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

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

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

where the atoms have the form:

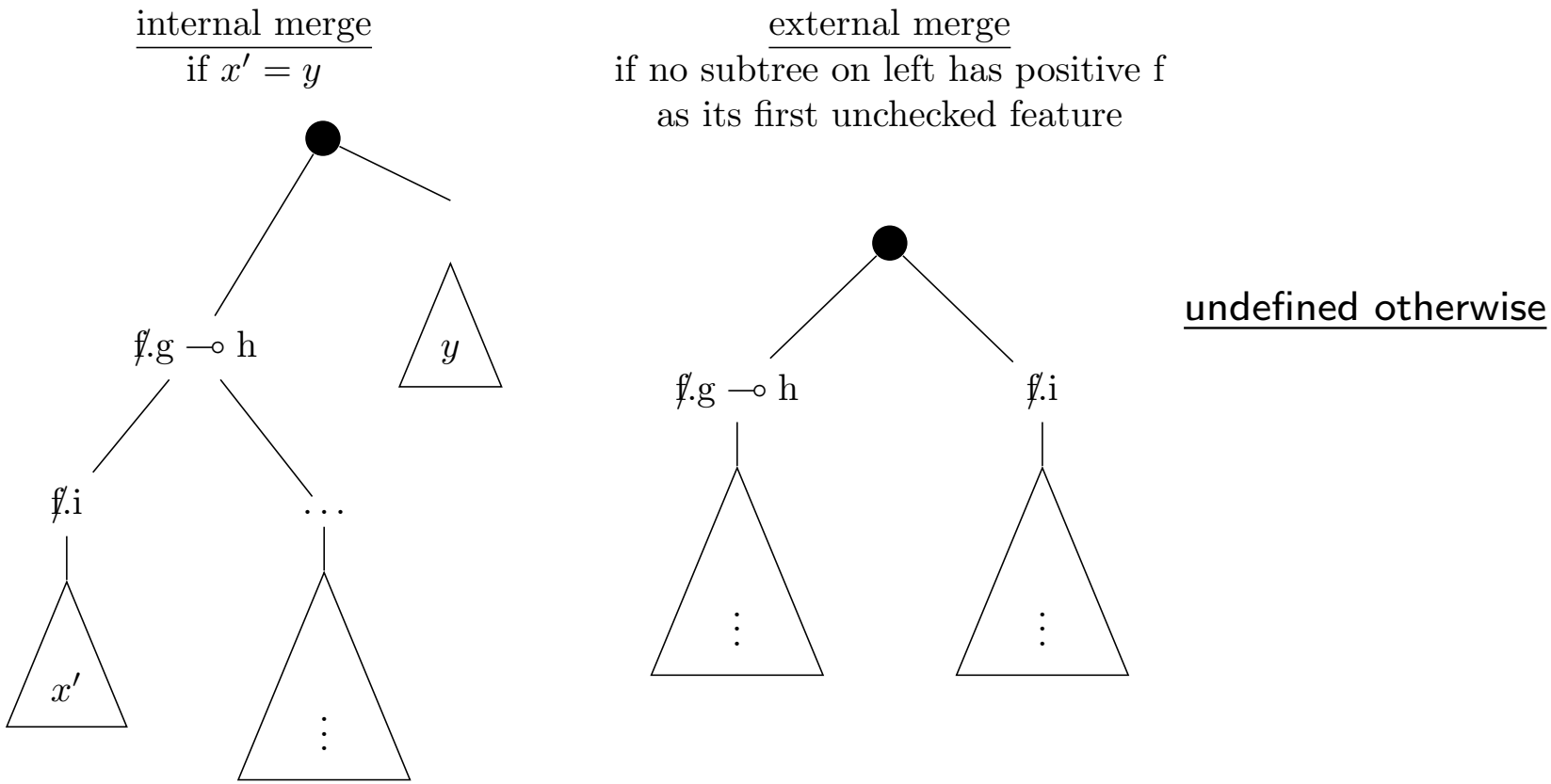
$$\begin{array}{lcll} \surd\text{destroy} & : & \text{D.D} & \multimap \quad \text{V} \\ \surd\text{destroy} & : & & \quad\ \text{N} \end{array}$$

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

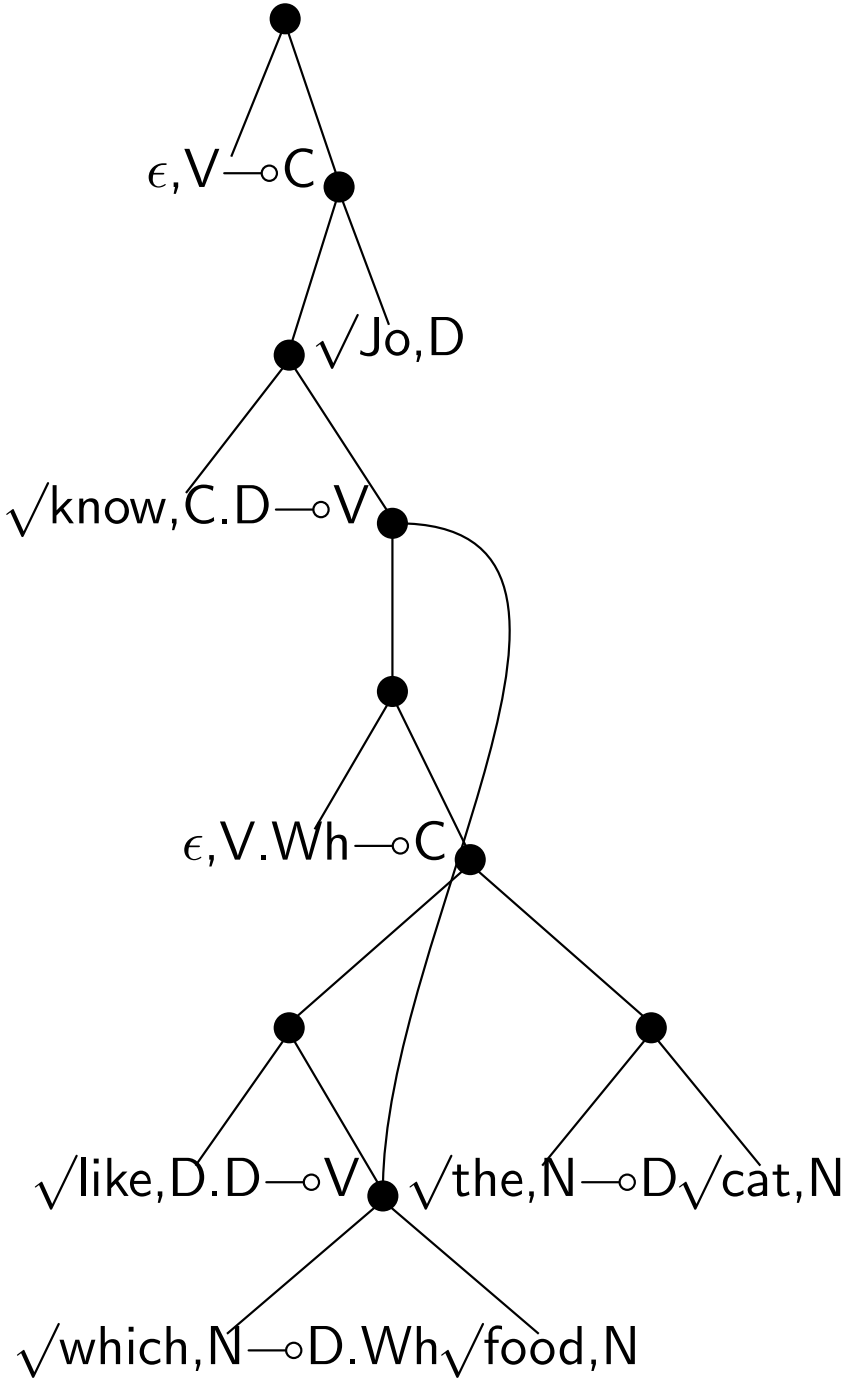
2. sel

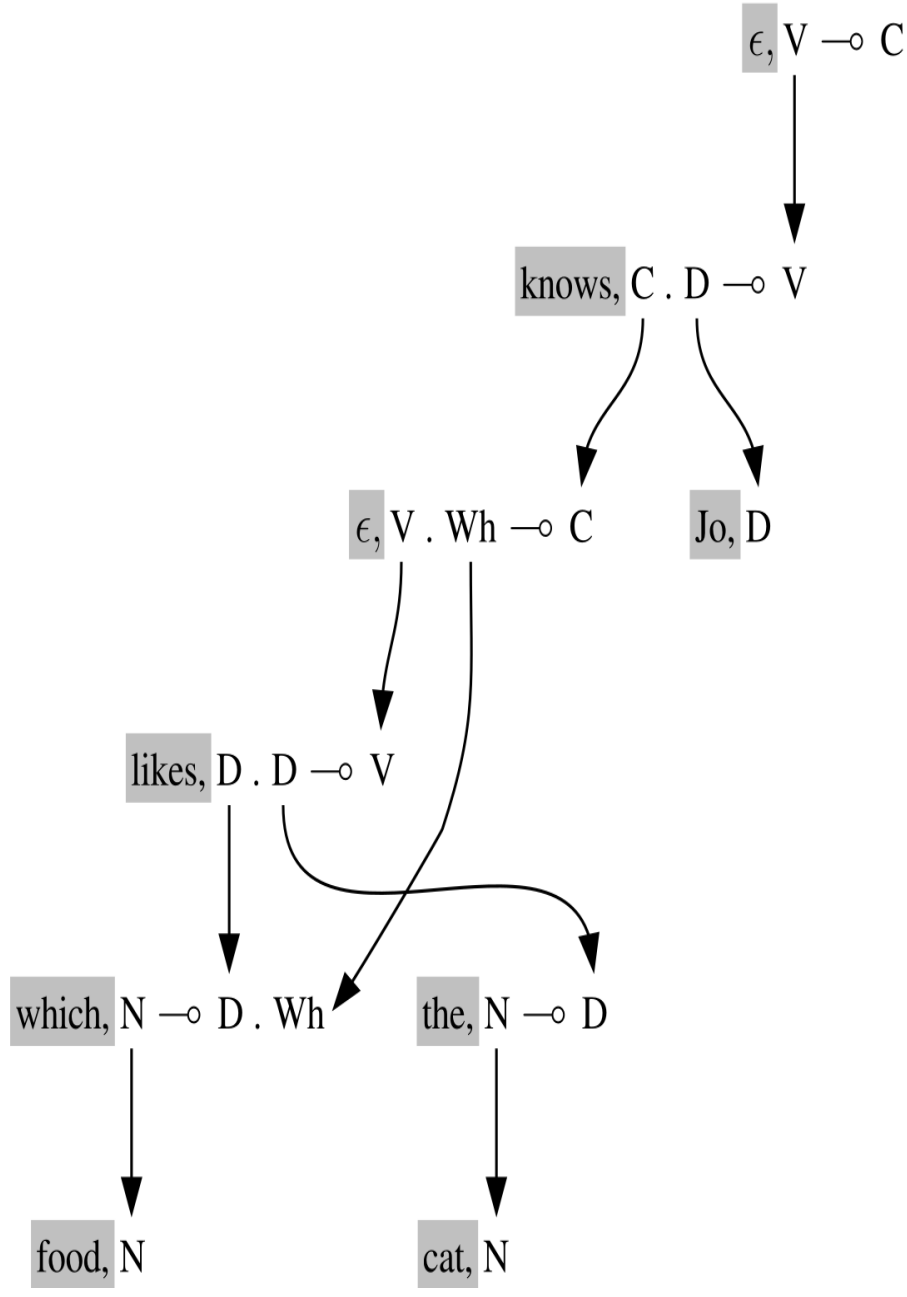
$$\text{sel } x = \begin{cases} x & \text{if } x \text{ is labelable} \\ \text{undefined} & \text{otherwise} \end{cases}$$

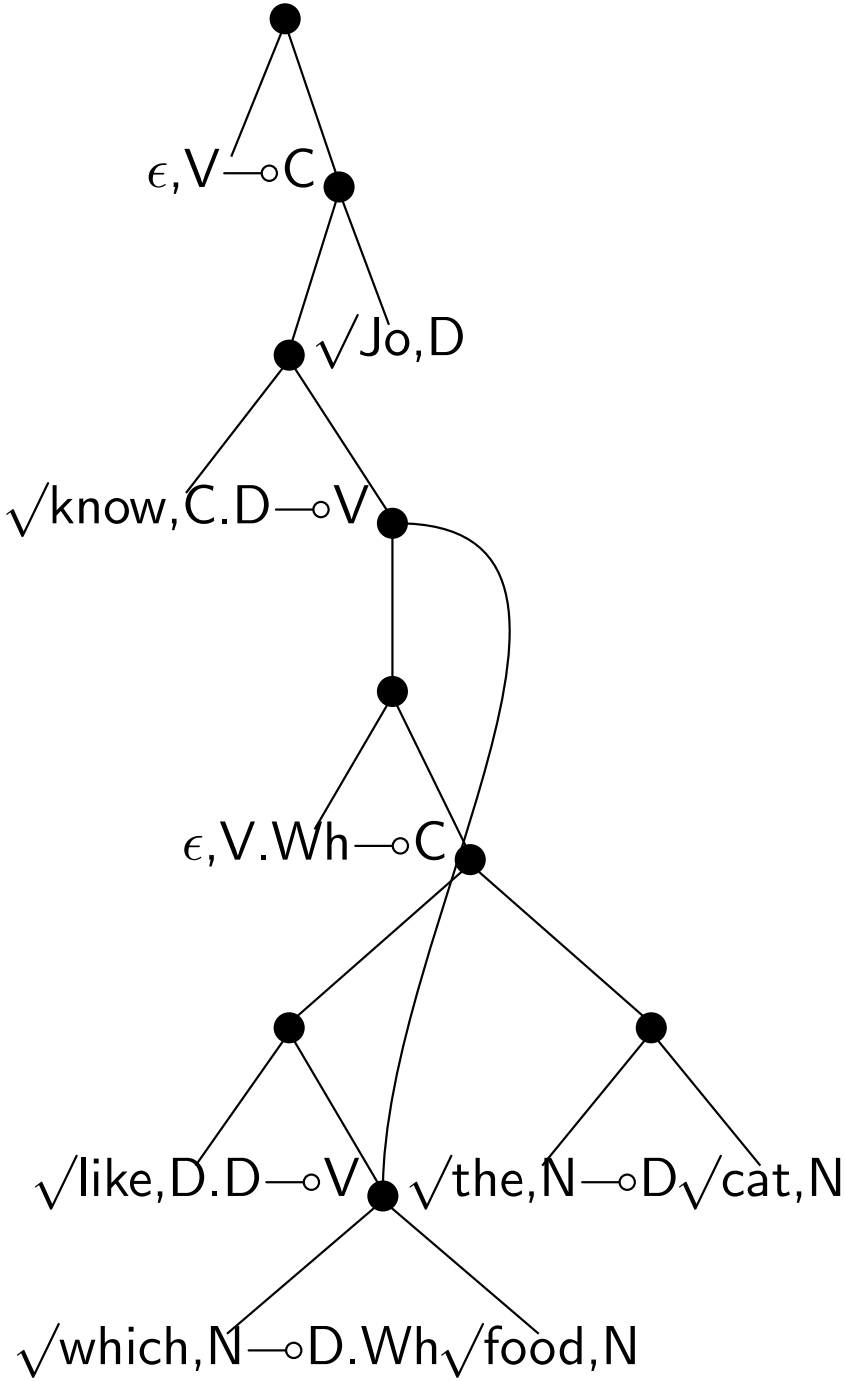
- Labeling checks features of opposite polarities, in order. . .

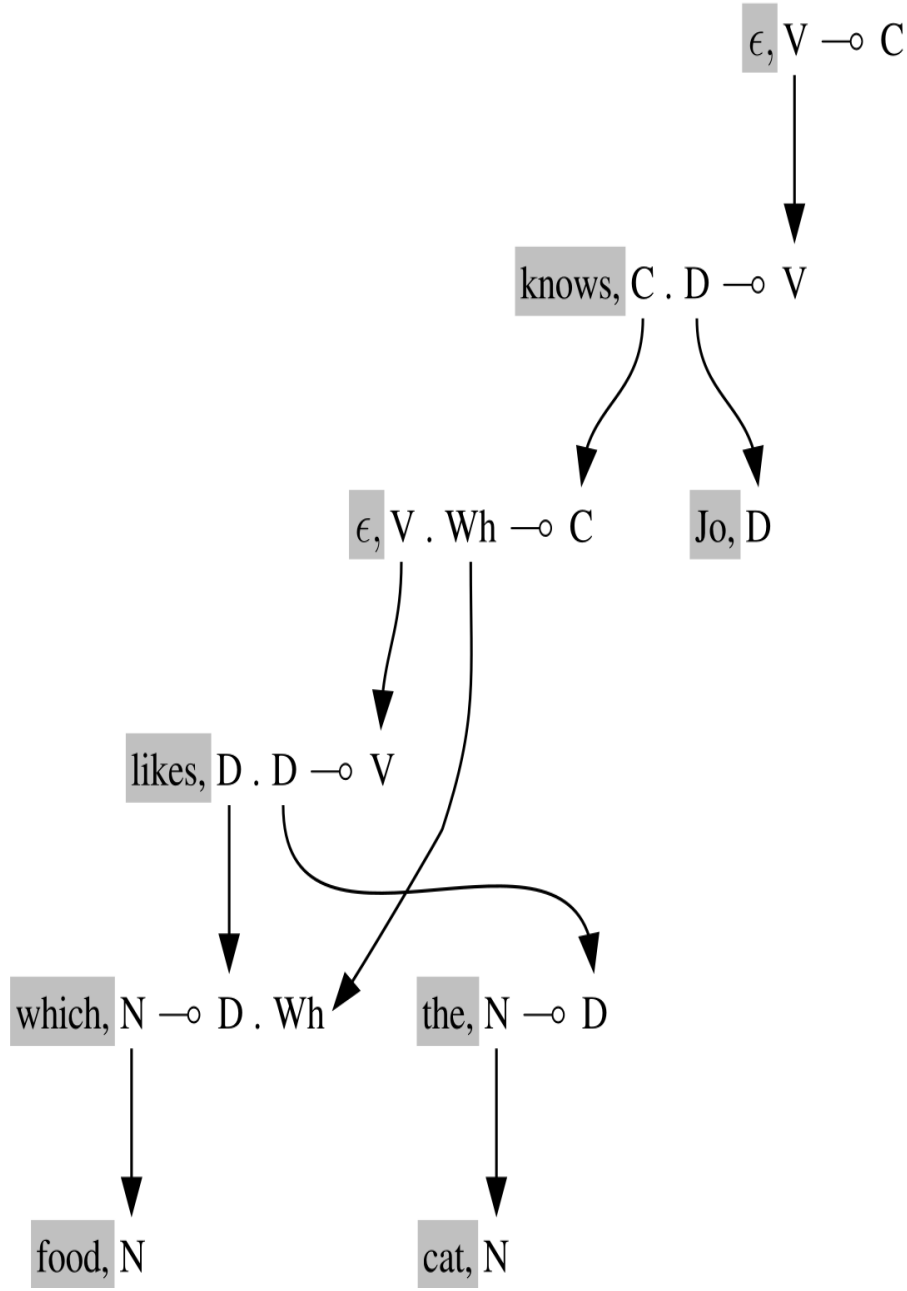


if (smc) no two subconstituents have same first unchecked feature.


ϵ,V—∘C
√Jo,D
√know,C.D—∘V
ϵ,V.Wh—∘C
√like,D.D—∘V
√the,N—∘D
√cat,N
√which,N—∘D.Wh
√food,N


ϵ, V —∘ C
knows, C . D —∘ V
ϵ, V . Wh —∘ C
Jo, D
likes, D . D —∘ V
which, N —∘ D . Wh
the, N —∘ D
food, N
cat, N





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

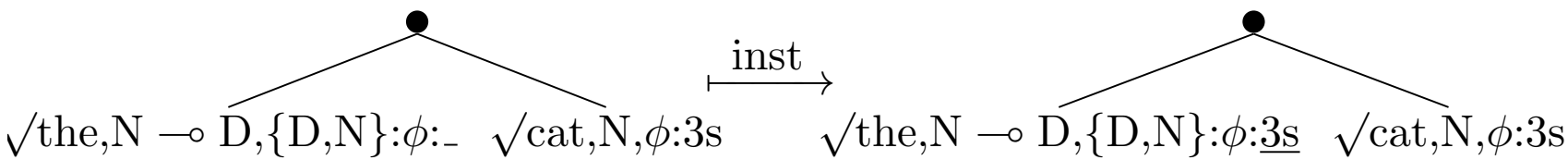
3. agr

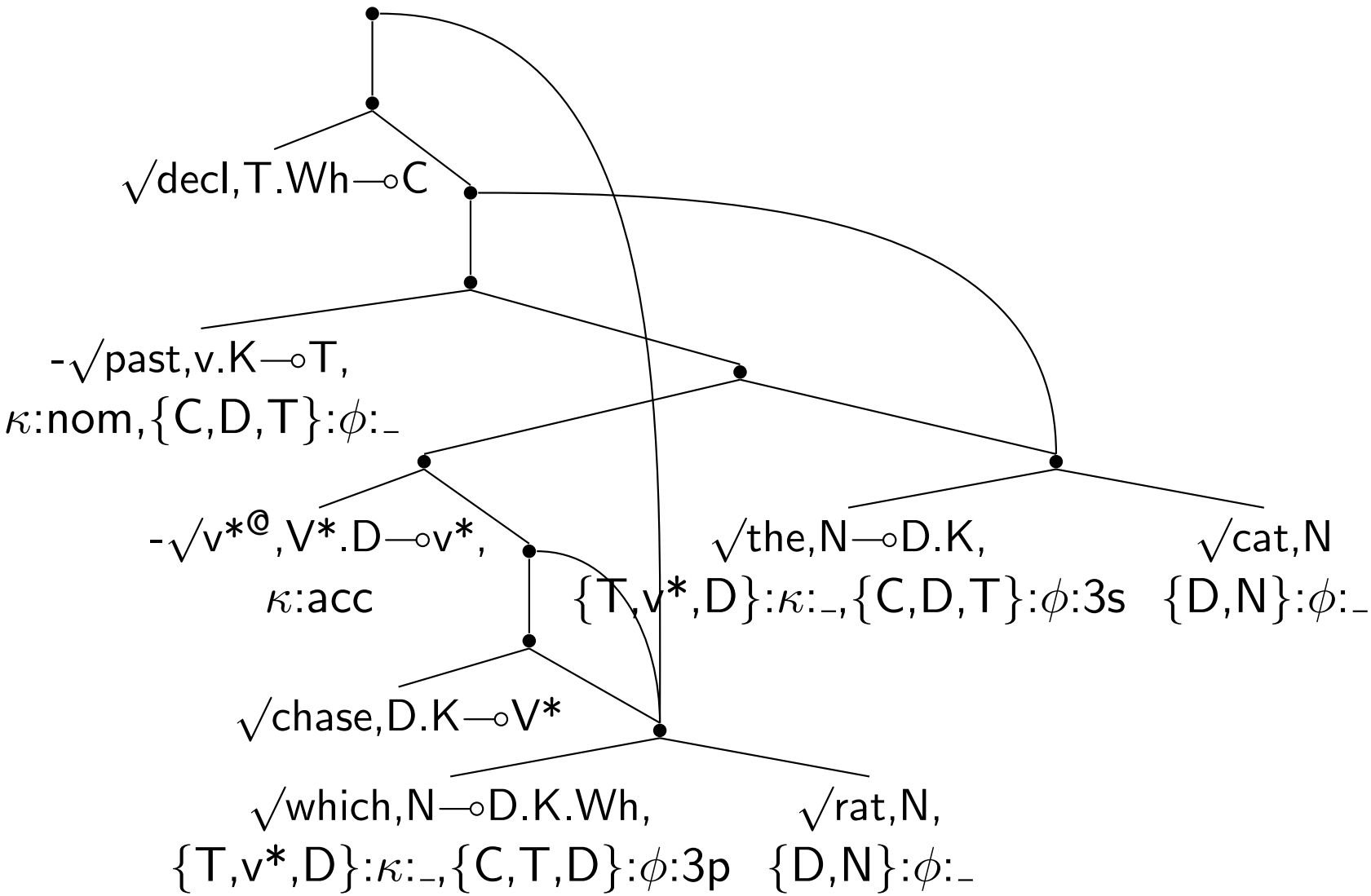
$$\text{agr t} = \begin{cases} \text{t} & \text{if t can be instantiated} \\ \\ \text{undefined} & \text{otherwise} \end{cases}$$

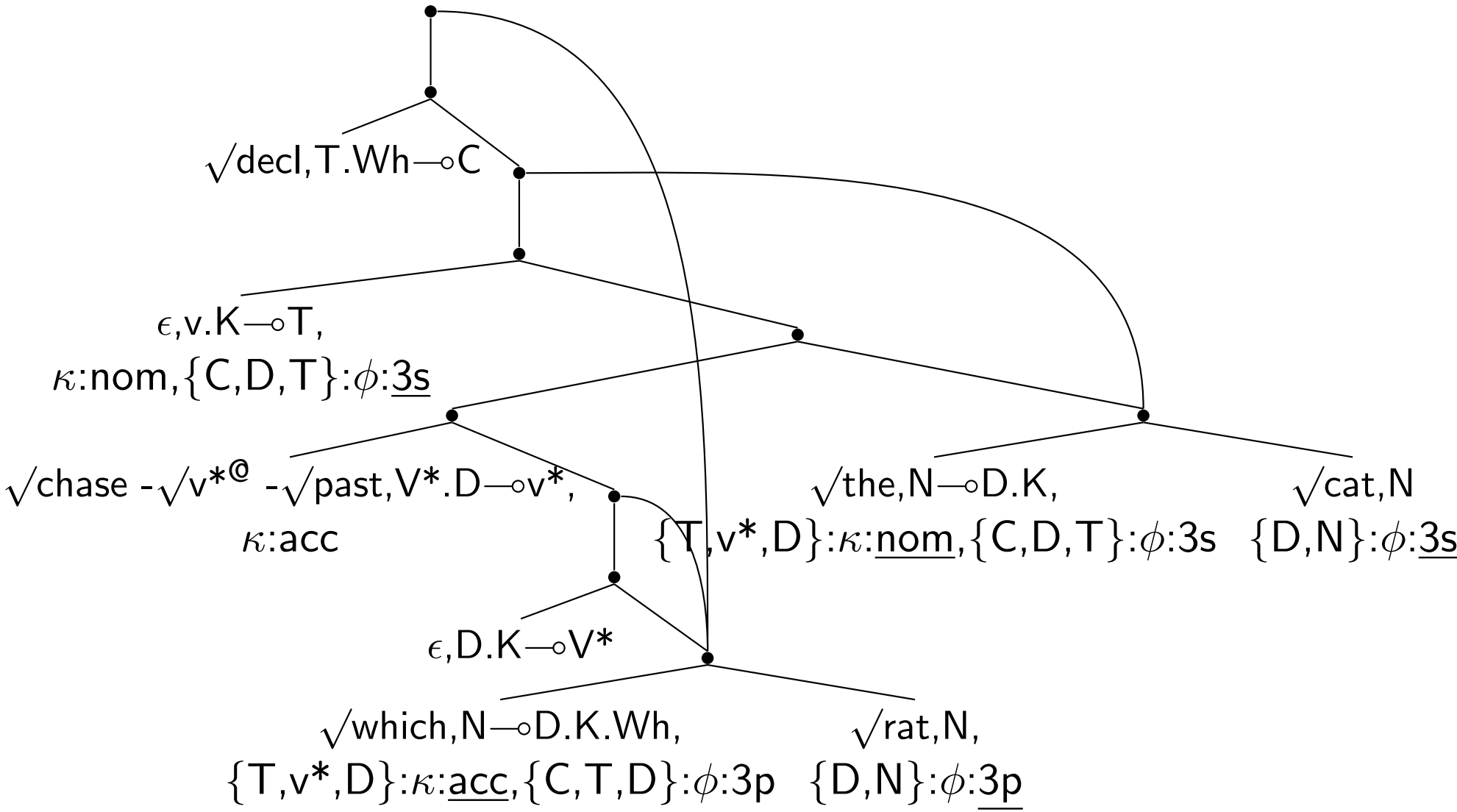
- Atoms given set of type:value features, each with a tier, where tier $\approx$ features of visible heads on spine
- 'probe' ϕ:_ or ϕ:<u>3s</u> instantiated by 'goal' ϕ:3s, in bottom-up order, in 'tier-adjacent' head

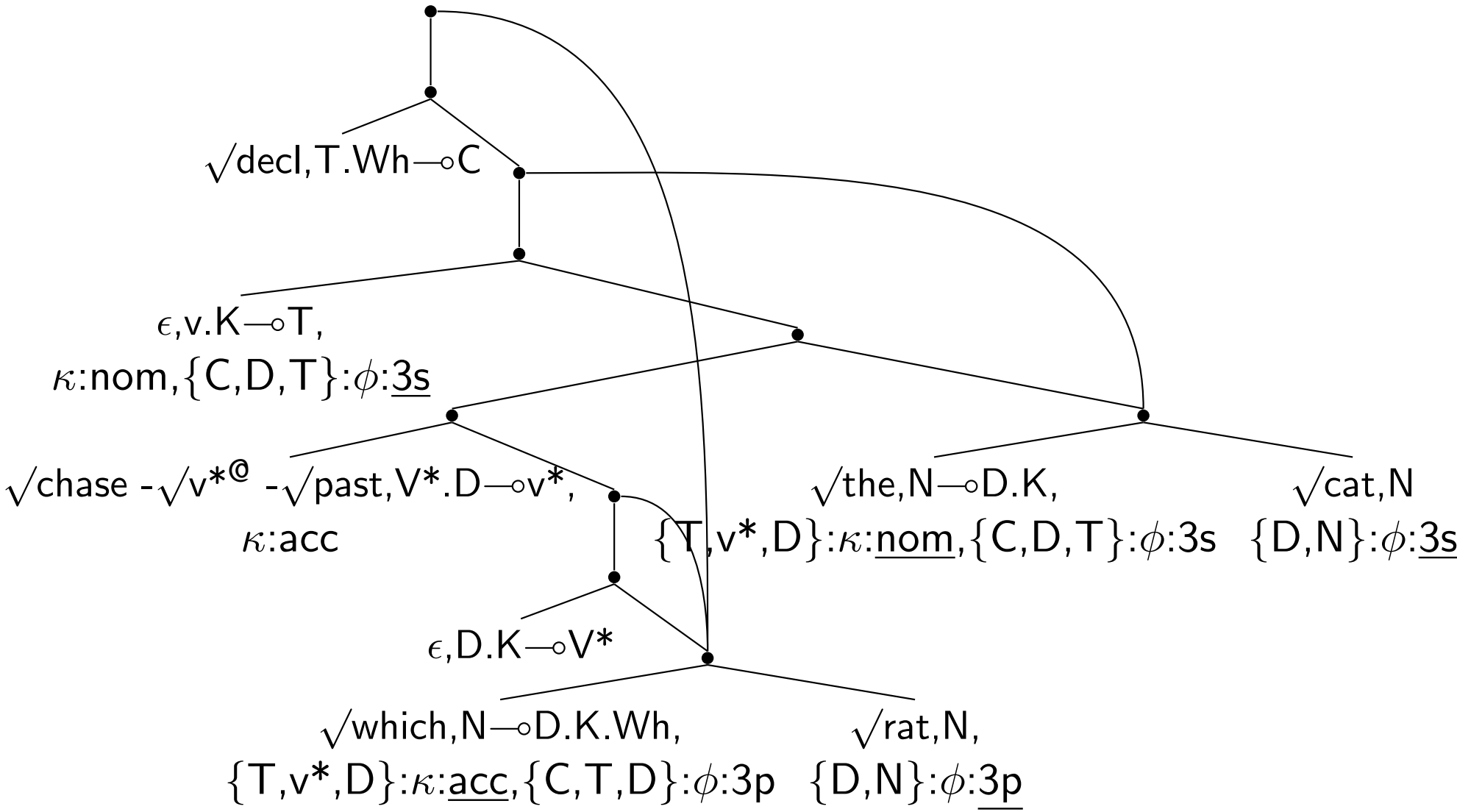
(Béjar&Rezak'09)

(Hanson'23)




√decl,T.Wh—∘C
-√past,v.K—∘T,
κ:nom,{C,D,T}:φ:_
-√v*@,V*.D—∘v*,
κ:acc
√the,N—∘D.K,
{T,v*,D}:κ:_,{C,D,T}:φ:3s
√cat,N
{D,N}:φ:_
√chase,D.K—∘V*
√which,N—∘D.K.Wh,
{T,v*,D}:κ:_,{C,T,D}:φ:3p
√rat,N,
{D,N}:φ:_

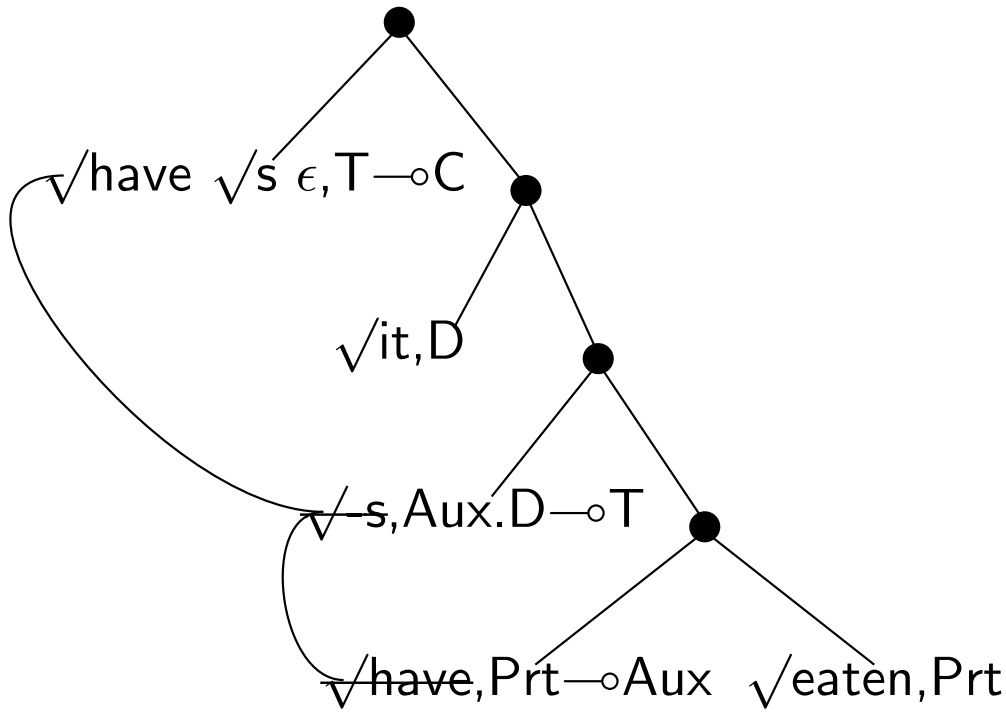

√decl,T.Wh—∘C
ϵ,v.K—∘T,
κ:nom,{C,D,T}:ϕ:3s
√chase -√v*@ -√past,V*.D—∘v*,
κ:acc
√the,N—∘D.K,
{T,v*,D}:κ:nom,{C,D,T}:ϕ:3s
√cat,N
{D,N}:ϕ:3s
ϵ,D.K—∘V*
√which,N—∘D.K.Wh,
{T,v*,D}:κ:acc,{C,T,D}:ϕ:3p
√rat,N,
{D,N}:ϕ:3p



Conjecture: adapting Ermolaeva&Kobele'22, instantiated trees are MCF

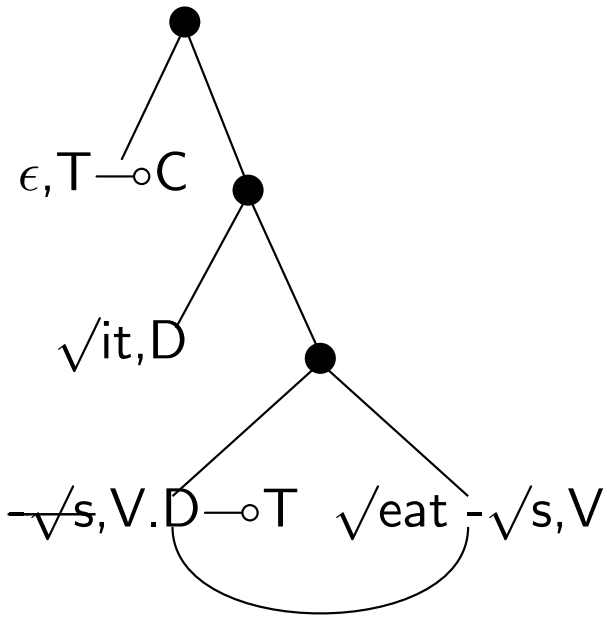
4. hm

Heads can be raised from the phrases they form



4. hm

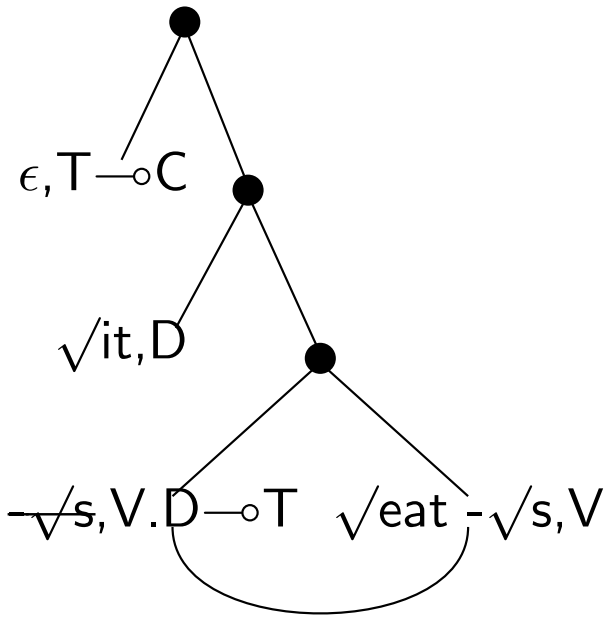
Heads can also be lowered from the phrases they form



Stabler'01: Folded in with mrg, sel, and lin: 5 MG rules $\Rightarrow$ 13

4. hm

Heads can also be lowered from the phrases they form

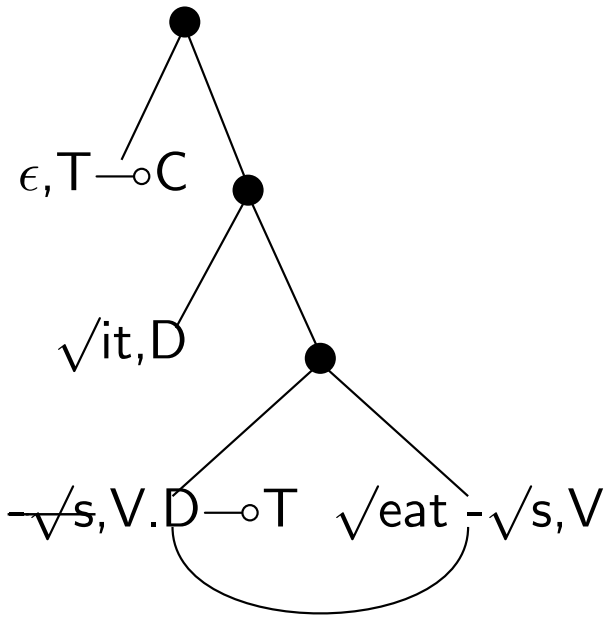


Stabler'01: Folded in with mrg, sel, and lin: 5 MG rules $\Rightarrow$ 13

Chomsky'01: There are some reasons to suspect that a substantial core of head-raising processes. . . may fall within the phonological component.

4. hm

Heads can also be lowered from the phrases they form



Stabler'01: Folded in with mrg, sel, and lin: 5 MG rules $\Rightarrow$ 13

Chomsky'01: There are some reasons to suspect that a substantial core of head-raising processes. . . may fall within the phonological component.

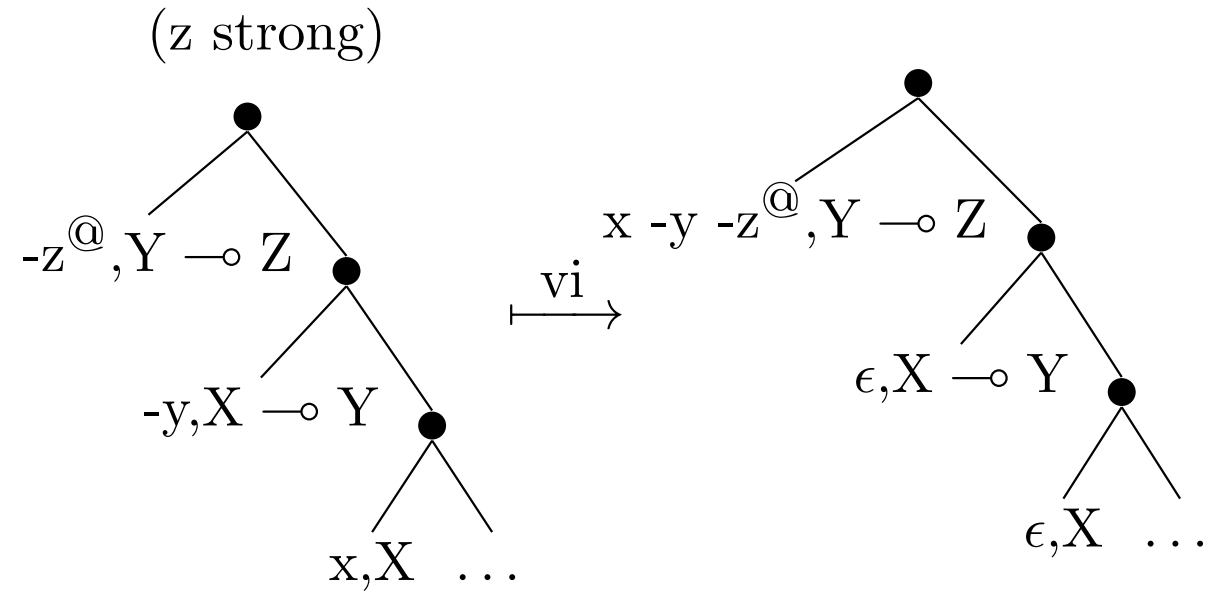
Chomsky&al'23: 'Head Movement' (or at least a subset of phenomena going under that name) may . . . be a post-syntactic operation,. . .

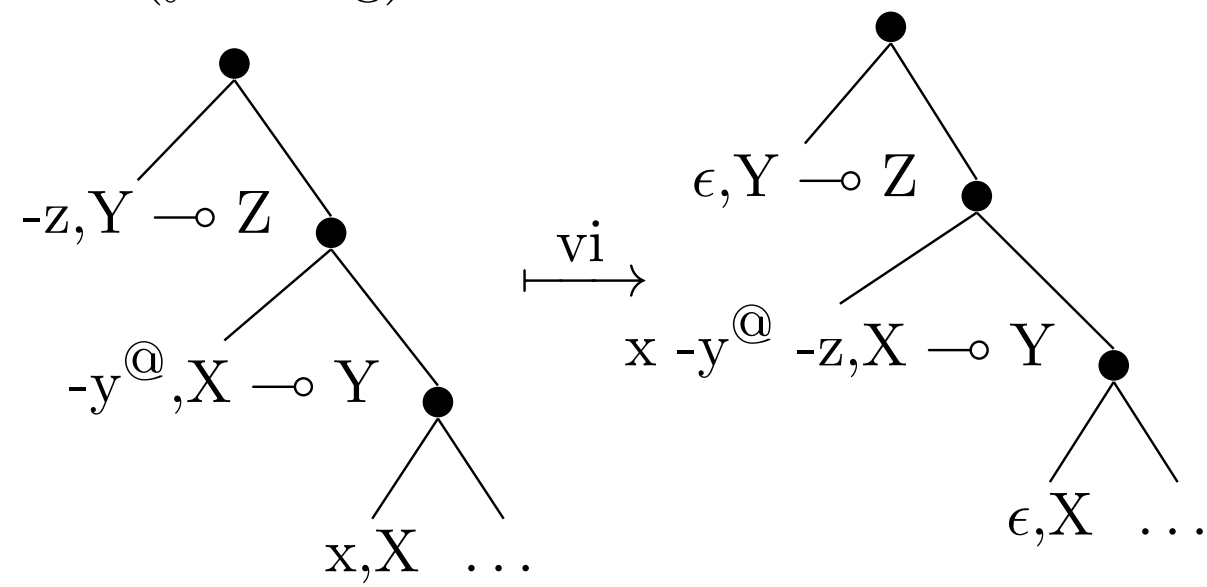
4. hm

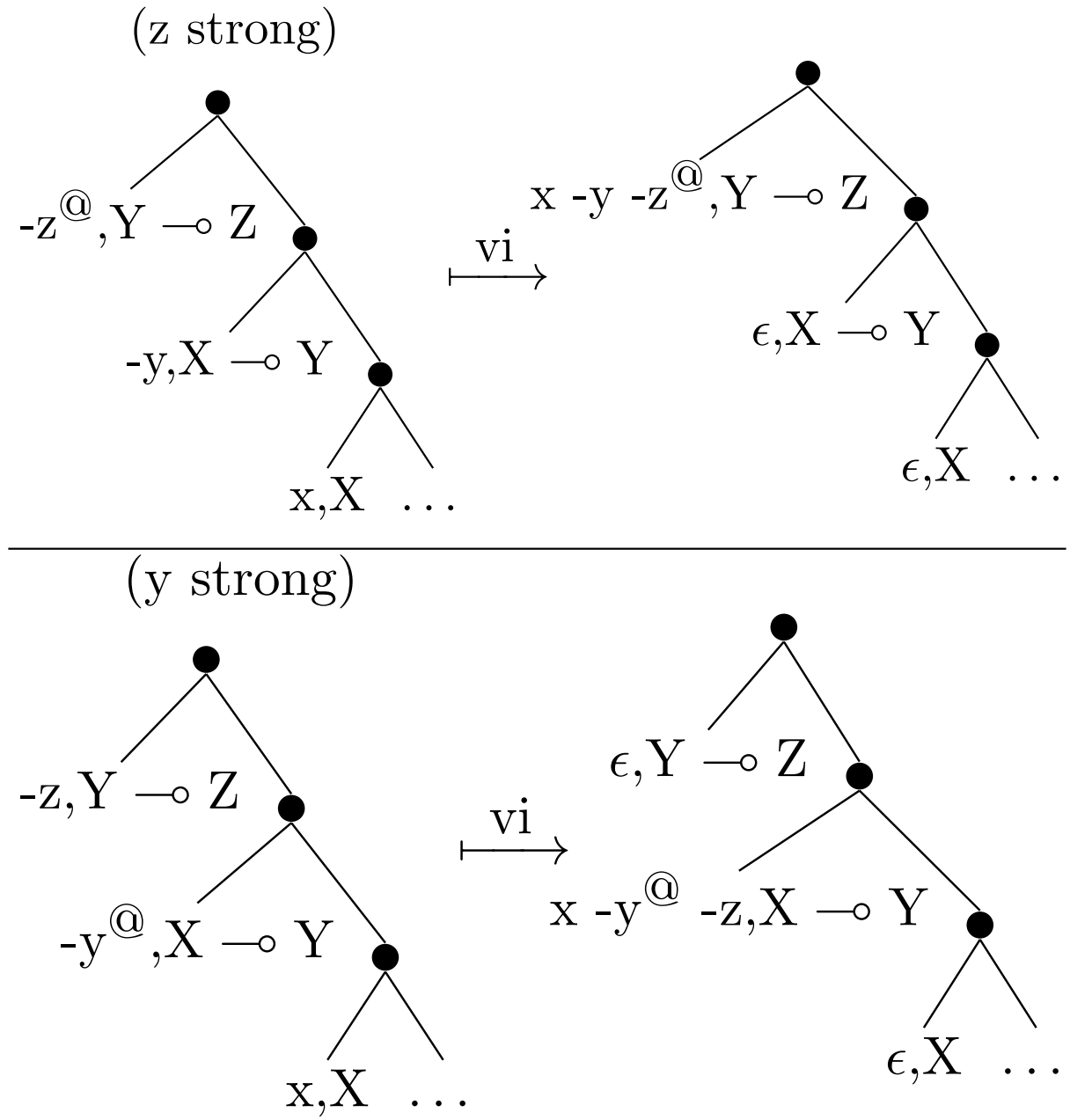
Arregi&Pietrazko'21: Head movement determined by 2 diacritics:

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

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


(z strong)
-z@,Y ⊸ Z
-y,X ⊸ Y
x,X …
vi
x -y -z@,Y ⊸ Z
ϵ,X ⊸ Y
ϵ,X …


(y strong)
-z,Y ⊸ Z
-y@,X ⊸ Y
x,X …
vi
ϵ,Y ⊸ Z
x -y@ -z,X ⊸ Y
ϵ,X …



Conjecture: adapting Michaelis'98, Kobele'02, head moved trees are MCF

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*

  $\text{lin} = \text{ord} \circ \text{del}$

* 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

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

  So we transform the leaf

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

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

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

- 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: vi instantiated trees MCF

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

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

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 $\rightarrow$ different pronunciations

The barbarians destroy the city

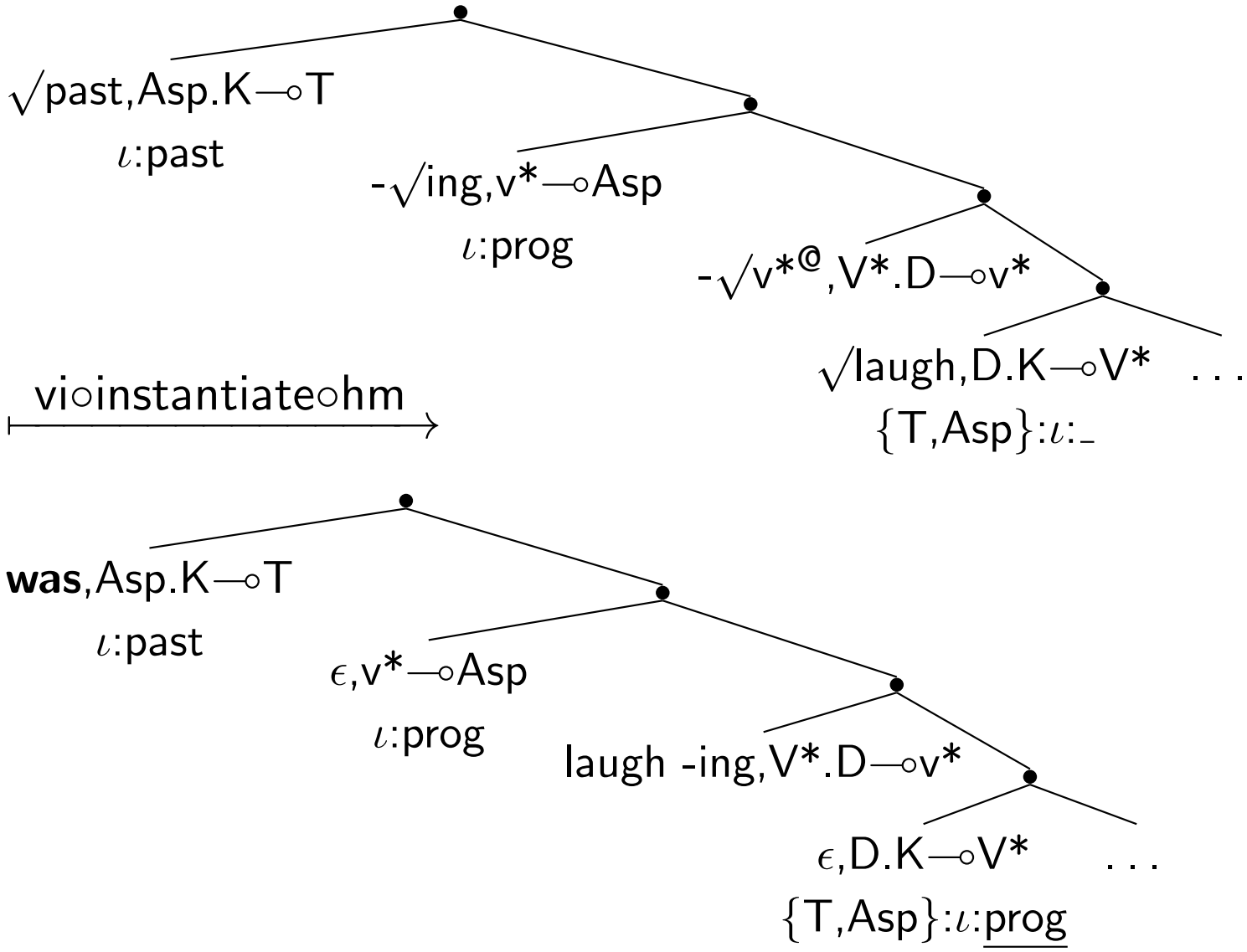
The destruction of the city by the barbarians

The oligarchs capture the markets

The capture of the markets by the oligarchs

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

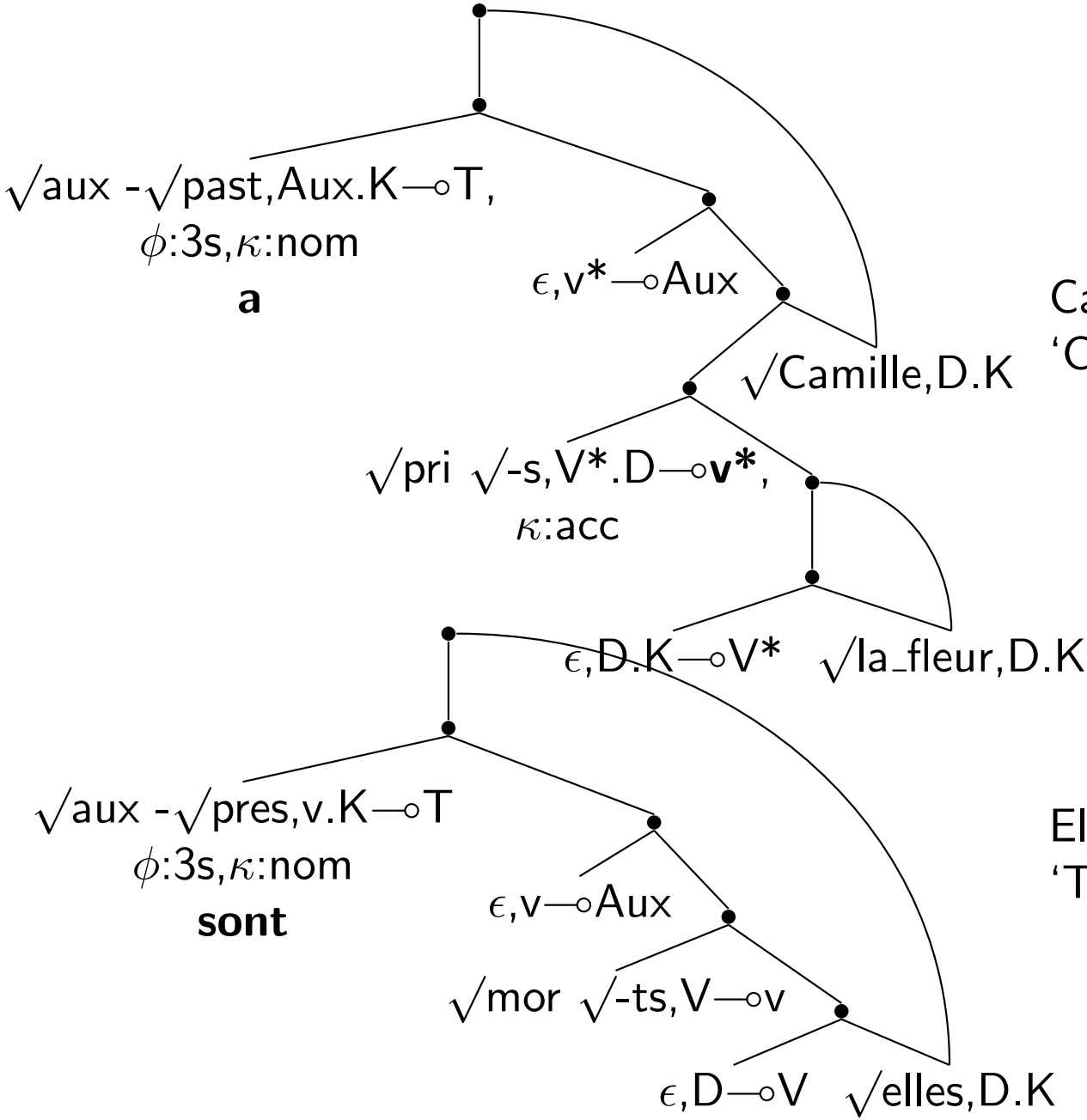
Auxiliaries: English



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

Auxiliaries: French

(Olivier'25, Bjorkman'11)

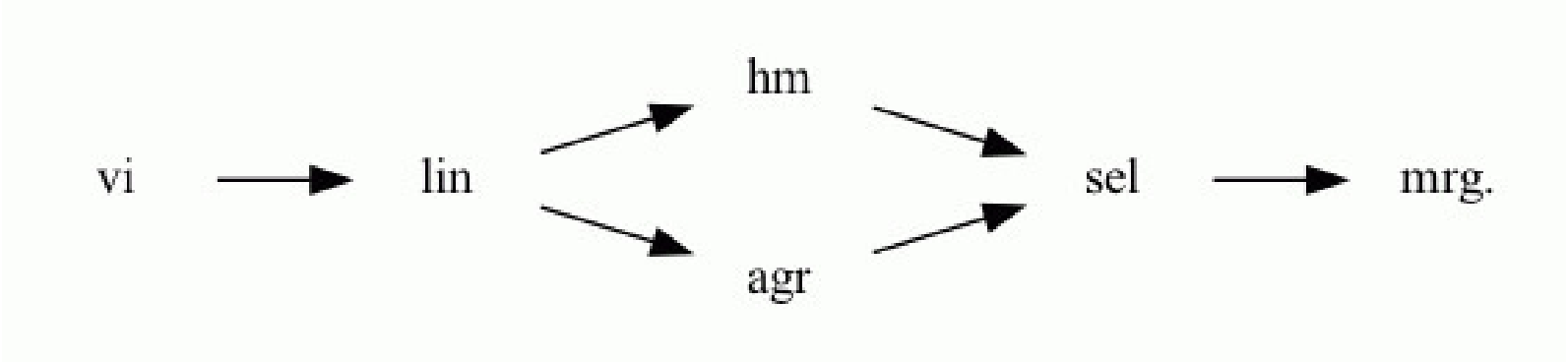


Camille a pri-s la fleur
'Camille took the flower'

Elles sont mor-ts
'They died'

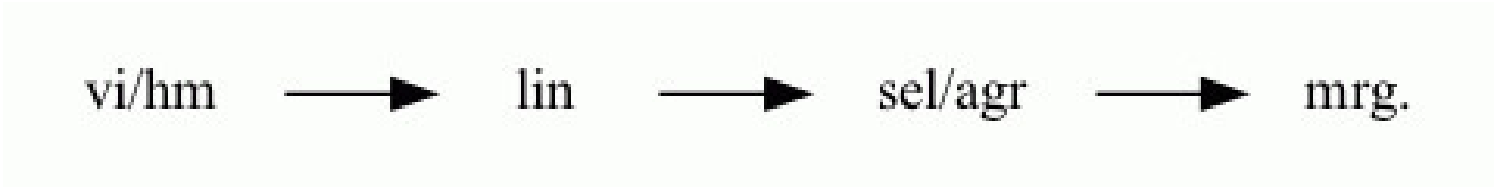
Dependencies among modules

- 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