Grammar Size and Quantitative Restrictions on Movement

Thomas Graf

Stony Brook University
mail@thomasgraf.net
http://thomasgraf.net

SCiL 2018
Jan 4–7, 2018

Outline

Minimalist Grammars (MGs)

- Minimalist grammars (MGs) are a formalization of Chomskyan syntax (Stabler 1997, 2011)
- Succinct formalism for defining MCFGs
- Operations: Merge and **Move**
- Grammar is just a finite list of feature-annotated lexical items (LIs)

Chemistry	Syntax
atoms	words
electrons	features
molecules	sentences

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...

the	men	like	which	men
N^+ D^-	N^-	D^+ D^+ V^-	N^+ D^-	N^-

- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...

the	men	like	which	men
N^+ D^-	N^-	D^+ D^+ V^-	N^+ D^-	N^-

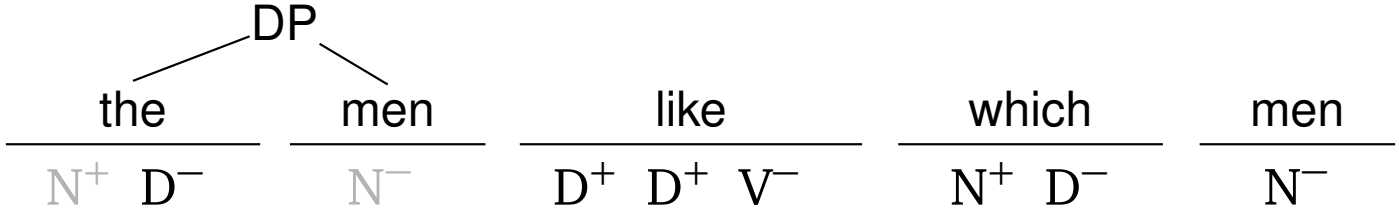
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...



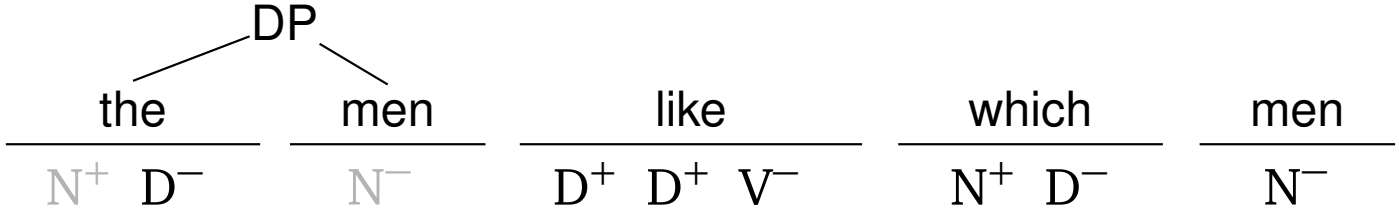
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...



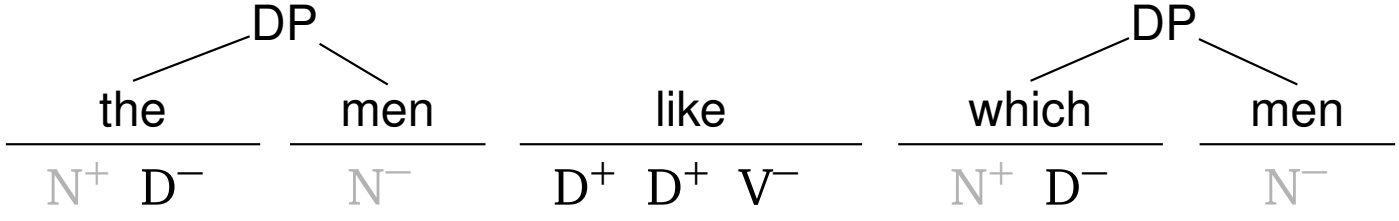
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...



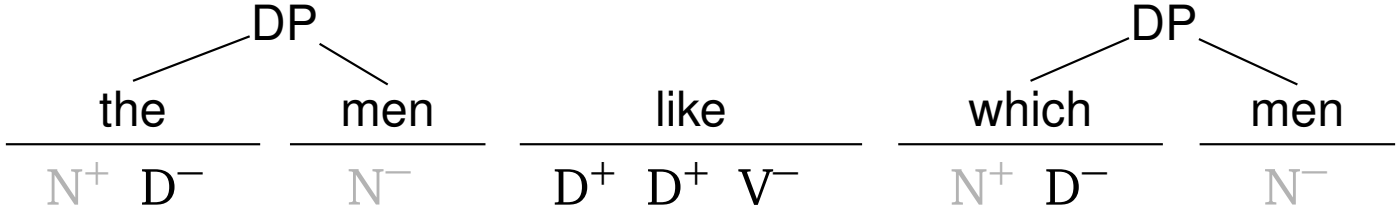
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...



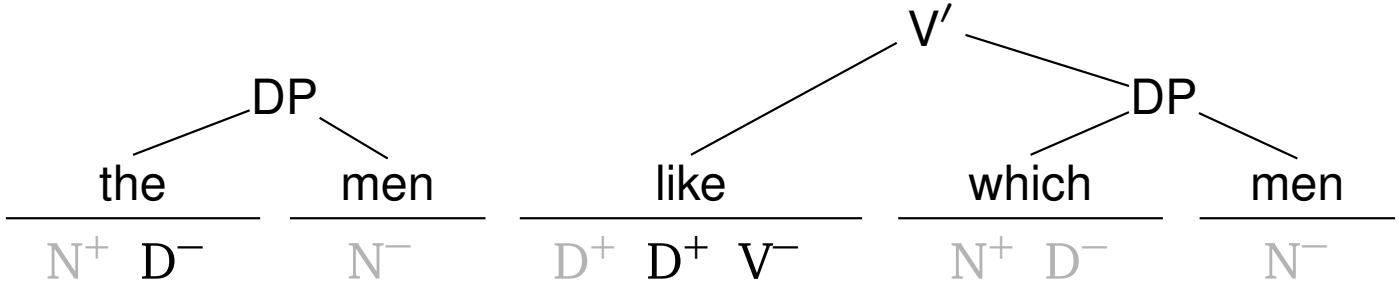
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , . . .

selector feature N^+ , V^+ , . . .



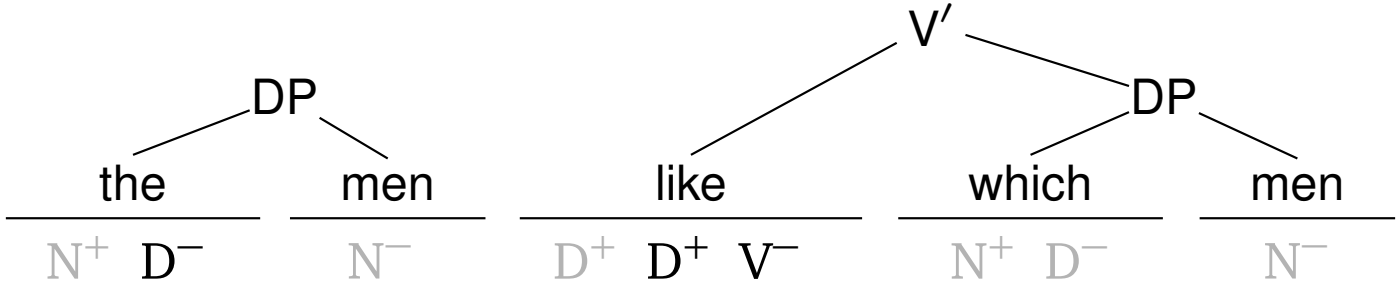
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...



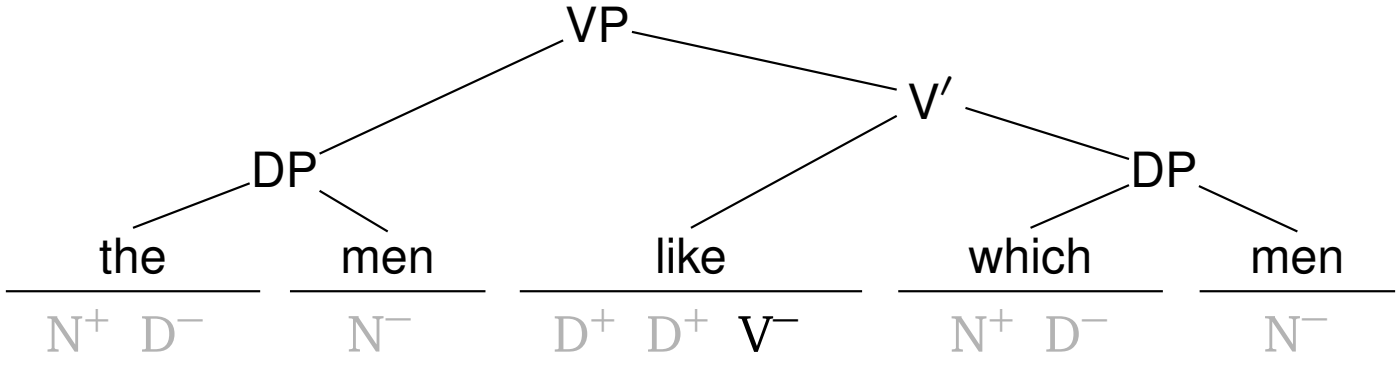
- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge

Merge combines subtrees to encode **head-argument dependencies**.

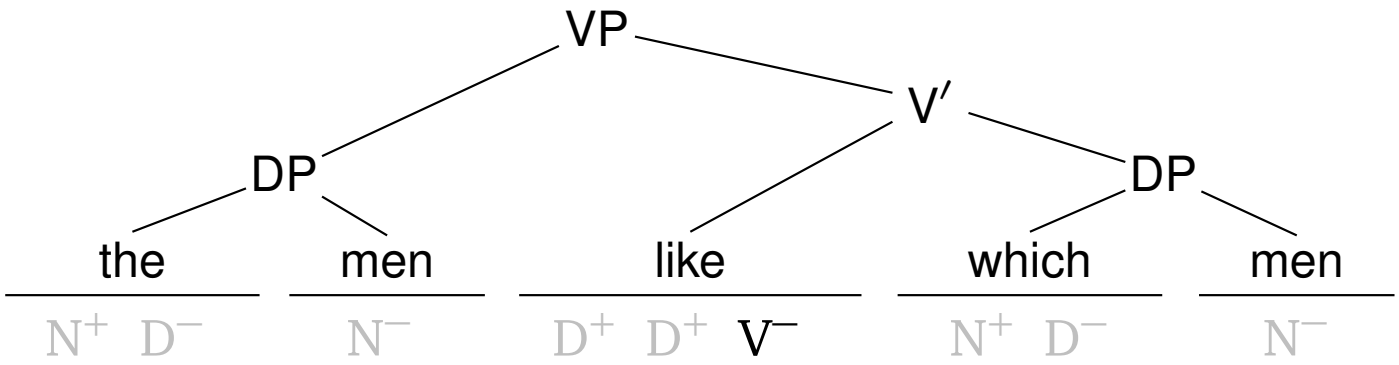
category feature N^- , V^- , ...

selector feature N^+ , V^+ , ...

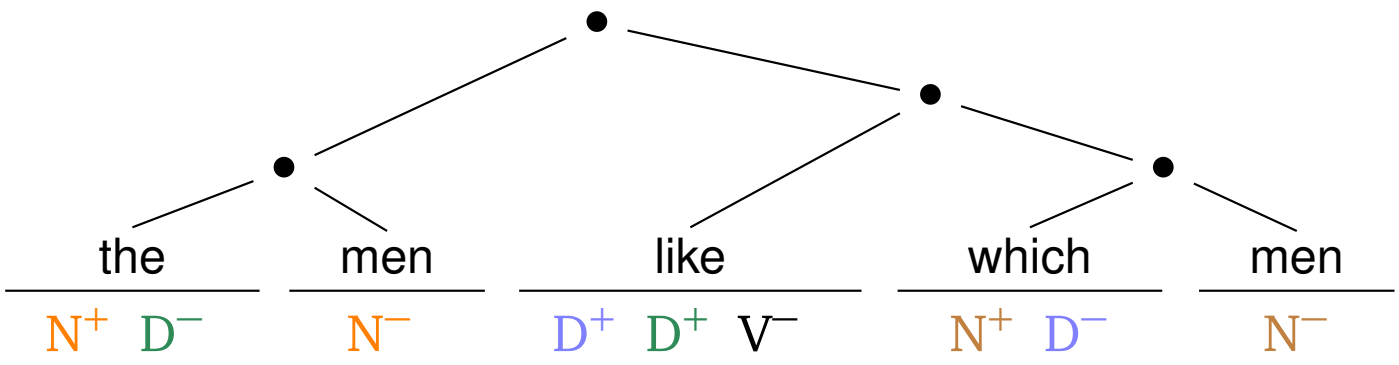


- *the* and *men* have matching features, triggering Merge
- same steps for *which men*
- *like* merged with *which men*
- *like* merged with *the men*

Merge in Derivation Trees



Derived Tree



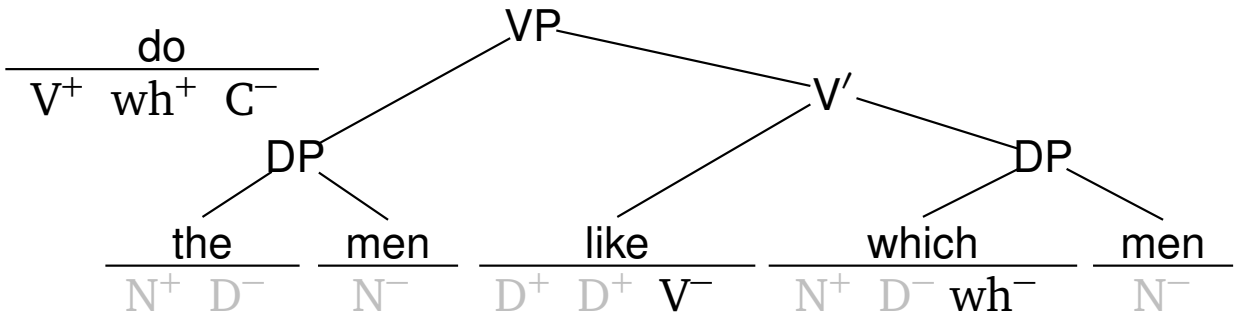
Derivation Tree

Move

Move displaces subtrees to derive the correct **linear order**.

licensee feature wh^- , top^- , ...

licensor feature wh^+ , top^+ , ...



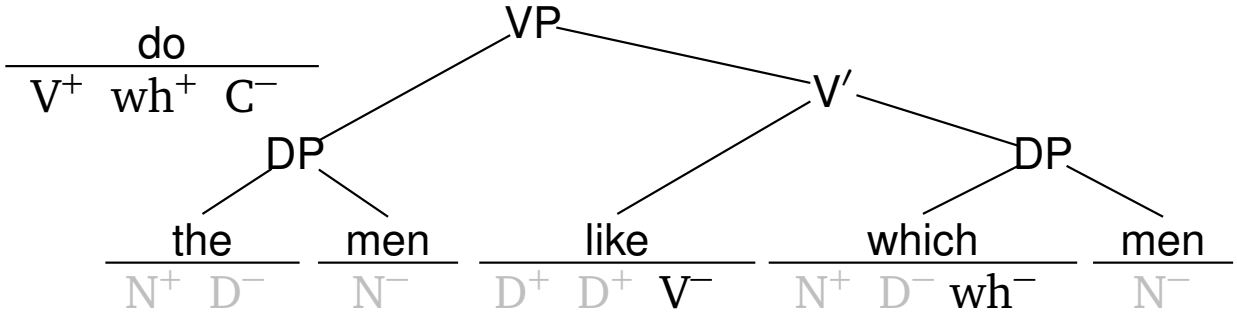
- Merge *do*
- Move triggered by features of opposite polarity

Move

Move displaces subtrees to derive the correct **linear order**.

licensee feature wh^- , top^- , ...

licensor feature wh^+ , top^+ , ...



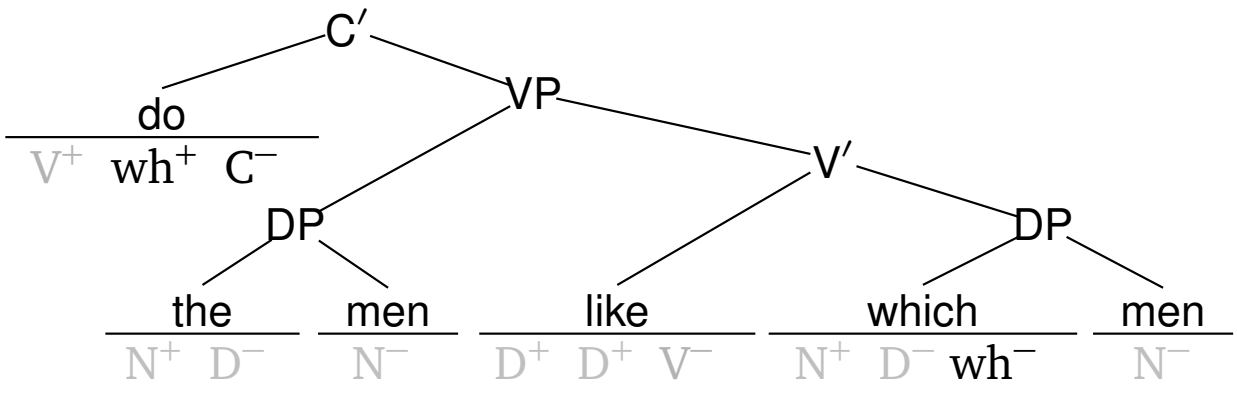
- Merge *do*
- Move triggered by features of opposite polarity

Move

Move displaces subtrees to derive the correct **linear order**.

licensee feature wh^- , top^- , ...

licensor feature wh^+ , top^+ , ...



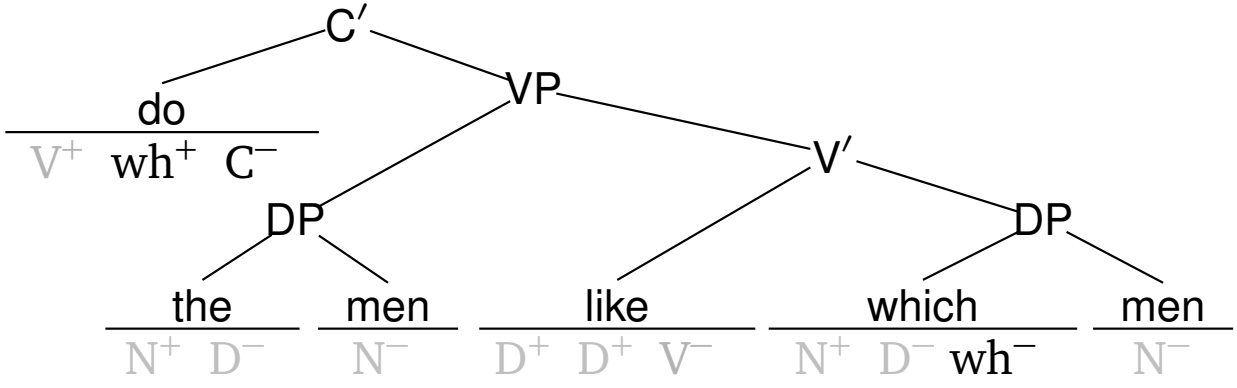
- Merge *do*
- Move triggered by features of opposite polarity

Move

Move displaces subtrees to derive the correct **linear order**.

licensee feature wh^- , top^- , ...

licensor feature wh^+ , top^+ , ...



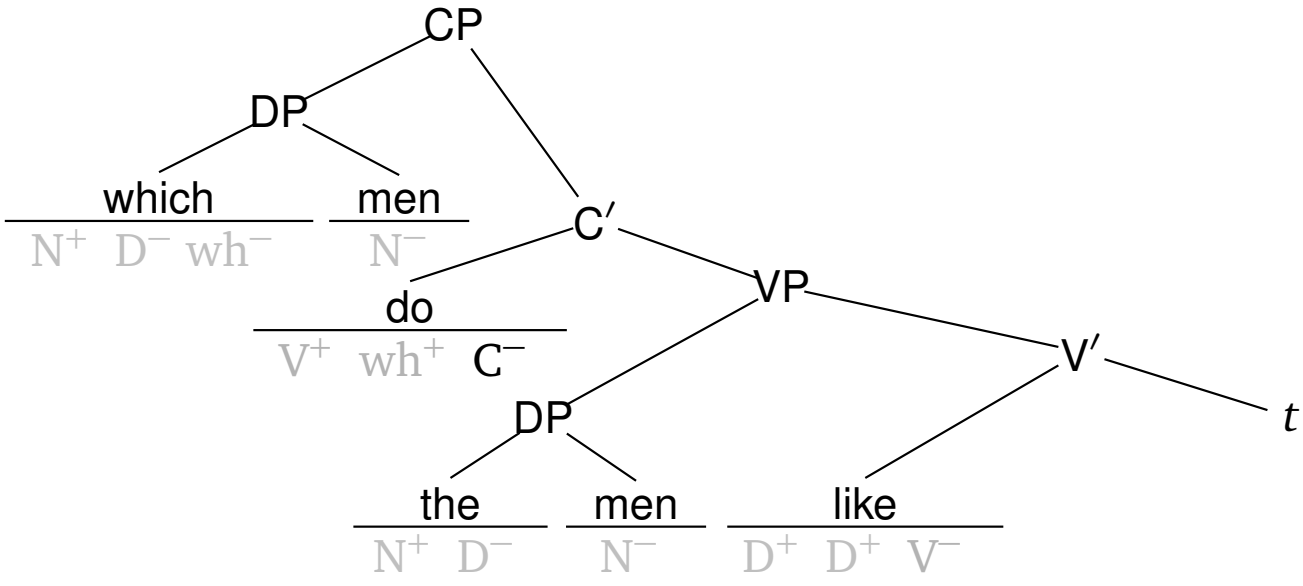
- Merge *do*
- Move triggered by features of opposite polarity

Move

Move displaces subtrees to derive the correct **linear order**.

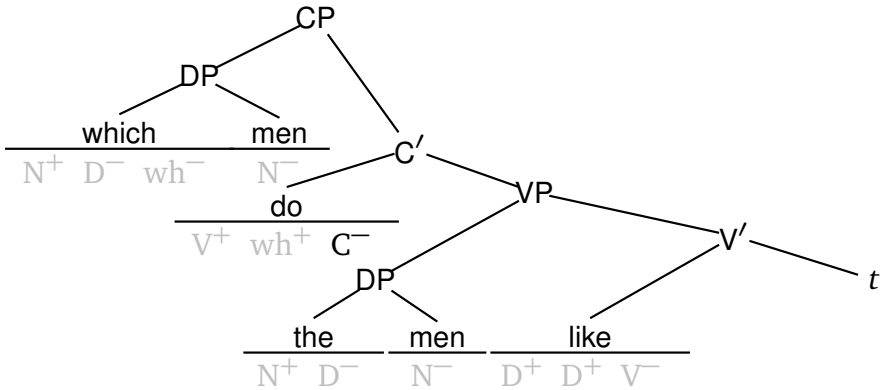
licensee feature wh^- , top^- , ...

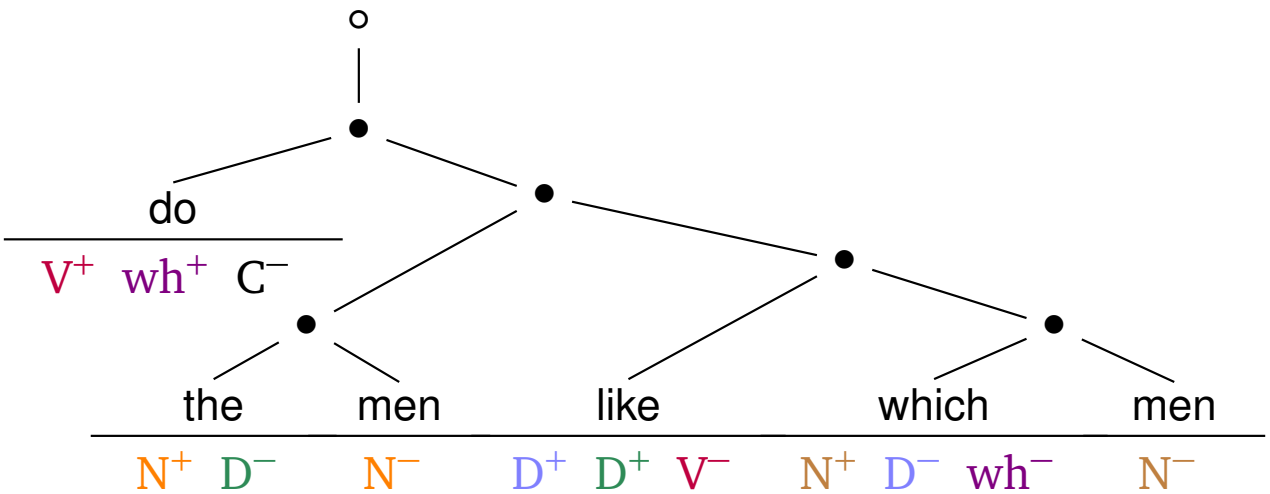
licensor feature wh^+ , top^+ , ...



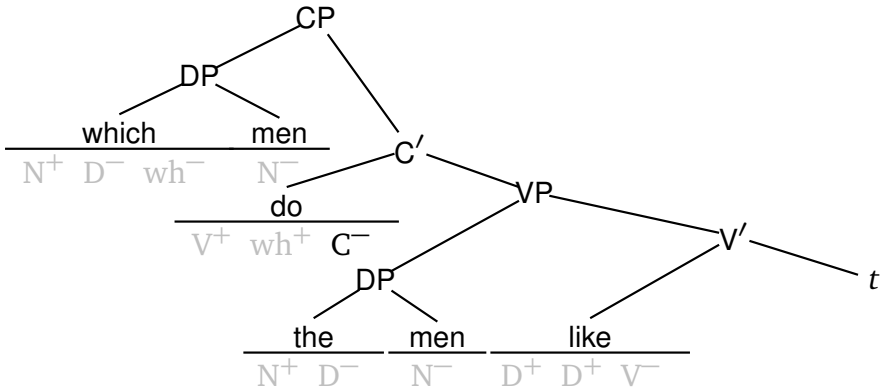
- Merge *do*
- Move triggered by features of opposite polarity

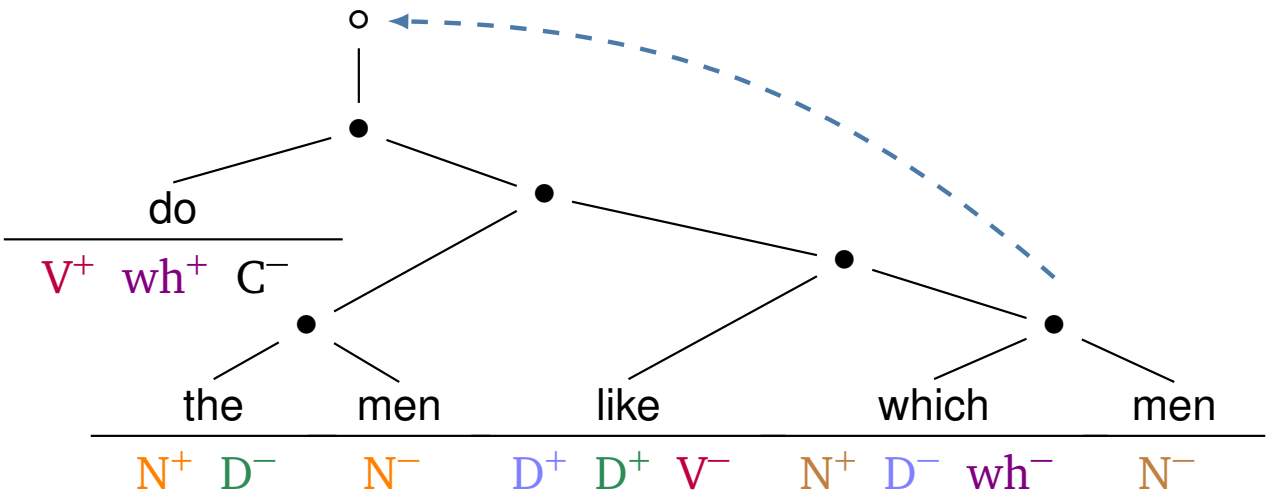
Move in Derivation Trees





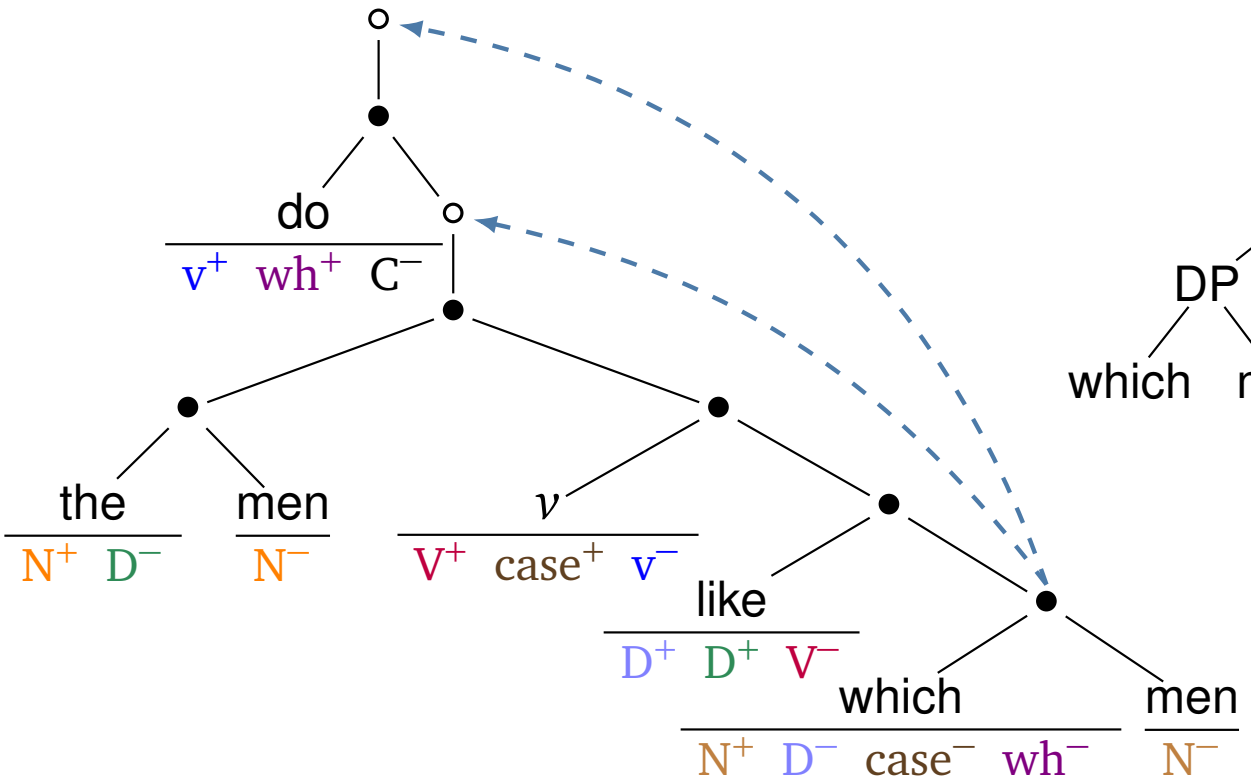
Move in Derivation Trees

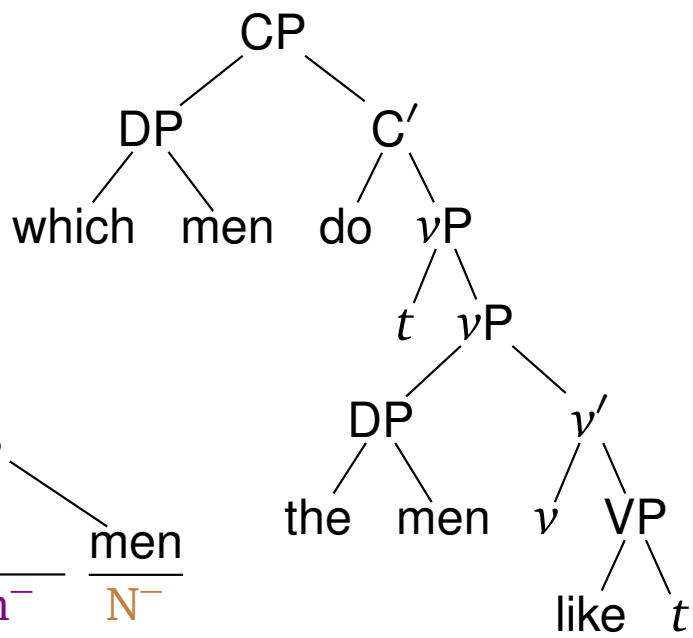




Intermediate Movement

Intermediate movement is possible, but has no effect on output.



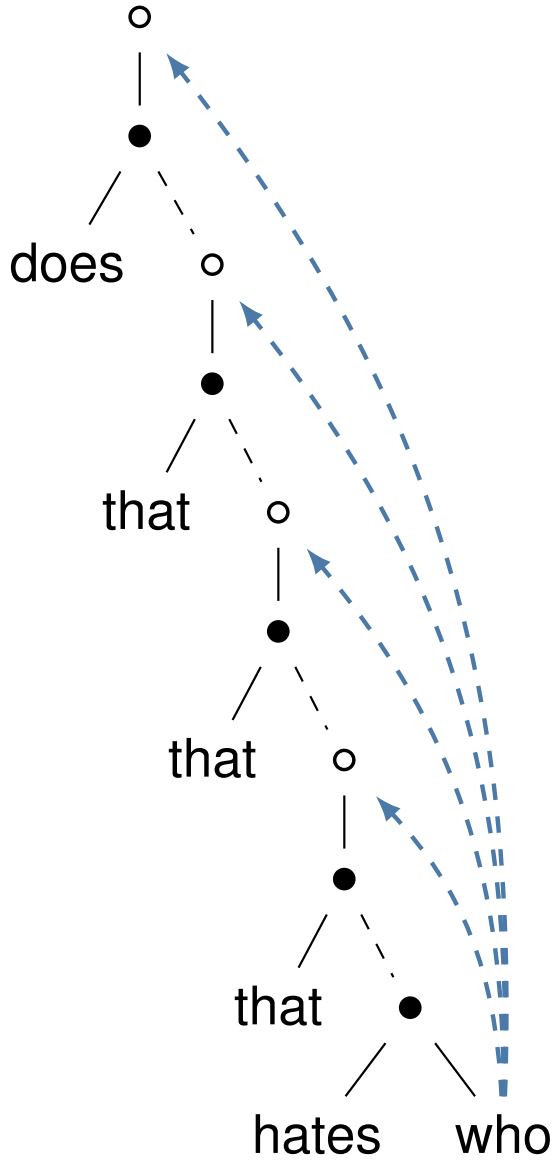


An Issue with Intermediate Movement

Minimalist analyses posit an unbounded number of intermediate landing sites.

(1) Who does John think
t that Mary believes
t that Sue said
t that . . .
t that Bill hates *t*?

But every LI can only carry **finitely many features**!

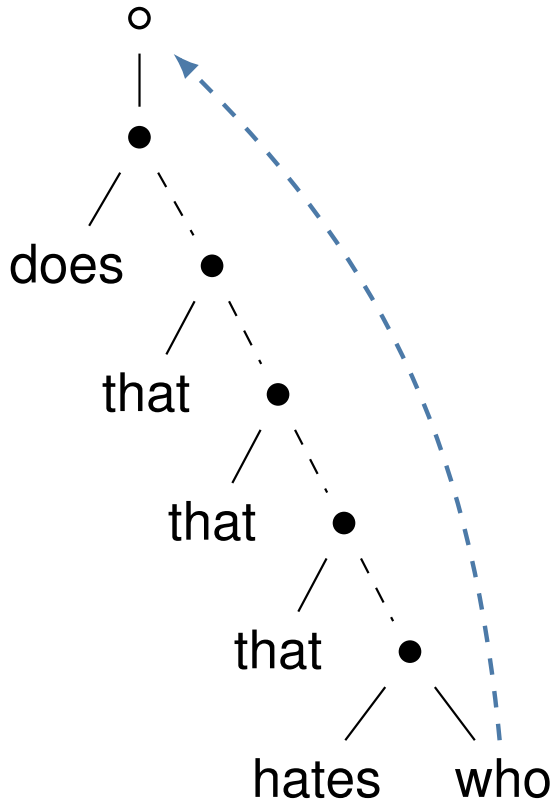


Derivational Solution (Kobele 2006)

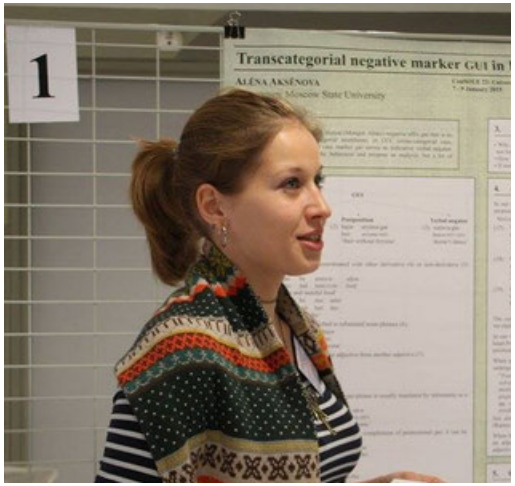
- Only final landing site has feature.
- Intermediate movement is inserted by mapping to phrase structure trees

Recipe for Successive Cyclic Movement

Add trace in every crossed Spec,CP.



Generalization: No Intermediate Movement (Graf et al. 2016)



Alëna Aksënova

Aniello De Santo

- Kobele's solution can be generalized.
- Intermediate movement never needs to be feature triggered.
- It is **derivationally redundant**.

Definition (Single Movement Normal Form)

An MG is in single movement normal form **(SMNF)** iff every LI has at most one licensee feature.

Theorem

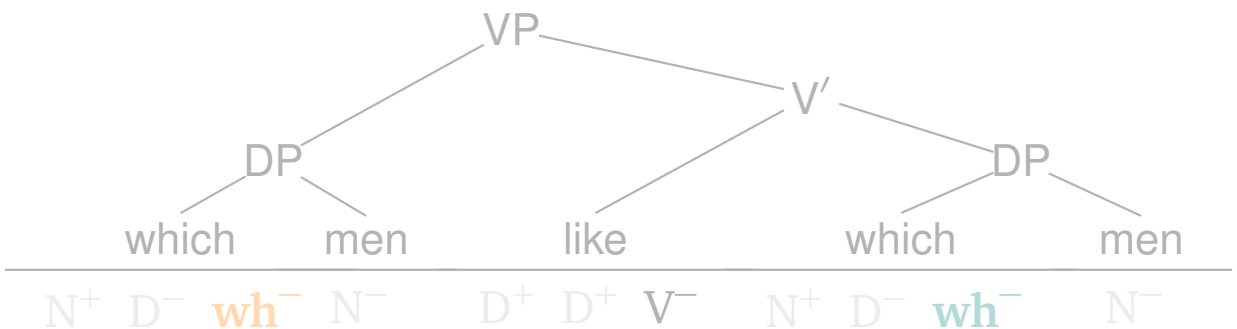
For every MG there is a strongly equivalent MG that is in SMNF.

An Annoying Complication

MGs have one central locality restriction on Move.

Shortest Move Constraint (SMC)

If two LIs in a tree both have a licensee feature as their first currently unchecked feature, then these features must be distinct.

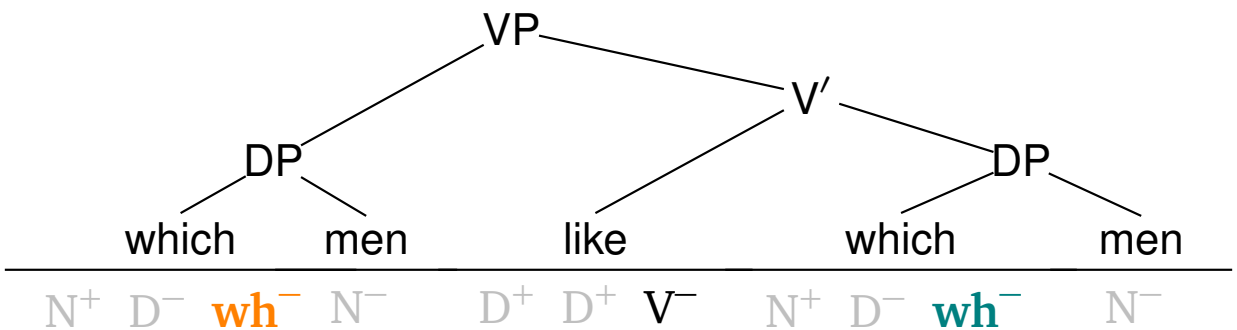


An Annoying Complication

MGs have one central locality restriction on Move.

Shortest Move Constraint (SMC)

If two LIs in a tree both have a licensee feature as their first currently unchecked feature, then these features must be distinct.

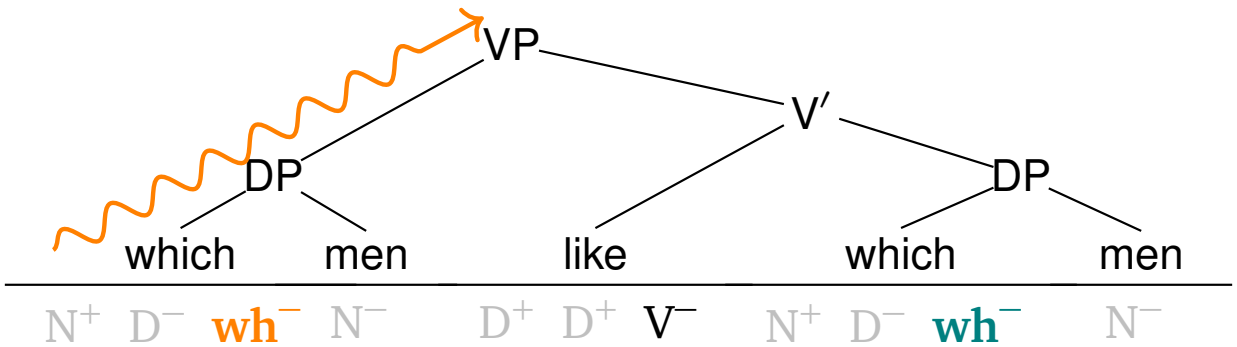


An Annoying Complication

MGs have one central locality restriction on Move.

Shortest Move Constraint (SMC)

If two LIs in a tree both have a licensee feature as their first currently unchecked feature, then these features must be distinct.

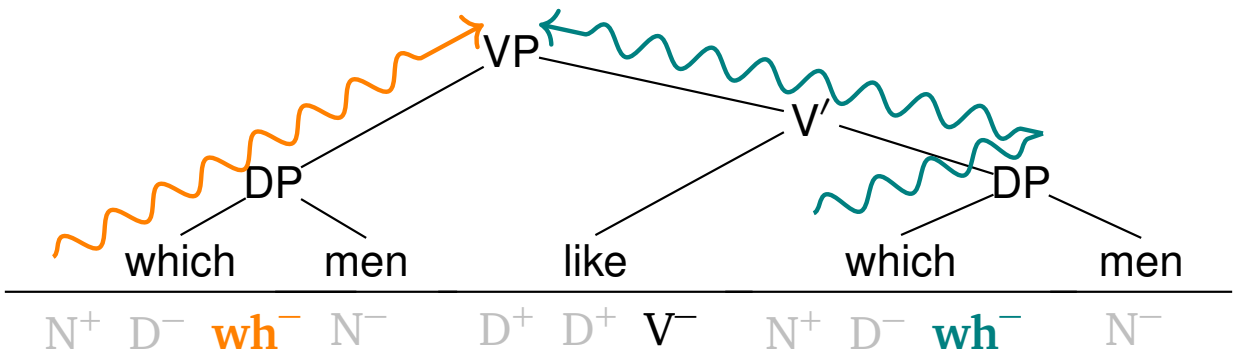


An Annoying Complication

MGs have one central locality restriction on Move.

Shortest Move Constraint (SMC)

If two LIs in a tree both have a licensee feature as their first currently unchecked feature, then these features must be distinct.

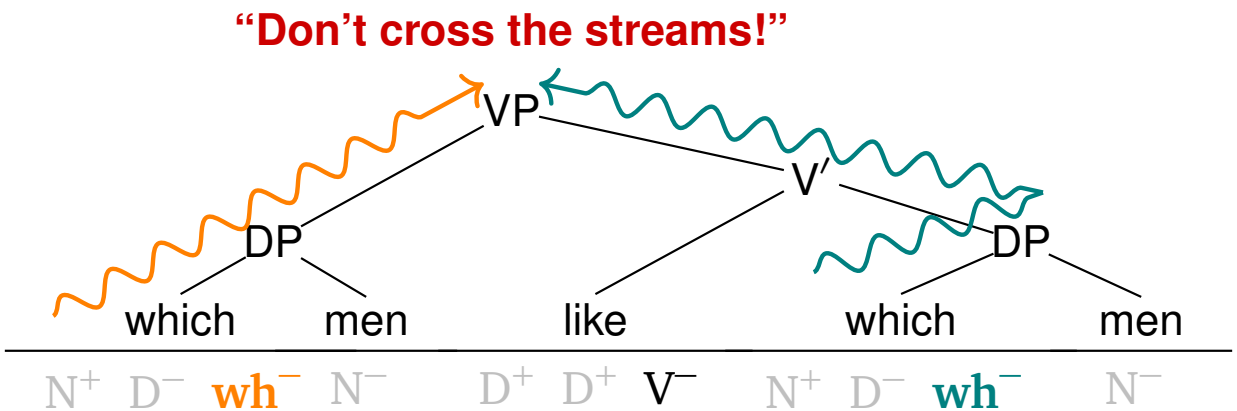


An Annoying Complication

MGs have one central locality restriction on Move.

Shortest Move Constraint (SMC)

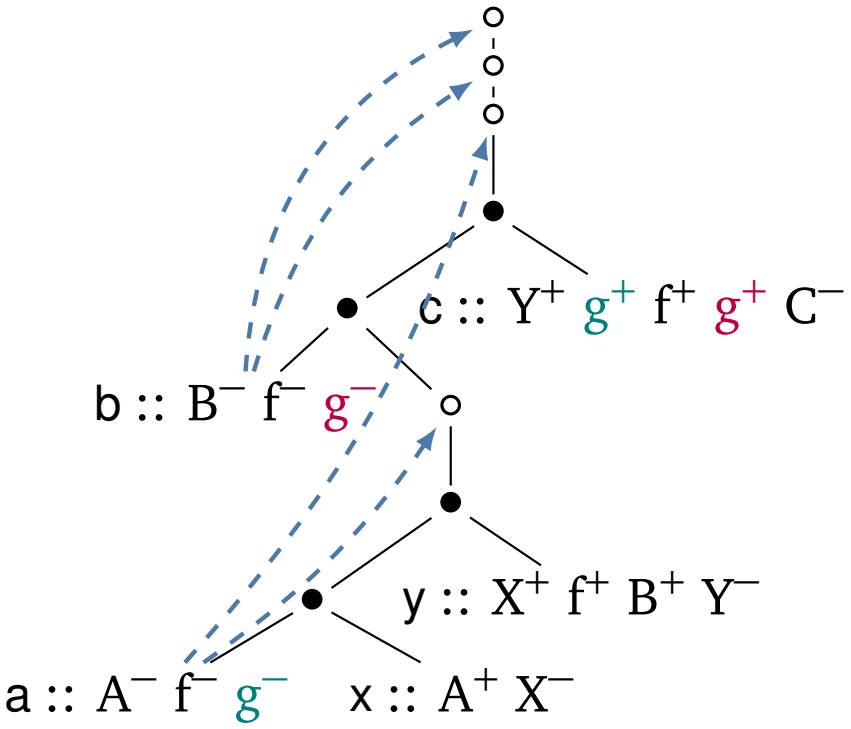
If two LIs in a tree both have a licensee feature as their first currently unchecked feature, then these features must be distinct.



A Work-Around with Subscripts

Feature Subscripting

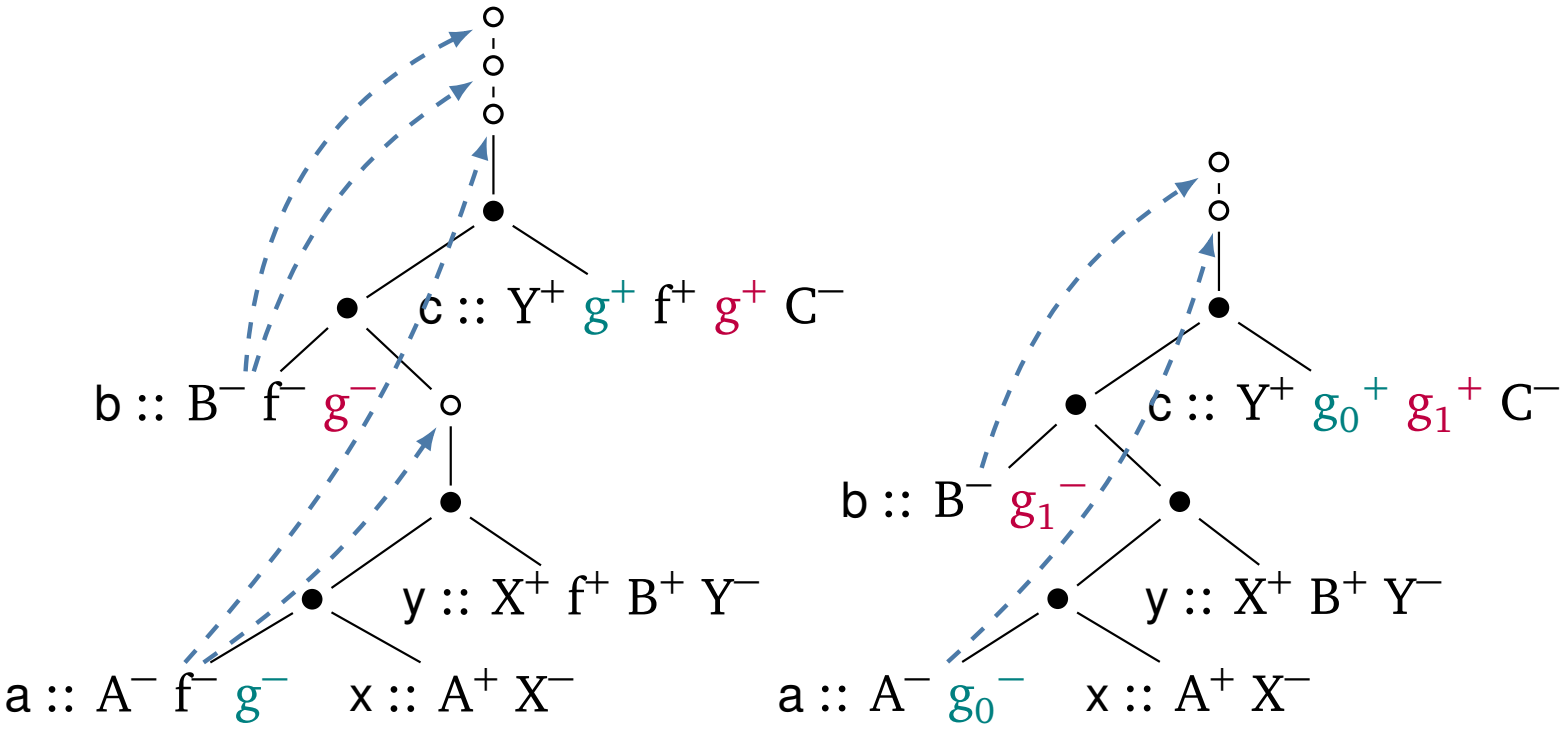
- For every LI l, only keep its last licensee feature.
- Add subscripts to licensee features to avoid SMC violations.



A Work-Around with Subscripts

Feature Subscripting

- For every LI l, only keep its last licensee feature.
- Add subscripts to licensee features to avoid SMC violations.



Lexical Blow-Up

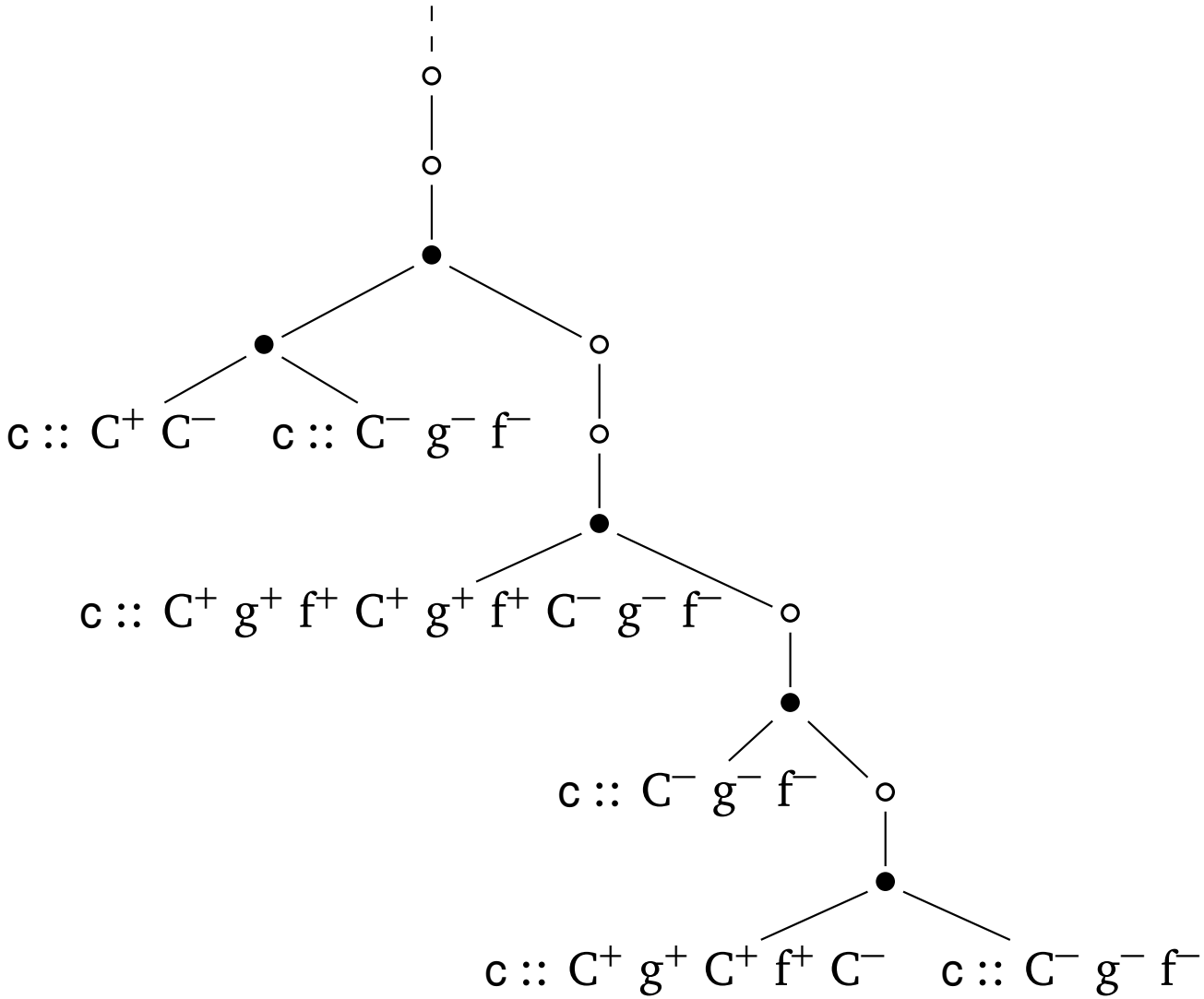
- SMNF translation induces **linear lexical blow-up**
- Effect varies a lot depending on movement configurations:

lower bound	linear size reduction(!),
	1:1 for non-redundant grammars
upper bound	large linear blow-up

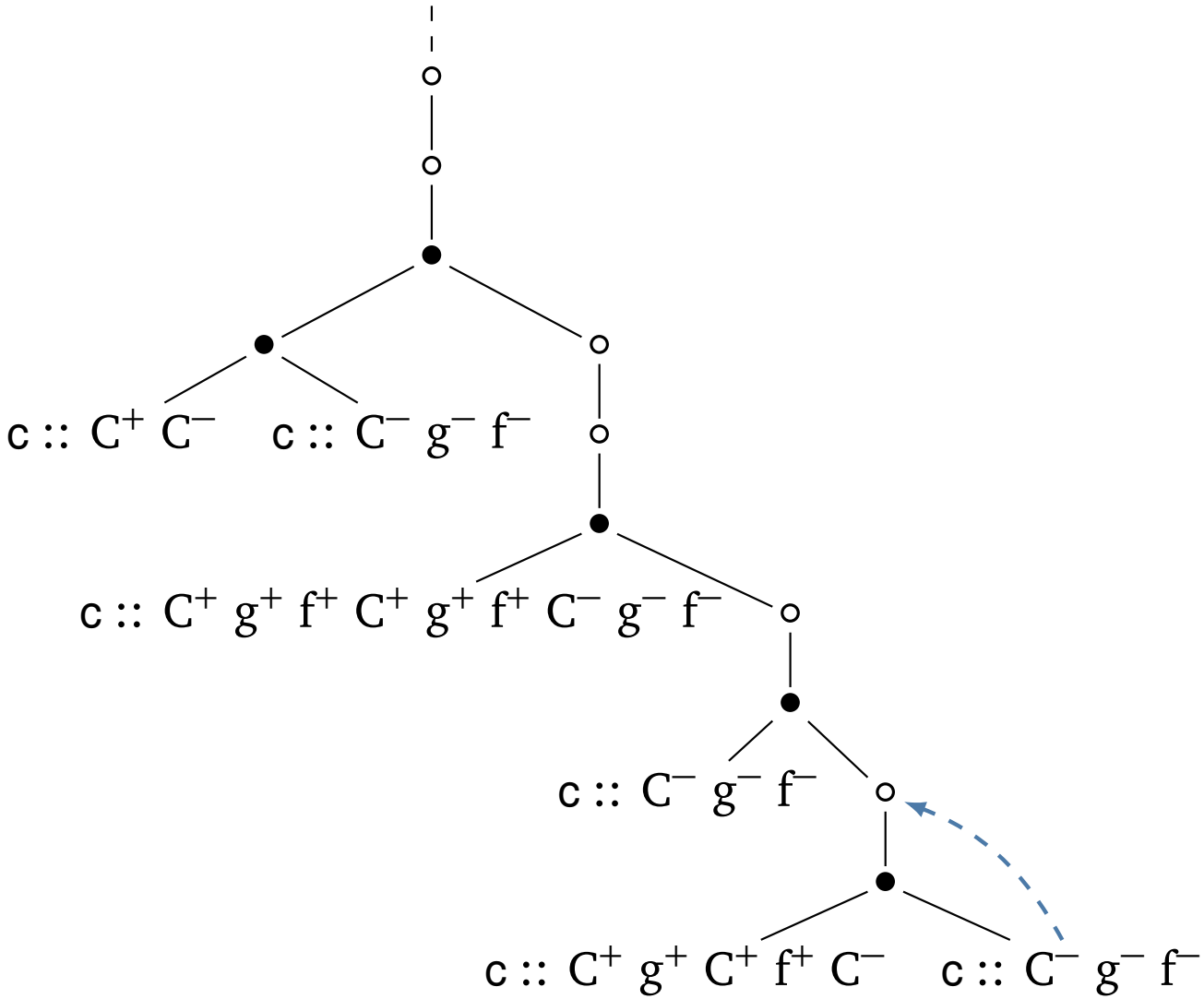
$$\sum_{l \in \mathbf{Lex}} \mu^{\gamma(l)+\delta(l)}$$

μ ... maximum number of required indices
$\gamma(l)$ ... number of licensor features of LI l in original grammar
$\delta(l)$ ... 1 if l has licensee features, 0 otherwise

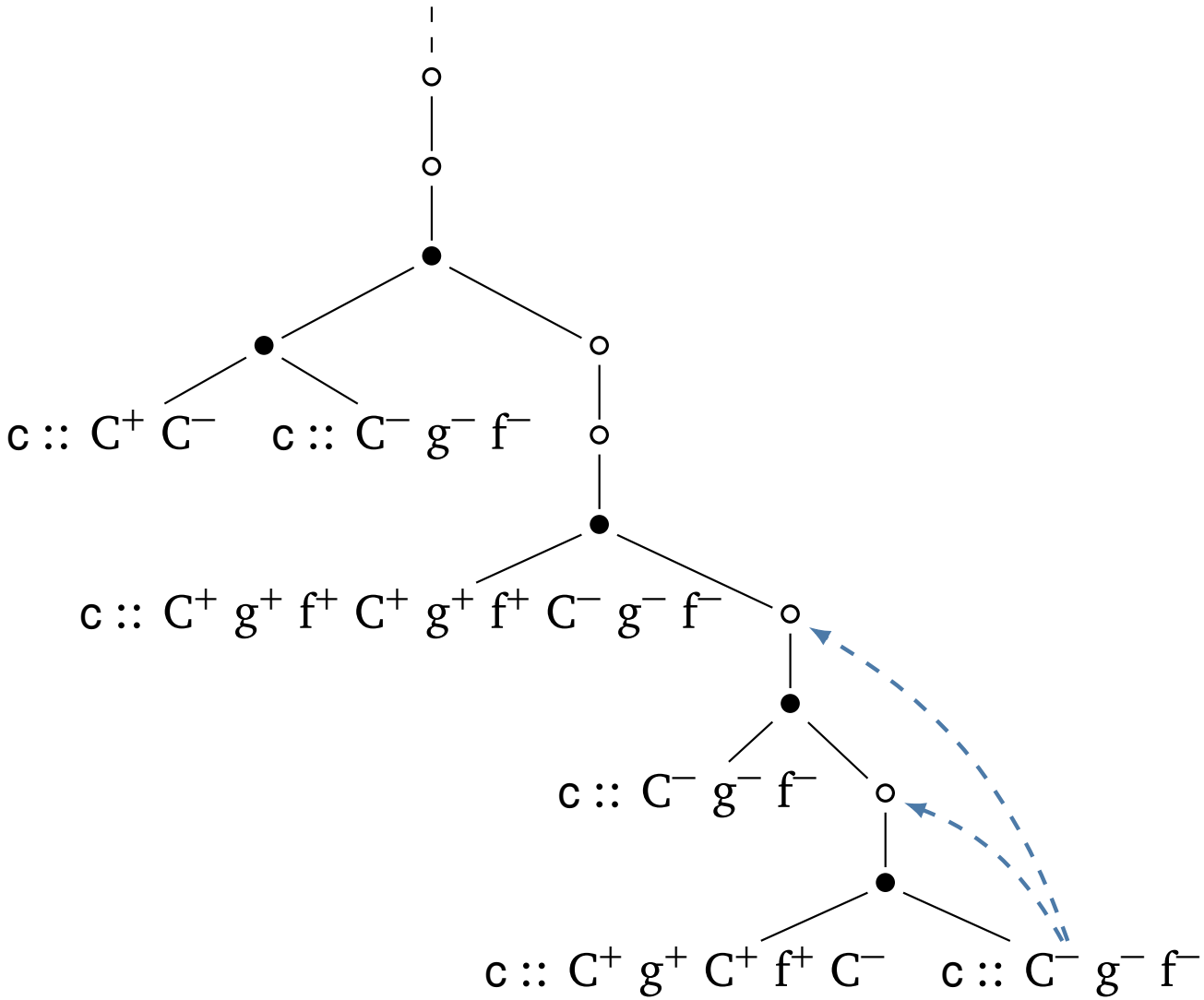
Abstract Example (Sketch)



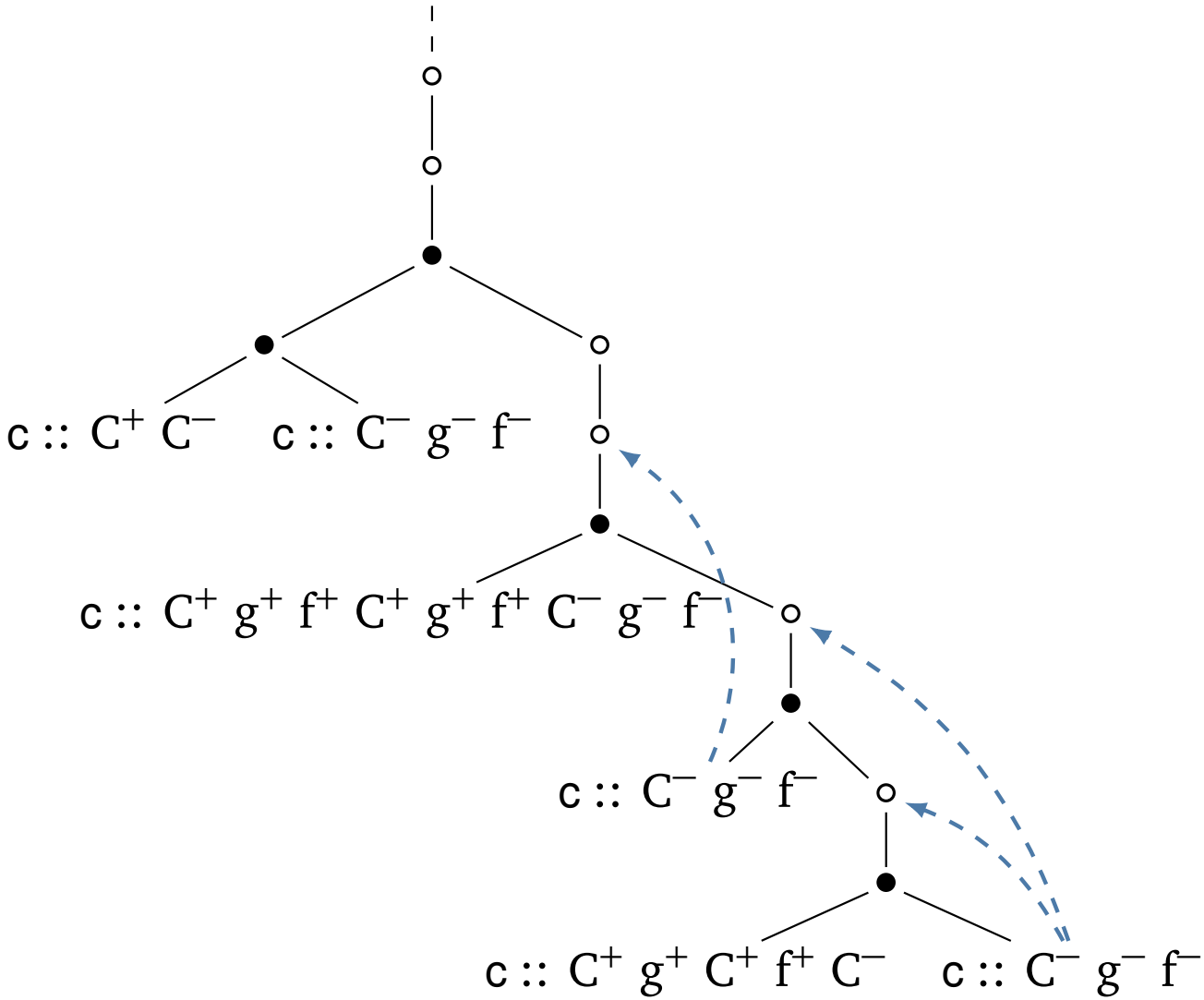
Abstract Example (Sketch)



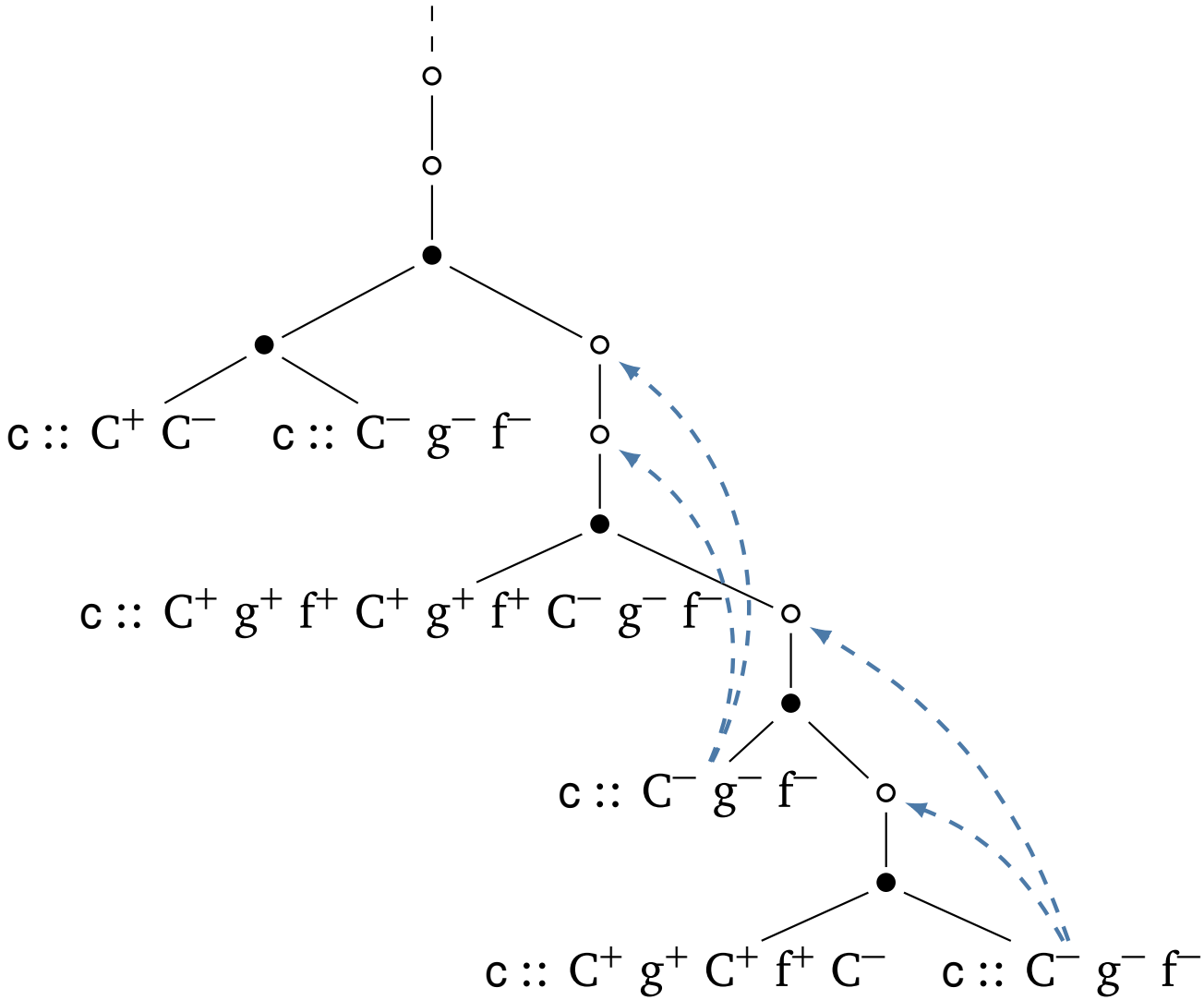
Abstract Example (Sketch)



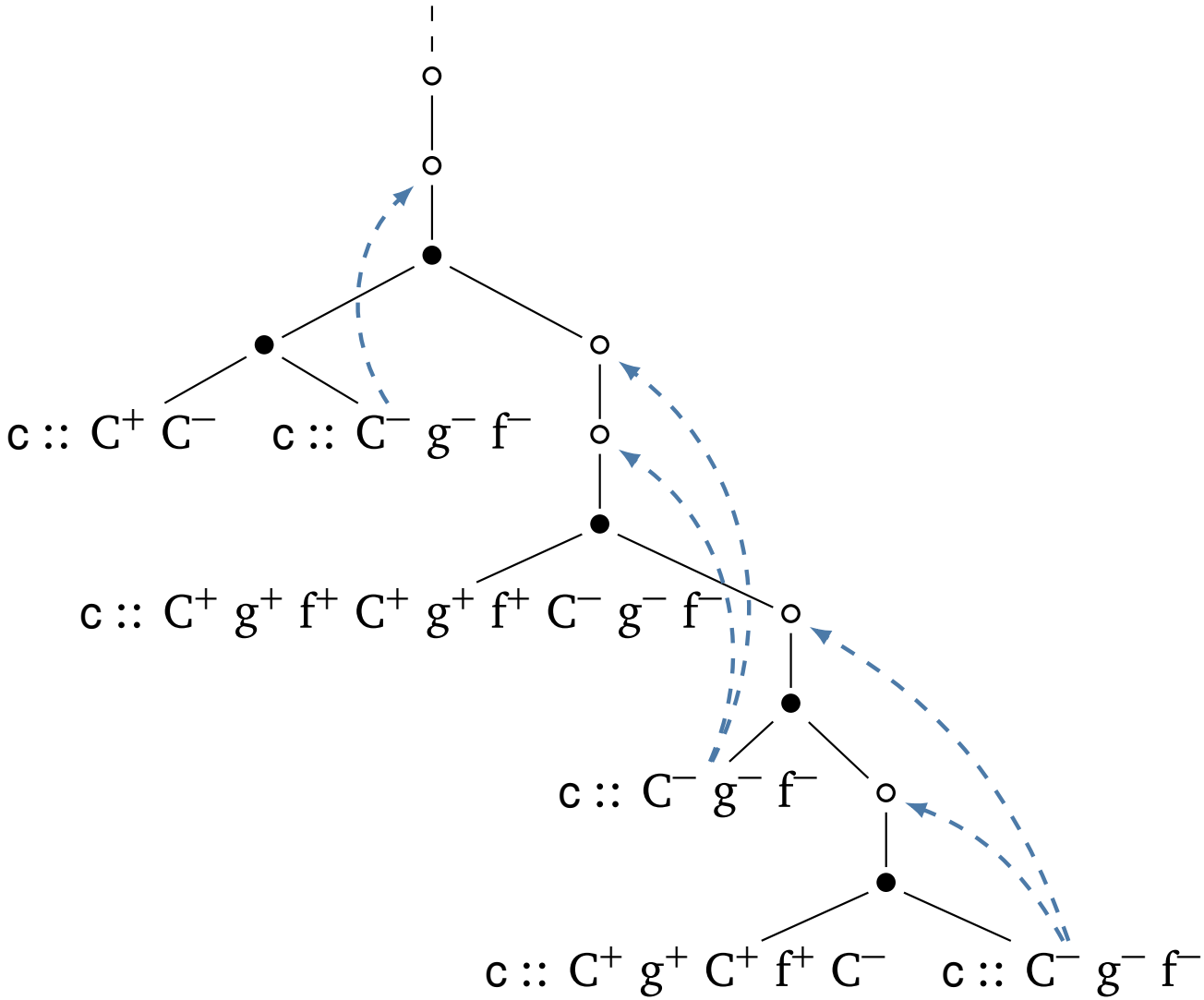
Abstract Example (Sketch)



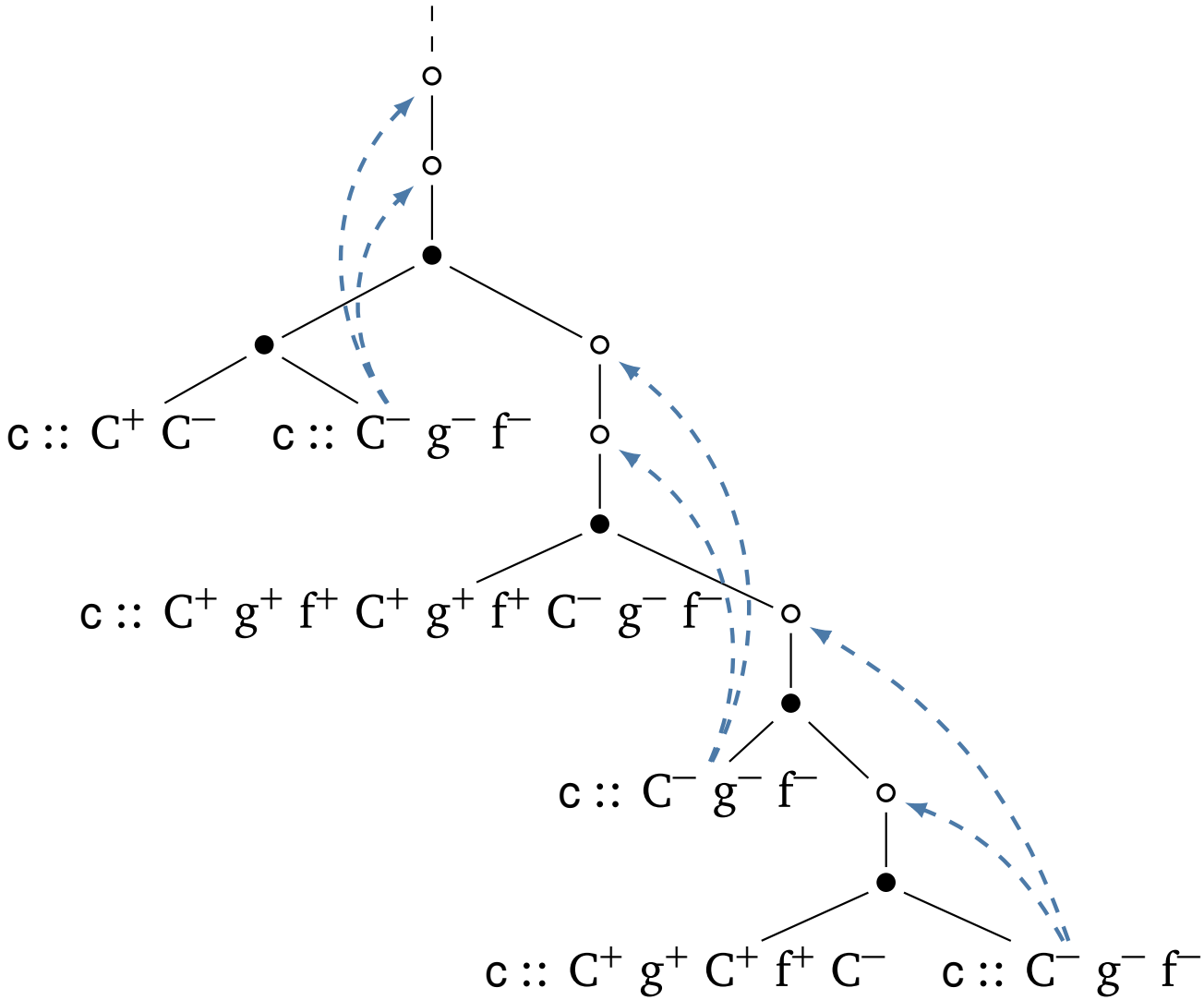
Abstract Example (Sketch)



Abstract Example (Sketch)



Abstract Example (Sketch)



Abstract Example (the Math)

- The original grammar contains a single LI.

$$c :: C^{+}\ g^{+}\ f^{+}\ C^{+}\ g^{+}\ f^{+}\ C^{-}\ g^{-}\ f^{-}$$

- The SMNF grammar contains **8** variants.

$$c :: C^{+}\ f_0^{+}\ C^{+}\ f_0^{+}\ C^{-}\ f_0^{-} \qquad c :: C^{+}\ f_0^{+}\ C^{+}\ f_0^{+}\ C^{-}\ f_1^{-}$$
$$c :: C^{+}\ f_0^{+}\ C^{+}\ f_1^{+}\ C^{-}\ f_0^{-} \qquad c :: C^{+}\ f_0^{+}\ C^{+}\ f_1^{+}\ C^{-}\ f_1^{-}$$
$$c :: C^{+}\ f_1^{+}\ C^{+}\ f_0^{+}\ C^{-}\ f_0^{-} \qquad c :: C^{+}\ f_1^{+}\ C^{+}\ f_0^{+}\ C^{-}\ f_1^{-}$$
$$c :: C^{+}\ f_1^{+}\ C^{+}\ f_1^{+}\ C^{-}\ f_0^{-} \qquad c :: C^{+}\ f_1^{+}\ C^{+}\ f_1^{+}\ C^{-}\ f_1^{-}$$

- We can get **n** variants of the LI by changing the phonetic exponent, so the grammar size increases at least by **8n**.
- But we can keep increasing number **m** of arguments:

$$2^{m+1} \times n$$

Interim Summary

- Every MG is a finite set of LIs.
- The more LIs, the larger the grammar.
- Derivation trees are the primary data structure.
- Intermediate movement is derivationally redundant and costly:
  - complicates proofs
  - increases computational complexity (Graf and Heinz 2015)
  - at odds with MG processing models (Graf et al. 2017)
- But SMNF MGs may be much larger, which is bad for
  - parsing
  - learning
  - explanatory adequacy

A New Empirical Puzzle

Are the movement configurations we find in natural language exactly those that **induce little lexical blow-up**?

Interim Summary

- Every MG is a finite set of LIs.
- The more LIs, the larger the grammar.
- Derivation trees are the primary data structure.
- Intermediate movement is derivationally redundant and costly:
  - complicates proofs
  - increases computational complexity (Graf and Heinz 2015)
  - at odds with MG processing models (Graf et al. 2017)
- But SMNF MGs may be much larger, which is bad for
  - parsing
  - learning
  - explanatory adequacy

A New Empirical Puzzle

Are the movement configurations we find in natural language exactly those that **induce little lexical blow-up**?

What Produces a Large Blow-Up in Grammar Size?

- Large blow-up occurs whenever there are multiple LIs s.t.
  1. they all have the same final movement feature, and
  2. they have overlapping movement paths, and
  3. their relative configuration is not fixed across derivations.
- That's easy to do with abstract examples,
  but **natural examples are tough**.

The Constraint-Grammar-Size Conspiracy

Patterns that would induce a large blow-up are independently forbidden.

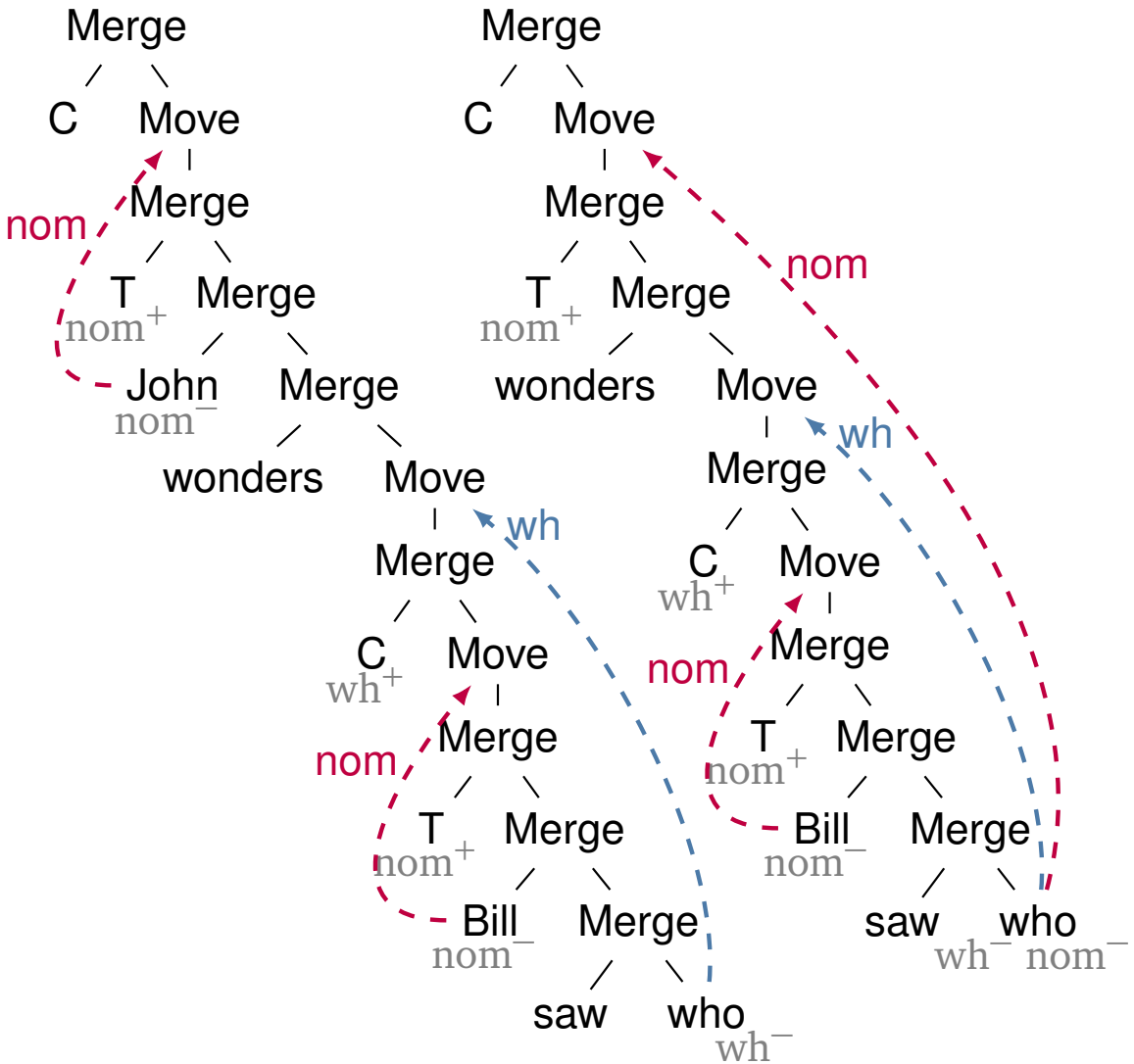
What Produces a Large Blow-Up in Grammar Size?

- Large blow-up occurs whenever there are multiple LIs s.t.
  1. they all have the same final movement feature, and
  2. they have overlapping movement paths, and
  3. their relative configuration is not fixed across derivations.
- That's easy to do with abstract examples,
  but **natural examples are tough**.

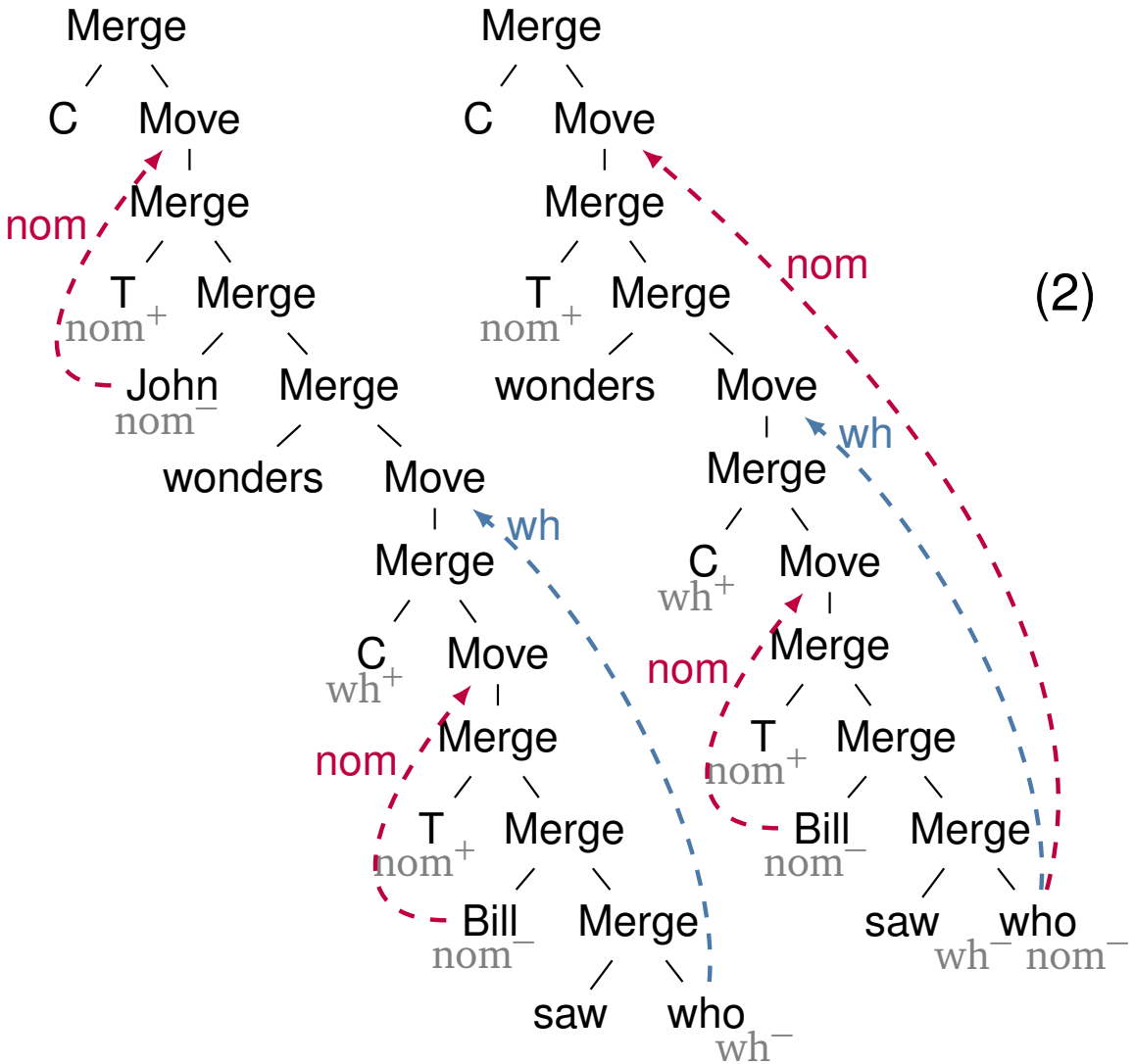
The Constraint-Grammar-Size Conspiracy

Patterns that would induce a large blow-up are independently forbidden.

Improper Movement

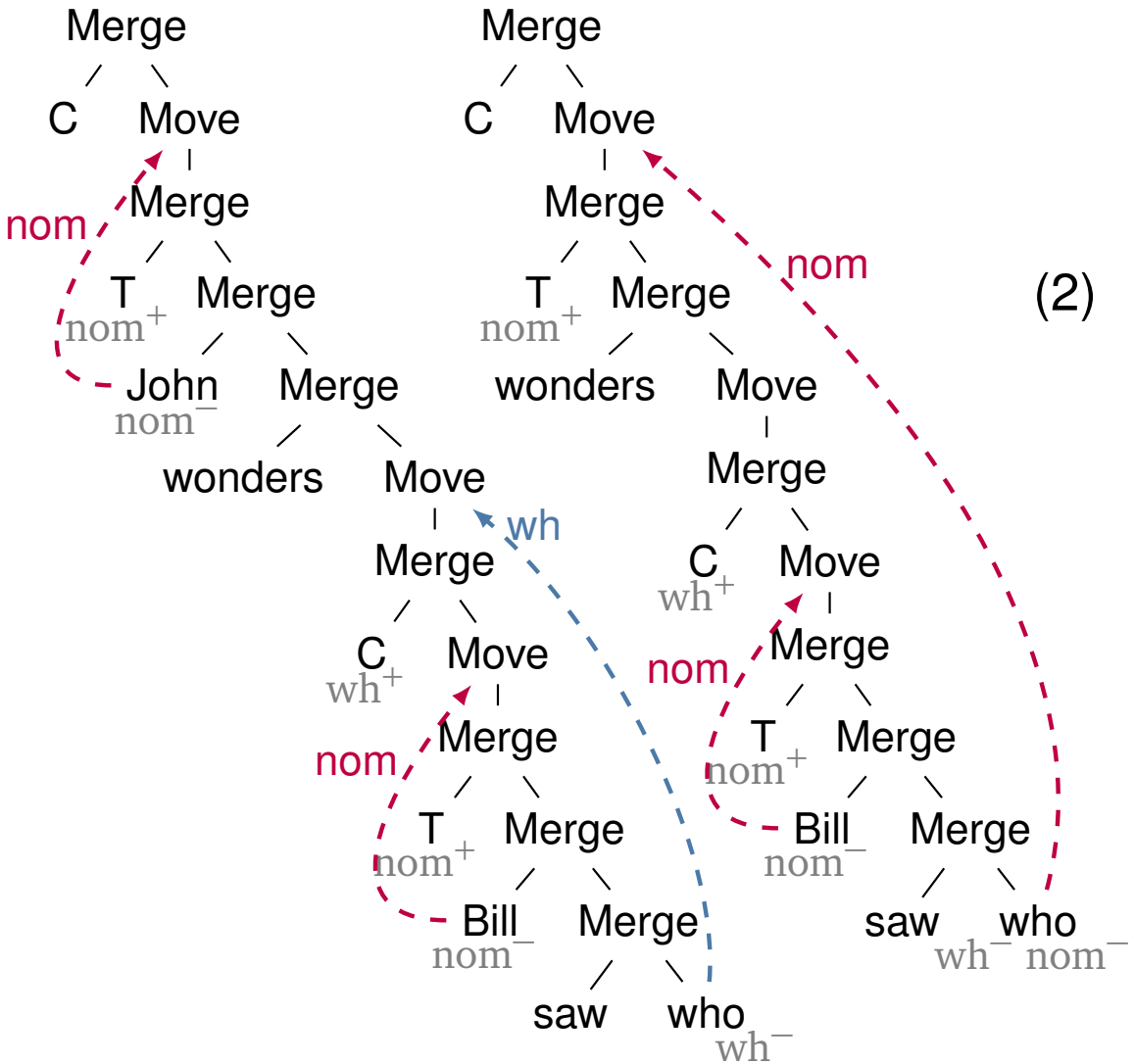


Improper Movement



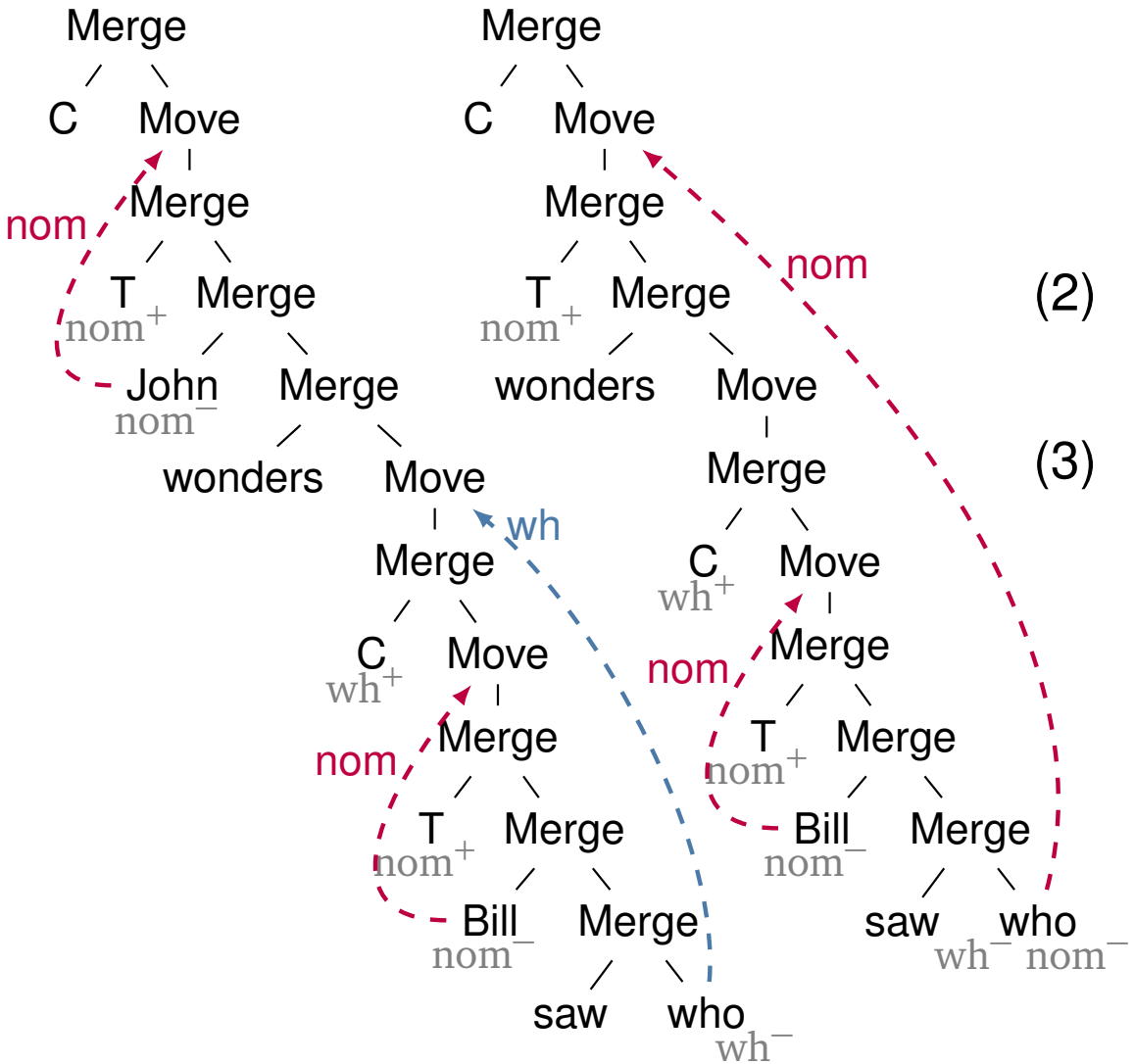
(2) John wonders who Bill saw.

Improper Movement



(2) John wonders who Bill saw.

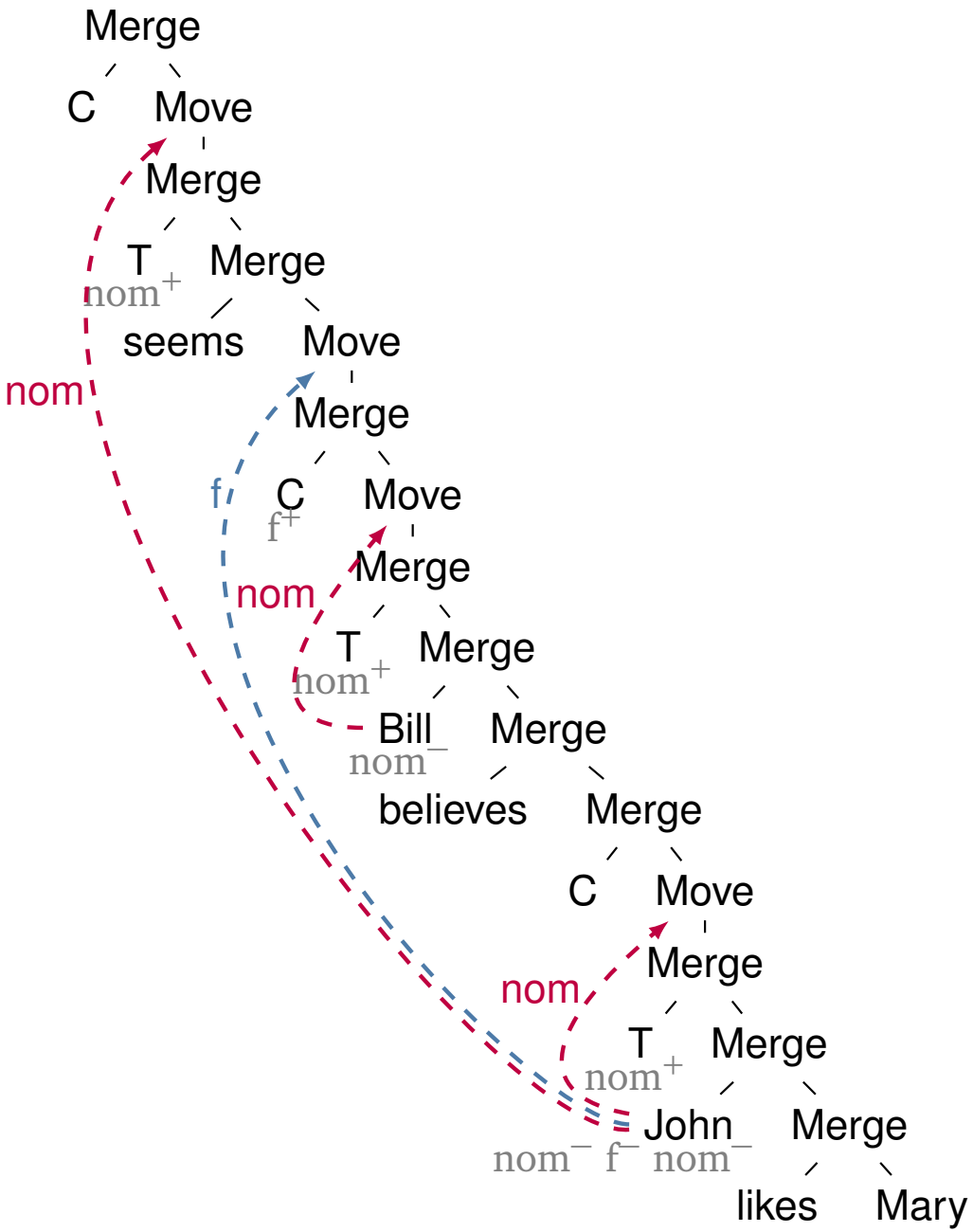
Improper Movement



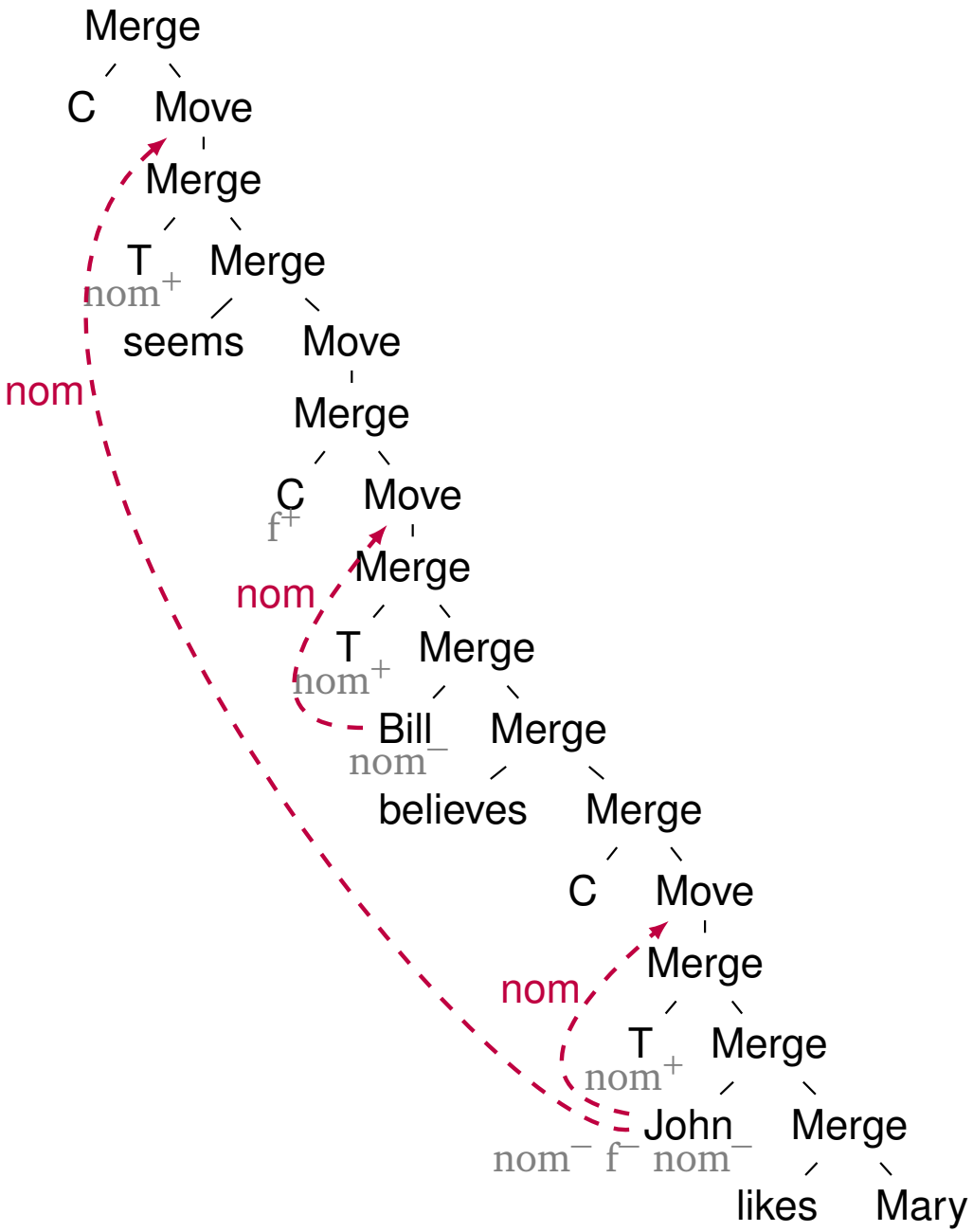
(2) John wonders who Bill saw.

(3) * Who wonders Bill saw?

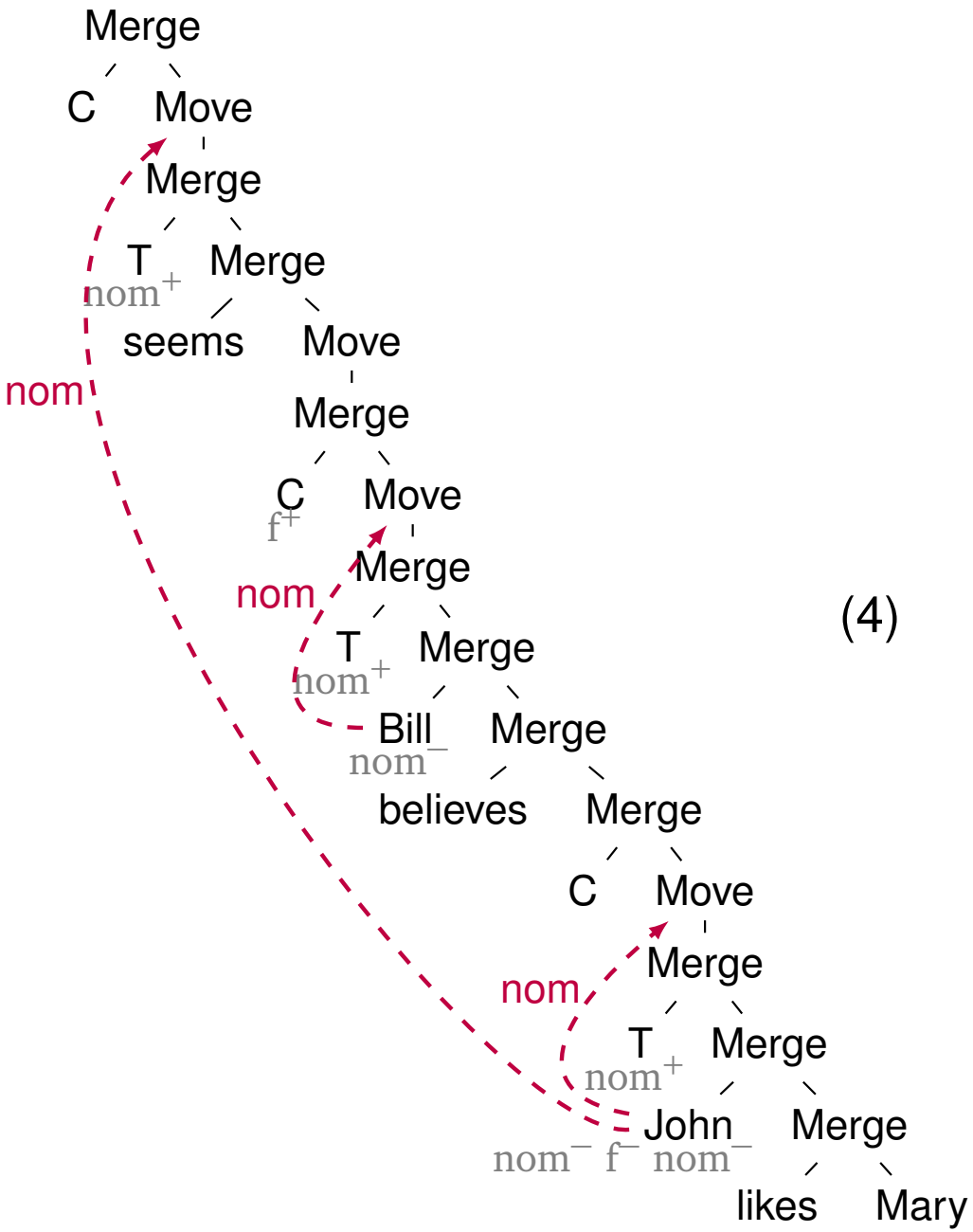
Superraising



Superraising



Superraising



(4) * John seems Bill believes likes Mary

Freezing Effects

- One could also get overlapping paths by extracting an f-mover from within an f-mover.
- But this produces freezing effects.

(5) It seems your comment about John annoys Sue.

(6) * John seems your comment about t annoys Sue.

(7) * Who don't you know [which pictures of t] Mary bought.

Wh-Islands

- Multiple A′-movers of the same type would also be an option.
- But here the wh-island constraint intervenes.

(8) a. What$_{wh}$ did John say Mary gave t_{wh} to Bill?

b. * What$_{wh0}$ did John say who$_{wh1}$ Mary gave t_{wh0} to t_{wh1}? (Wh-island violation)

c. * What$_{wh0}$ did Bill think which$_{wh1}$ man t_{wh1} says who$_{wh1}$ Mary gave t_{wh0} to t_{wh1}?

d. * What$_{wh0}$ did Sue claim who$_{wh1}$ Bill thinks which$_{wh2}$ man t_{wh2} says Mary gave t_{wh0} to t_{wh1}?

Taking Stock

- MGs are all about two structure-building operations: **Merge** and **Move**.
- Intermediate movement complicates formalism
- SMNF simplifies MGs, but at the risk of larger lexicons.
- Realistic grammars block the truly dangerous configurations.
- Unclear whether this is **coincidence or conspiracy**

References I

Graf, Thomas, Alëna Aksënova, and Aniello De Santo. 2016. A single movement normal form for Minimalist grammars. In *Formal Grammar : 20th and 21st International Conferences, FG 2015, Barcelona, Spain, August 2015, Revised Selected Papers. FG 2016, Bozen, Italy, August 2016*, ed. Annie Foret, Glyn Morrill, Reinhard Muskens, Rainer Osswald, and Sylvain Pogodalla, 200–215. Berlin, Heidelberg: Springer. URL `https://doi.org/10.1007/978-3-662-53042-9_12`.

Graf, Thomas, and Jeffrey Heinz. 2015. Commonality in disparity: The computational view of syntax and phonology. Slides of a talk given at GLOW 2015, April 18, Paris, France.

Graf, Thomas, James Monette, and Chong Zhang. 2017. Relative clauses as a benchmark for Minimalist parsing. *Journal of Language Modelling* 5:57–106. URL `http://dx.doi.org/10.15398/jlm.v5i1.157`.

Kobele, Gregory M. 2006. *Generating copies: An investigation into structural identity in language and grammar*. Doctoral Dissertation, UCLA. URL `http://home.uchicago.edu/~gkobele/files/Kobele06GeneratingCopies.pdf`.

Stabler, Edward P. 1997. Derivational Minimalism. In *Logical aspects of computational linguistics*, ed. Christian Retoré, volume 1328 of *Lecture Notes in Computer Science*, 68–95. Berlin: Springer. URL `https://doi.org/10.1007/BFb0052152`.

Stabler, Edward P. 2011. Computational perspectives on Minimalism. In *Oxford handbook of linguistic Minimalism*, ed. Cedric Boeckx, 617–643. Oxford: Oxford University Press.