Realizational Morphology
in a Modular Minimalist Grammar

Edward Stabler, stabler@ucla.edu

Mathematics of Language 2025

- **Modular minimalist grammars with realizational morphology
  are briefly motivated and defined**

  modular = relatively independent components separated

  realizational = atoms of syntax not pronounced words

- **Captures generalizations missed by previous MGs
  and many other generative grammars**

Paper, slides, code: https://github.com/epstabler/mol25

1

Something generative grammars systematically miss?

Chomsky&Lasnik'77, Lasnik'00: "*Syntactic Structures* makes the claim that there could be another language just like English but where Affix Hopping is optional. The theory we're looking at now . . . makes the claim that there couldn't be any such language.

Affix Hopping and DO-Support. . . describe but don't capture the. . . generalization: *A stranded affix is no good.* "

Bresnan'00: "To counter the fact that DO is ungrammatical elsewhere, there must be a constraint that penalizes its presence"

Grimshaw'97, Sag'11: "[DO is] necessary whenever it is possible"

Non-lexicalist, late insertion, realizational theories

Kayne'93: "There is no auxiliary selection rule"

Bjorkman'11: "BE is not directly selected for, but is instead inserted to support inflectional material that was unable to combine with a main verb"

Olivier'25: "HAVE and BE are allomorphs"

Kalin&Weisser'25:
"Combining all the evidence. . . the most adequate model:"
   non-lexicalist (syntactic word-building),
   post-syntactic (syntactic atoms have no phonology),
   phonology 'realizes' features but not in lexical increments

Modular vs monostratal

Let **monostratal** = each rule application builds one piece of structure

Examples: CFG, MCFG, CG, CCG, TAG, 1990's MG,...

Let **modular** = each piece of structure respects a number of relatively independent, separately stated constraints

Examples: Most of theoretical linguistics

Here: A (simple, preliminary) modular formulation of MG, with realizational morphology, close to recent Chomskian proposals

A modular grammar

mrg:	accepts/transduces binary trees over finite set of atoms
sel:	checks match of selection features
agr:	checks match of agreement features
hm:	move heads to highest accessible 'strong' positions
lin:	linearize
vi:	vocabulary insertion

The derived **(tree) language** is the range of the composition g

$$g = vi \circ lin \circ hm \circ agr \circ sel \circ mrg$$

1. mrg

mrg: an identity trandsuction on binary trees

$$
\text{mrg } t = \begin{cases} t & \text{if t is an atom} \\ \\ S \text{ (mrg x) (mrg y)} & \text{if } t = S \times y \end{cases}
$$

where the atoms have the form:

$$
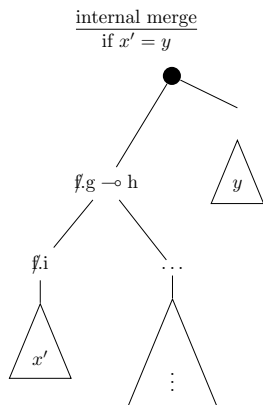\sqrt{}\text{destroy} \quad : \quad \text{D.D} \quad \multimap \quad \text{V}
$$
$$
\sqrt{}\text{destroy} \quad : \quad \quad \quad \quad \text{N}
$$

$\multimap$: 'negative occurrences' of features conjoined on left, positive on right. First pos = 'category'
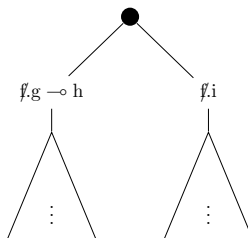
2. sel

$$\text{sel } x = \begin{cases} x & \text{if } x \text{ is labelable} \\ \text{undefined} & \text{otherwise} \end{cases}$$

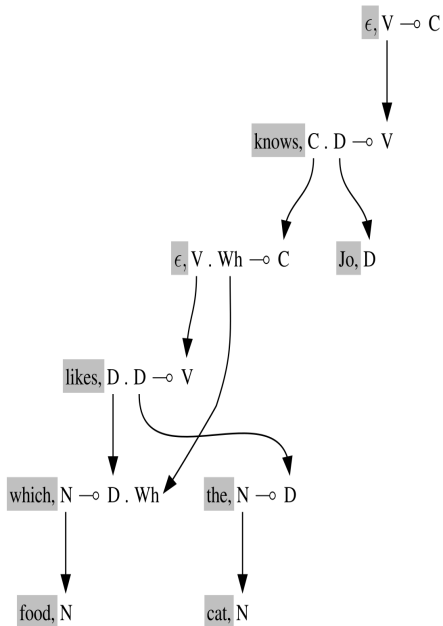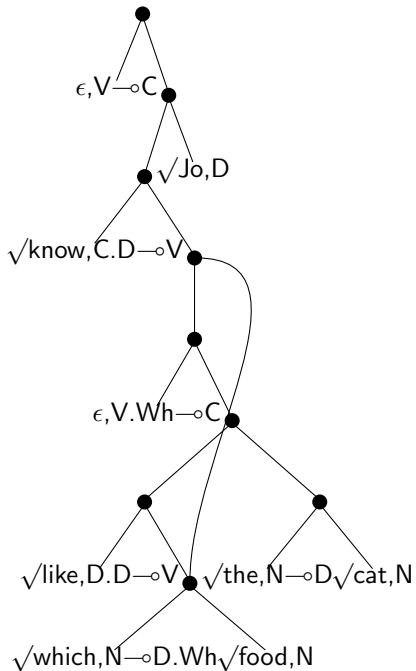• Labeling checks features of opposite polarities, in order...

$$\underline{\text{internal merge}}$$
$$\text{if } x' = y$$

$$\underline{\text{external merge}}$$
if no subtree on left has positive f
as its first unchecked feature

$$\underline{\text{undefined otherwise}}$$

if (smc) no two subconstituents have same first unchecked feature. ₇

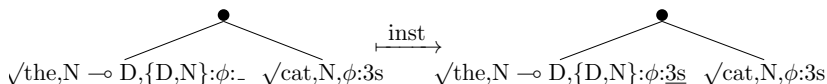note: labeling is efficient, but smc blocks multiple wh extraction – too strict
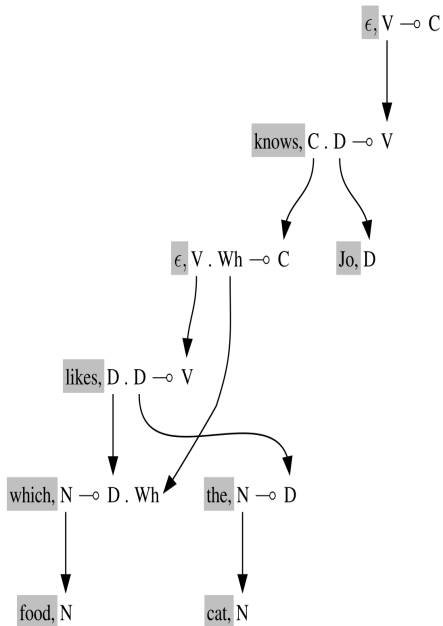
3. agr

$$\text{agr } t = \begin{cases} t & \text{if t can be instantiated} \\ \\ \text{undefined} & \text{otherwise} \end{cases}$$

- Atoms given set of type:value features, each with a tier,
  where tier $\approx$ features of visible heads on spine
- 'probe' $\phi{:}_-$ or $\phi{:}\underline{3s}$ instantiated by 'goal' $\phi{:}3s$,
  in bottom-up order,   (Béjar&Rezak'09)
  in 'tier-adjacent' head   (Hanson'23)

$$\sqrt{\text{the}},\text{N} \multimap \text{D},\{\text{D,N}\}{:}\phi{:}_- \quad \sqrt{\text{cat}},\text{N},\phi{:}3s \quad \overset{\text{inst}}{\longmapsto} \quad \sqrt{\text{the}},\text{N} \multimap \text{D},\{\text{D,N}\}{:}\phi{:}\underline{3s} \quad \sqrt{\text{cat}},\text{N},\phi{:}3s$$
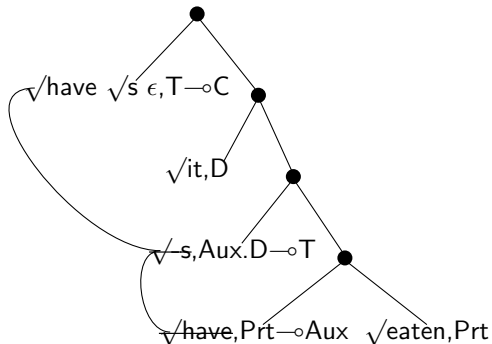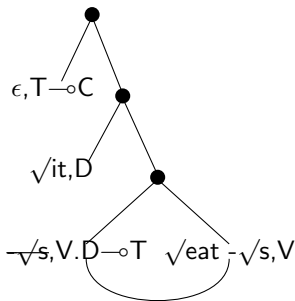
4. hm

Heads can be raised from the phrases they form

4. hm

Heads can also be lowered from the phrases they form

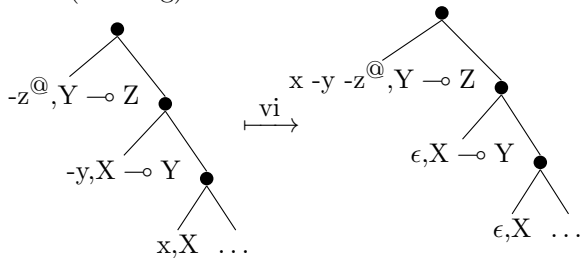Folded in with mrg, sel, and lin, Stabler'01: 5 MG rules $\Rightarrow$

Chomsky&al'23: 'Head Movement' (or at least a subset of phenomena going under that name) may . . . be a post-syntactic operation,. . .

4. hm

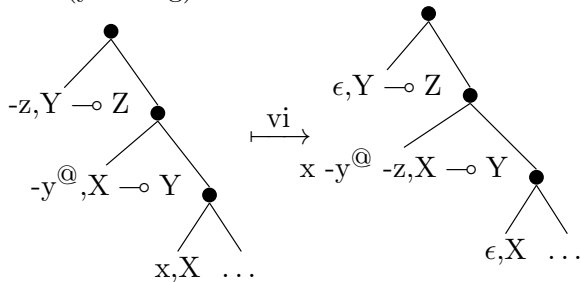Arregi&Pietrazko'21: Head movement determined by 2 diacritics:

- Diacritics: $-\sqrt{}$root is 'dependent', $\sqrt{}$root$^{@}$ is 'strong'

- In a maximal, non-recursive sequence of first-merged heads:
      $-h_0, -h_1, \ldots h_n$
  – Heads 'roll up' span, adjoining to form $h_n, \ldots -h_0$.
  – If no head marked @, heads placed in highest position, h0;
  – Else: all to highest @-marked head.

(z strong)

-z$^{@}$,Y $\multimap$ Z $\overset{\text{vi}}{\longmapsto}$ x -y -z$^{@}$,Y $\multimap$ Z

-y,X $\multimap$ Y ϵ,X $\multimap$ Y

x,X ... ϵ,X ...

(y strong)

-z,Y $\multimap$ Z $\overset{\text{vi}}{\longmapsto}$ ϵ,Y $\multimap$ Z

-y$^{@}$,X $\multimap$ Y   x -y$^{@}$ -z,X $\multimap$ Y

x,X ... ϵ,X ...

5. lin

- (ord) Put first merges head-first, else head-final.
  (Kayne'94,'20, Chomsky'95, Cinque'23)

  (del) Delete non-final internally merged elements[*]

  lin = ord ∘ del

[*] del is too simple –
  e.g. Yuan'25: deletion of non-final copies blocked
  when those copies are needed to host affixes, etc

6. vi

- Vi rules phonologially instantiate roots, based on context

$$\sqrt{\text{cat}} \quad \rightarrow \quad \text{cat}$$
$$\phi{:}3\text{p} \quad \rightarrow \quad \text{-s}$$

  So we transform the leaf

$$(\sqrt{\text{cat}},\text{N},\phi{:}3\text{p}) \quad \Rightarrow \quad (\text{cat -s},\text{N},\phi{:}3\text{p}).$$

- More specific rules take precedence, 'blocking' simpler ones:

$$\sqrt{\text{mouse}},\phi{:}3\text{s} \rightarrow \text{mice}.$$

- vi rules can apply to a complex formed by head movement:

$$\sqrt{\text{chase}} \; \sqrt{\text{past}} \quad \rightarrow \quad \text{chase -ed}.$$

- vi rules can also target spans – sequences of first-selected heads, not marked 'dependent' but in the domain of a vi rule:

$$(\sqrt{\text{de}}) \; (\sqrt{\text{el}}) \quad \rightarrow \quad \text{del}.$$

(Halle&Marantz'93,Embick&Marantz'08,Svenonius'16,Haugen&Siddiqi'16,i.a.)

6. vi

- defaults are negative conditions –
  new in MGs, but not new in grammars

  (Buszkowski'95, Groenink'95, Boullier'98, Kracht'98,...)

- MCFGs lack negative conditions –
  so expressible in Horn clauses, parsable in Datalog

  (Kanazawa)

- But here, the competition domains for defaults are finite in
  number and scope, so still: MCFG equivalent

The modular grammars

mrg:	accepts binary trees over finite set of atoms
sel:	check match of selection features
agr:	check match of agreement features
hm:	move heads to highest accessible 'strong' positions
lin:	linearize
vi:	vocabulary insertion

$$g = \mathbf{vi} \circ \mathbf{lin} \circ \mathbf{hm} \circ \mathbf{agr} \circ \mathbf{sel} \circ \mathbf{mrg}$$

Each particular g specified by <u>atoms</u> and <u>vi rules</u>

mrg, sel and agr are identities; hm, lin and vi preserve hierarchy.

morphology is non-lexicalist, post-syntactic, and realizational.

Nominalization

Chomsky'70: common underlying form → different pronunciations
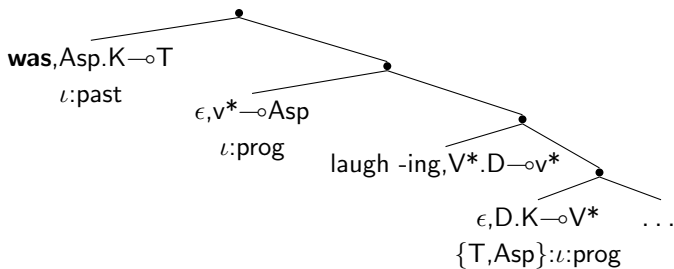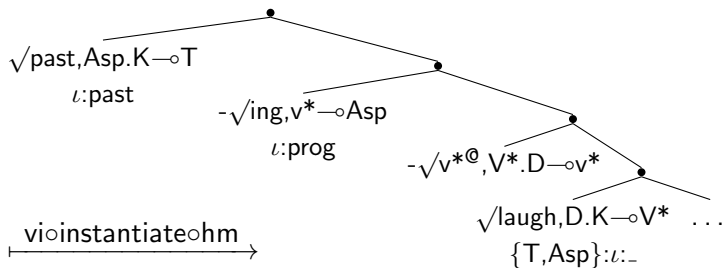
The barbarians destroy the city

The destruction of the city by the barbarians

The oligarchs capture the markets

The capture of the markets by the oligarchs

$$
\begin{aligned}
(\sqrt{capture}) &\rightarrow \text{capture} \\
(\sqrt{destroy}\ V) &\rightarrow \text{destroy} \\
(\sqrt{destroy}\ N) &\rightarrow \text{destruction.}
\end{aligned}
$$

Auxiliaries: English

√past,Asp.K⊸T
ι:past

-√ing,v*⊸Asp
ι:prog

-√v*@,V*.D⊸v*

√laugh,D.K⊸V* ...
{T,Asp}:ι:_

vi∘instantiate∘hm
⟼⟶

was,Asp.K⊸T
ι:past

ϵ,v*⊸Asp
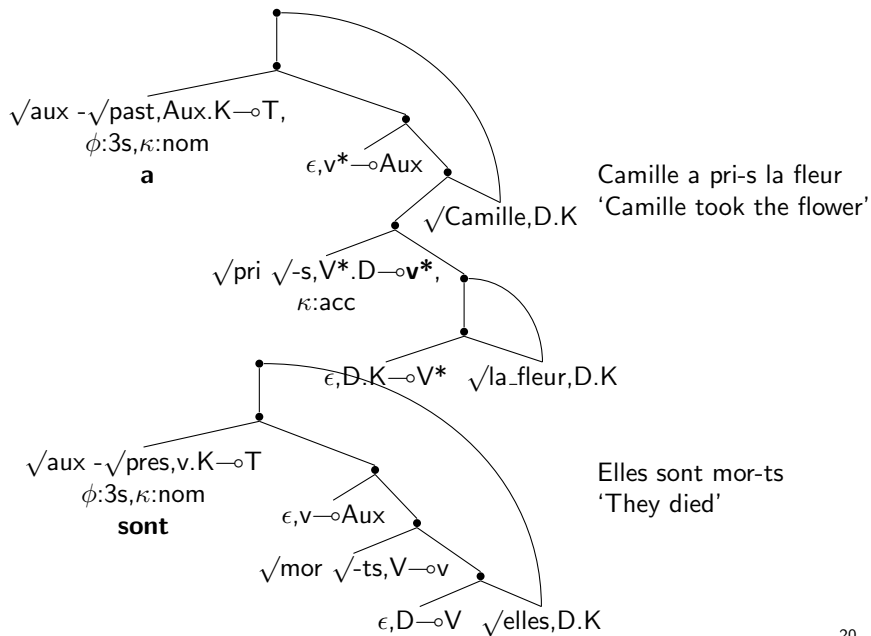ι:prog

laugh -ing,V*.D⊸v*

ϵ,D.K⊸V* ...
{T,Asp}:ι:<u>prog</u>

(Bjorkman'11, Arregi&Klecha'15, Fenger'19, Cruschina&Calabrese'21, i.a.)

Auxiliaries: French (Olivier'25, Bjorkman'11)

√aux -√past,Aux.K⊸T,
ϕ:3s,κ:nom
a

ϵ,v*⊸Aux

√Camille,D.K

√pri √-s,V*.D⊸**v***,
κ:acc

ϵ,D.K⊸V* √la_fleur,D.K

Camille a pri-s la fleur
'Camille took the flower'

√aux -√pres,v.K⊸T
ϕ:3s,κ:nom
sont

ϵ,v⊸Aux

√mor √-ts,V⊸v

ϵ,D⊸V √elles,D.K

Elles sont mor-ts
'They died'

Dependencies among modules

- Dependencies more explicit than in rule-based MGs:
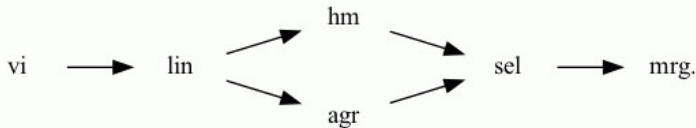
$$\mathbf{g} \;=\; \text{vi} \circ \text{lin} \circ \underline{\text{hm}} \circ \underline{\text{agr}} \circ \text{sel} \circ \text{mrg}$$
$$\text{vi} \circ \text{lin} \circ \underline{\text{agr}} \circ \underline{\text{hm}} \circ \text{sel} \circ \text{mrg}$$

- Future work: unify sel/agr in labeling; vi/hm in interface

  – lin is structurally trivial, so dependency arguments weak

Conclusions

MG properties adjusted

- modular; roots in atoms; sel, agr, lin, hm, vi post-mrg

MG properties preserved: Two conjectures

- Modular grammars (with smc,del) weakly MCF
  – major components already treated (sel, agr, lin); todo: hm, vi
- MCF parsing/learning may extend easily to del, smc replacements
  – re del: easy extension from MCFGs to 'parallel' MCFGs
  – re smc: distinct, stipulated condition can be replaced

New capture of prominent generalizations

- Halle&Marantz: roots, allomorphy, last-resorts
- Chomsky&Lasnik, Bjorkman, Olivier: 'overflow' auxiliaries
- But not: 'Stray affixes are no good'

Alexiadou&Borer'20 Introduction, Nominalization: 50 Years on from Chomsky's Remarks
Bjorkman'11 BE-ing Default: The Morphosyntax of Auxiliaries
Branigan'23 The Grammar of Multiple Head Movement
Bruening'18 The lexicalist hypothesis: Both wrong and superfluous
Boullier'04 Range concatenation grammars
Buszkowski'96: Categorial grammars with negative information
Chomsky'57 Syntactic Structures
Chomsky'70 Remarks on nominalization
Chomsky&Lasnik'77 Filters and control
Collins'02 Eliminating labels
Collins&Kayne Towards a theory of morphology as syntax
Engelfriet,Lilin&Maletti'09 Extended multi bottom–up tree transducers
Ermolaeva&Kobele'22 Agree as information transmission over dependencies
Giannoula'25 Deciphering mirror principle violations
Goto&Ishii'25 Seeking an optimal design of search and merge
Graf'22 Typological implications of tier-based strictly local movement
Graf'23 Subregular tree transductions, movement, copies, traces, and the ban on improper movement
Graf&Kostyszyn'21 Multiple wh-movement is not special
Grimshaw'97: Projection, heads, and optimality
Groenink'95: Literal movement grammars
Halle&Marantz'93 Distributed morphology and the pieces of inflection
Halpert&Zeijlstra'24 Off phases: It's all relative(ized)
Hanson'25 Tier-based strict locality and the typology of agreement
Harizanov&Gribanova'19 Whither head movement?
Haugen&Siddiqi'16 Towards a restricted realization theory
Holmberg'17 The final-over-final condition and linearization in generative grammar

Kalin&Weisser'25 Minimalism and morphology
Kanazawa'07 Parsing and generation as Datalog queries
Kanazawa'09 The pumping lemma for well-nested multiple context-free languages
Kanazawa'17 Parsing and generation as Datalog query evaluation
Kayne'93 Toward a modular thoery of auxiliary selection
Kayne'20 Antisymmetry and externalization
Ke'24 Can agree and labeling be reduced to minimal search?
Keine'25 Φ feature sharing
Kobele&Liu'25 Formalizing feature inheritance
Kracht'98 Strict compositionality and literal movement grammars
Lasnik'00 Syntactic Structures Revisited
Marantz'98 No escape from syntax
Olivier'25 A syntactic account of auxiliary selection in French
Sag'10 Sex, lies, and the English auxiliary system. Stanford
Sag&al'20 Lessons from the English auxiliary system
Stabler'01 Recognizing head movement
Stabler'11 Computational perspectives on minimalism
Stanojević'19 On the computational complexity of head movement and affix hopping
Vanden Wyngaerd&al'15 Late insertion and root suppletion