived in Tbilisi together with him-ites.'

39 **Inbound anaphora in Japanese** (Kageyama 2001):
Nagai-san wa [[Washington,-syuu] no kookoo]-de 3-nen-kan osie, ima wa
Nagai-Mr. TOP Washington-state GEN high.school-at 3-years taught now TOP
[[dooo,-syuu] no daigaku]-de manande iru
ANAPH-state GEN university-at studying is
'Mr. Nagai taught at a high school in the state of Washington for three years and is now
studying at university in the aforementioned state.'

*Anaphoric reference below the word-level does seem to be possible.*

7.2 **Deletion**

- Deletion can apply below the word-level (Chaves 2008).
- There are examples of **Right Node Raising:**

40 John loves __, whereas Mary hates books about the Second World War.

41 **[Pre-revolutionary]** and **[post-revolutionary]** France were very different from each other.

- Deletion in the second conjunct is also possible:

42 I thought that your [half-brother] and [half-sister] were living with their common biolog-
ical father.

- This resembles **gapping** at the phrasal level:

43 [I read books] and [you read magazines]

- VP ellipsis requires a matching antecedent VP to be licensed:

44 I [VP enjoy swimming] and you do [VP enjoy swimming] too.

- A deverbal category can license VP ellipsis (Hardt 1993):

45 a. David Begelman is a great laugh-er, and when he does [VP laugh], his eyes crinkle
at you the way Lady Brett’s did in the *The Sun Also Rises.*

b. Today, there is little or no harass-ment of lesbians and gays by the national govern-
ment, although autonomous governments might [VP harass them]

46 a. *That man is a rob-er, and when he does [VP rob], he tries not to make any noise.

b. *That man is a thief, and when he does [VP steal], he tries not to make any noise.*

- If VP ellipsis is licensed by a matching VP node, then deverbal nouns must also contain one:

47

\[
\begin{align*}
\text{N}^0 & \quad \text{N}^0 \\
\text{VP} & \quad \text{VP} \\
\text{V}^0 & \quad \text{V}^0 \\
\text{-er} & \quad \text{-er} \\
\text{rob} & \quad \text{laugh}
\end{align*}
\]

- Is it just about matching strings?

48 a. *My computer never can [VP compute] even the most complicated calculations.

b. *My boat’s propeller didn’t [VP propel] while I was trying to escape.*

8 **Conclusion**

Lexicalist Hypothesis
Words and phrases are built in separate parts of the grammar

- Pre-syntactic word formation predicts:
  - No phrases as input to word formation
  - No application of syntactic rules at the sub-word level (Lexical Integrity)
- We saw that there are challenges to a strong versions of lexicalism.
- However, does a naïve, strictly syntactic view of word formation predict no difference between
  phrasal and word-level processes?

8.1 **A compromise?**

A lexicalist model could, for example, assume that there is a feedback loop between syntax and
morphology (Kiparsky 1982; Stiebels & Wunderlich 1994)
• This would allow for phrases to be the input to word formation.

• Lexical Integrity (i.e. inaccessibility of sub-word units) can be derived by assuming that X⁰ is a boundary for syntax.

• However, there is some mediation between word-level and phrase-level processes (Williams 2007) (i.e. percolation of properties).

(49) $N^0_{[V/ANAPH]}$

 destruction

• This is sometimes referred to as the Firewall Theory (Lieber & Scalise 2007).

• Ultimately, this might be how we can reconcile some of the counter-examples with a theory of Lexical Integrity.

References


