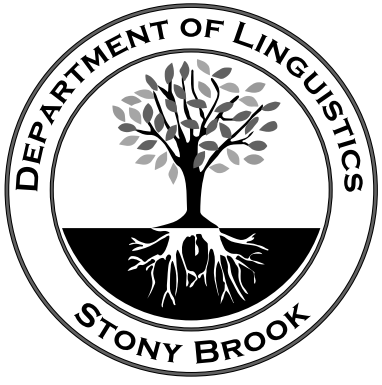
Subregular Tree Transductions, Movement, Copies, Traces, and the Ban on Improper Movement

(most succinct title ever)

Thomas Graf



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

June 16, 2023

The talk in a nutshell

- Movement is **tier-based strictly local (TSL)**.
  1. as a syntactic dependency (Graf 2018)
  2. as a structure building operation (this talk)
- There's some interesting wrinkles.
  - **conceptual:** copies (often) easier than traces
  - **empirical:** natural languages obey additional conditions that are needed for traces/linearization

Formal core contribution

Input-strictly local tree-to-tree transductions with lexical TSL tests (ISL[lexTSL])

Linguistic core contribution

Output-oriented
Ban on Improper Movement (OOBoIM)

Outline

1. Movement is local over tiers
2. Subregular transductions for syntax
   - ISL transductions for Merge
   - Adding tiers for movement
3. The challenge of linearization
4. Extra stuff for the MG nerds

Overall framework

- **Subregular program**
  - identify subclasses of finite-state machinery over strings & trees for phenomena in phonology, morphology, syntax, semantics (Heinz 2018; Chandlee 2022; Graf 2022a,b)
  - TSL an important subregular class

Jeff Heinz

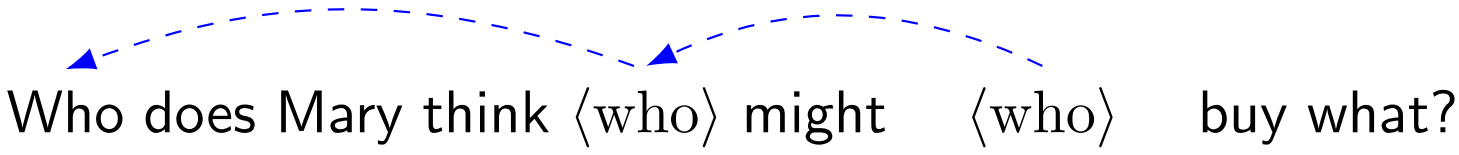
- **Minimalist grammars (MGs)**
  - formalization of Minimalist syntax (Stabler 1997, 2011)
  - allows us to stick closely to syntactic proposals
  - subregular syntax uses a modified version (sets of licensee features, not strings)

Ed Stabler

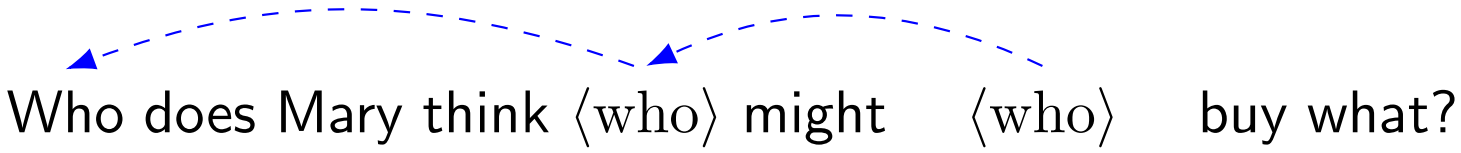
Minimalist grammars and movement

Who does Mary think might buy what?

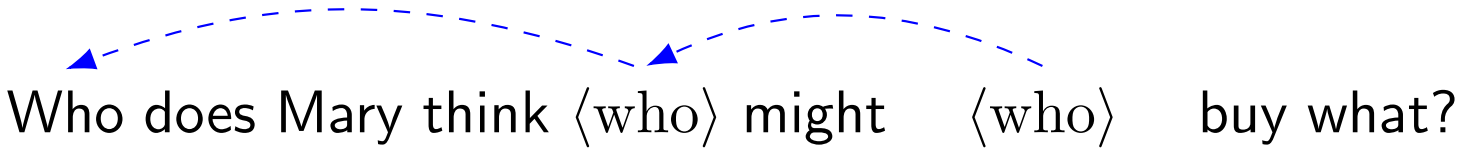
Minimalist grammars and movement



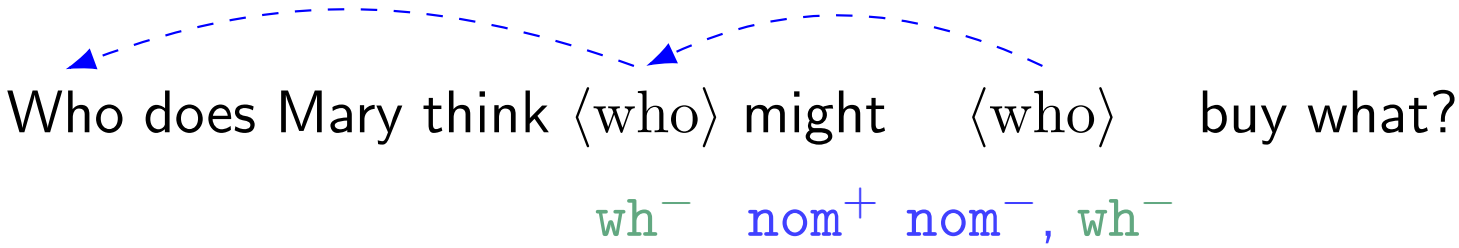
Minimalist grammars and movement



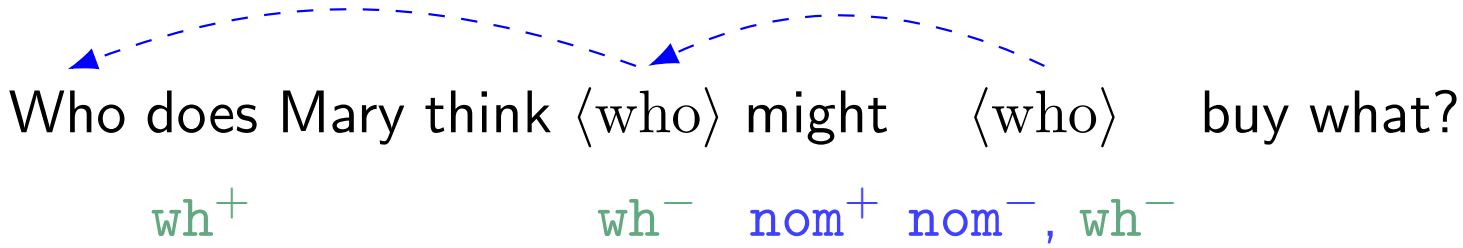
Minimalist grammars and movement



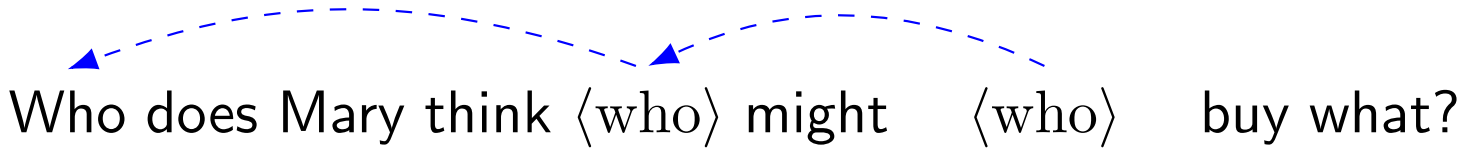
Minimalist grammars and movement



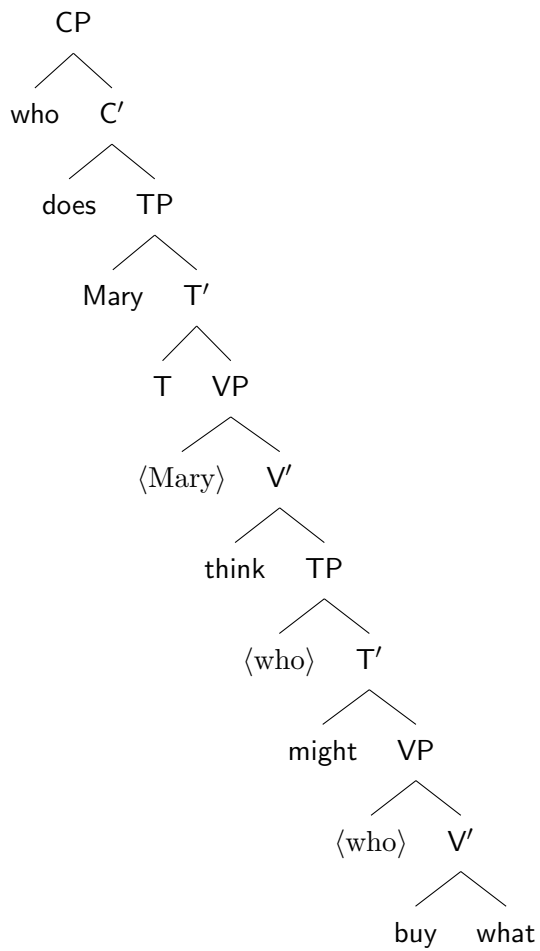
Minimalist grammars and movement



Minimalist grammars and movement



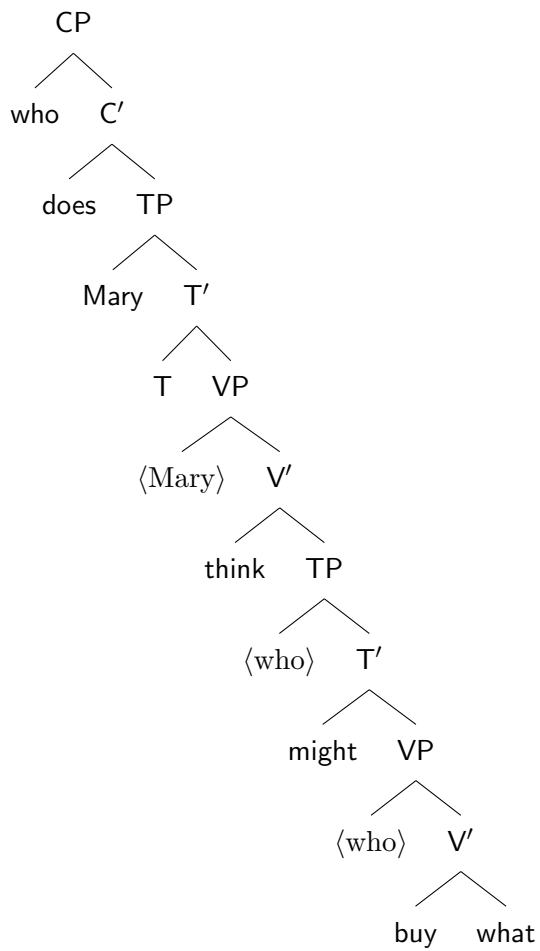
$\texttt{wh}^+$ $\texttt{wh}^-$ $\texttt{nom}^+$ $\texttt{nom}^-$, $\texttt{wh}^-$

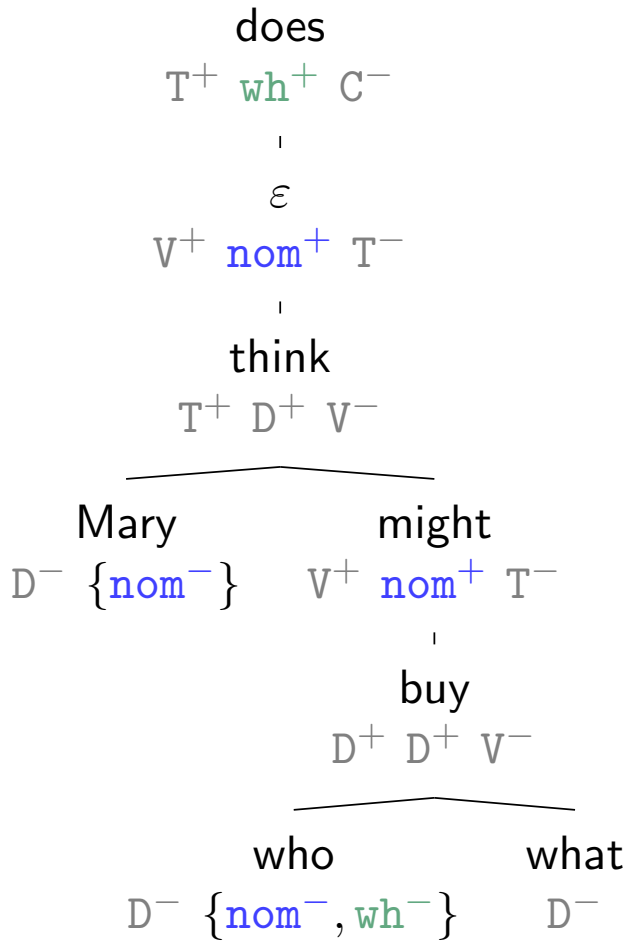


Minimalist grammars and movement

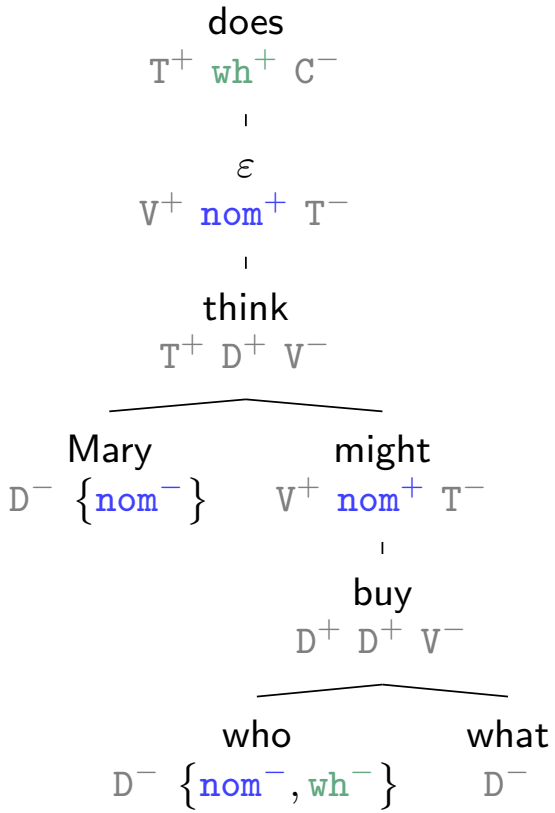
Who does Mary think ⟨who⟩ might ⟨who⟩ buy what?

$\texttt{wh}^+$ $\texttt{wh}^-$ $\texttt{nom}^+$ $\texttt{nom}^-$, $\texttt{wh}^-$



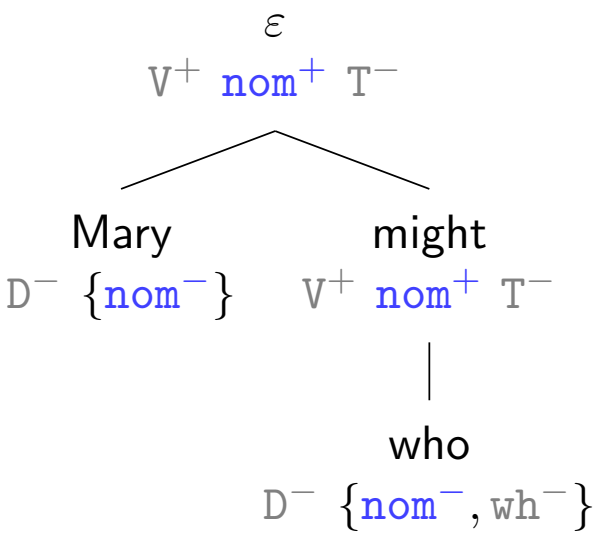


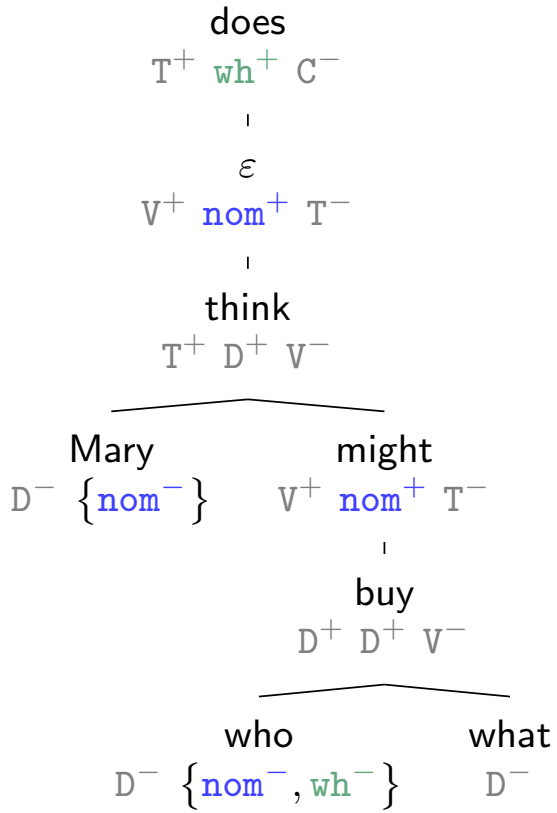
Movement in MGs is tier-based strictly local (TSL)



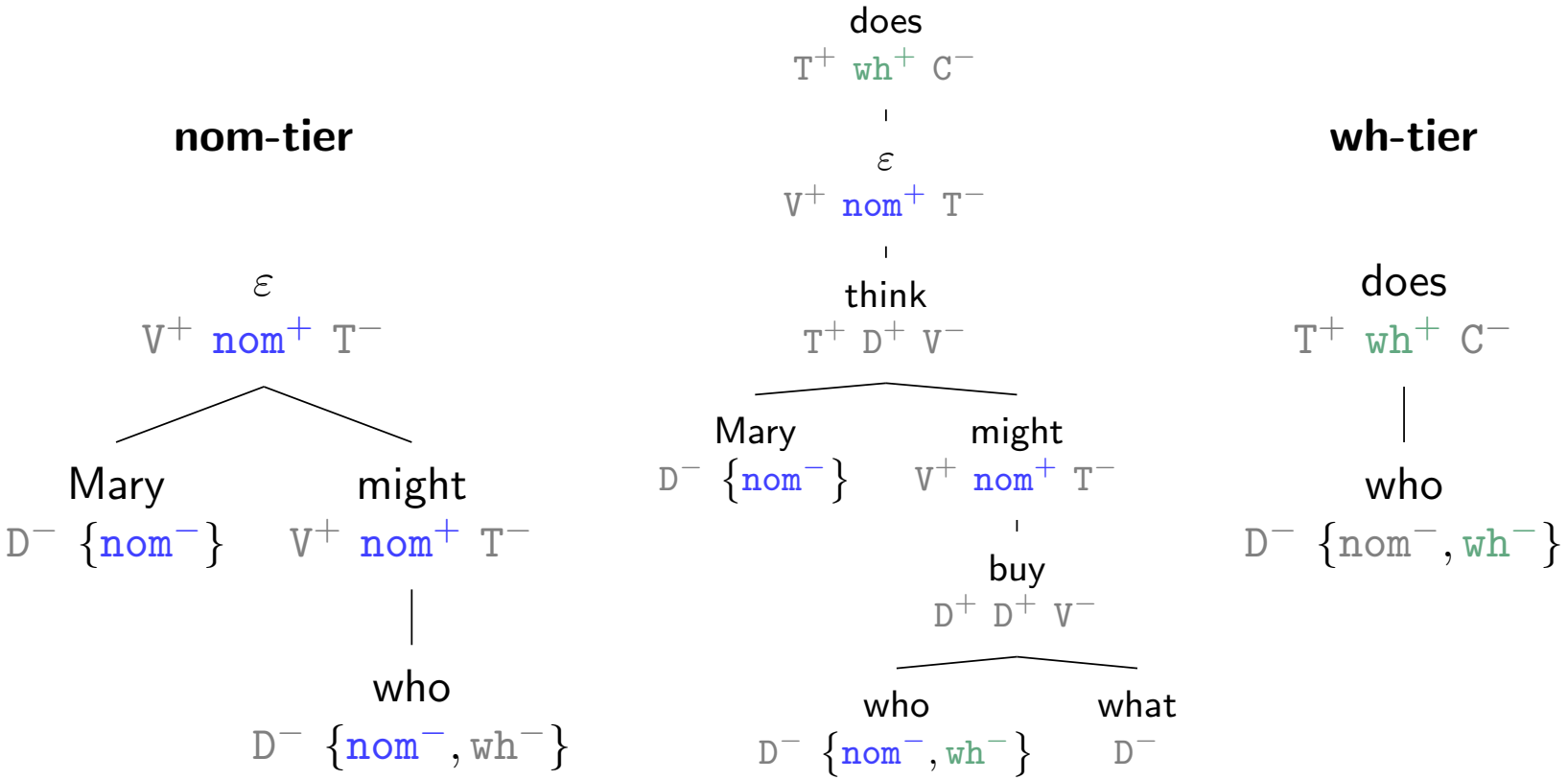
Movement in MGs is tier-based strictly local (TSL)

nom-tier

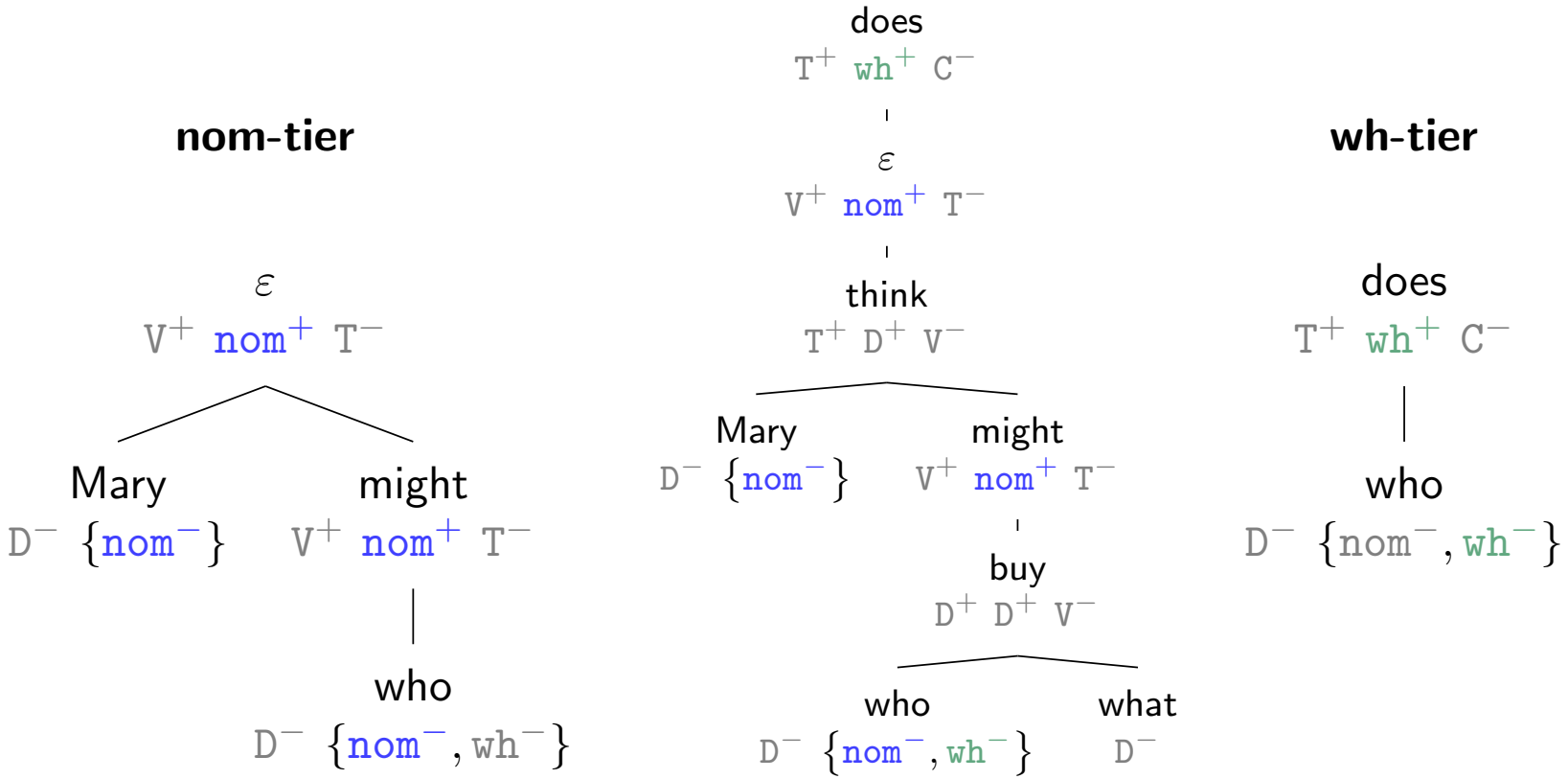




Movement in MGs is tier-based strictly local (TSL)



Movement in MGs is tier-based strictly local (TSL)



Feature checking as a constraint on each f-tier

1. Every f^+ has **exactly 1** f^- among its daughters.
2. Every f^- has f^+ **as its mother**.

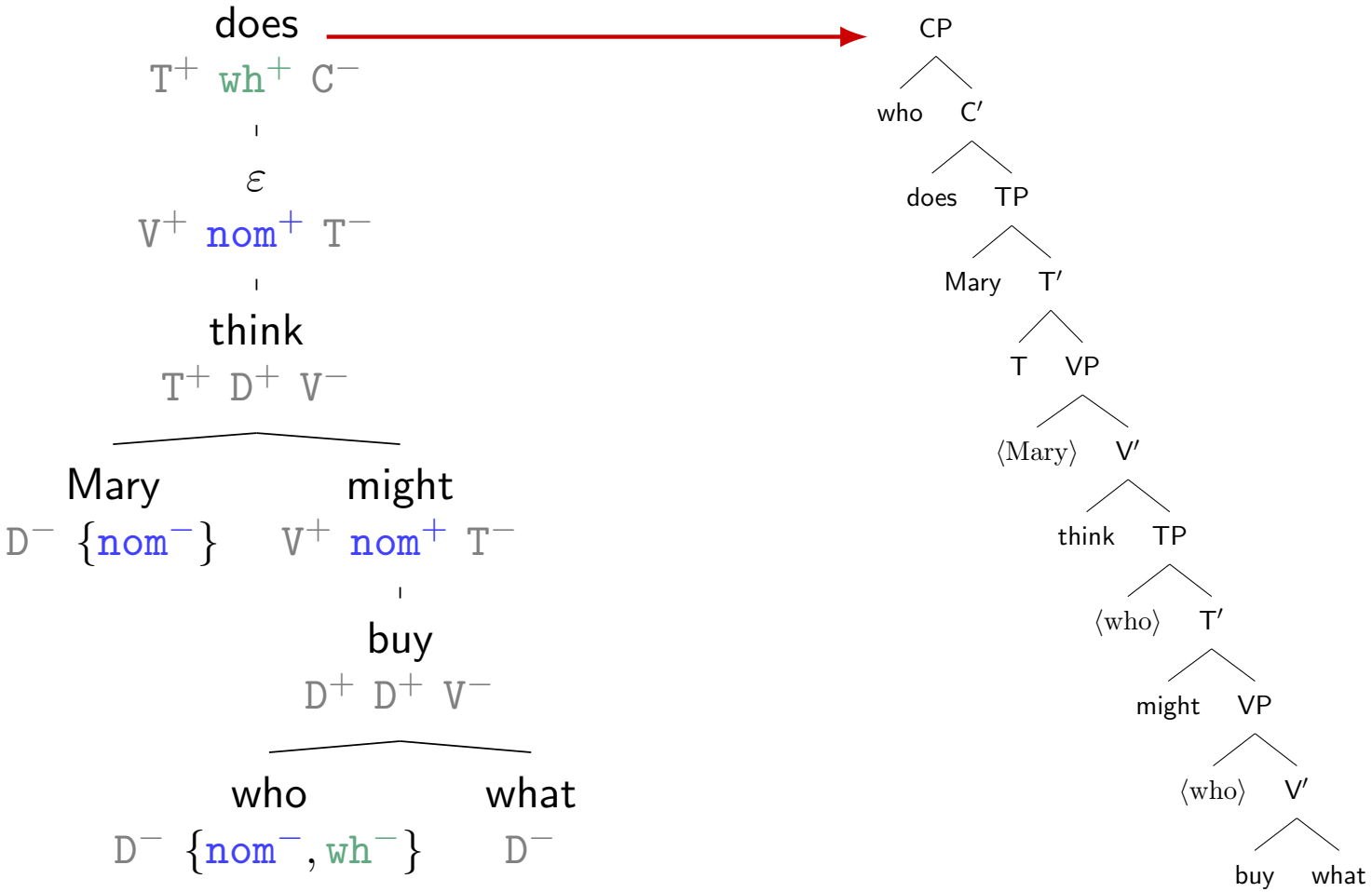
Who cares?

All of this is part of **a bigger program**.

1. **Learning**
   exploit subregular restrictions for new learning algorithms
   string extension learning, BUFIA
2. **Typology**
   explain crosslinguistic absence of certain types of patterns
   no "cowardly" island effects
3. **Universals**
   derive old/new universals from limits of machinery
   ISL-recoverability of category features, OOBoIM
4. **Theoretical unification**
   unify seemingly unrelated patterns
   vowel harmony & extraction morphology
5. **Cognitive parallelism**
   computational common core across language modules
   central role of SL and TSL

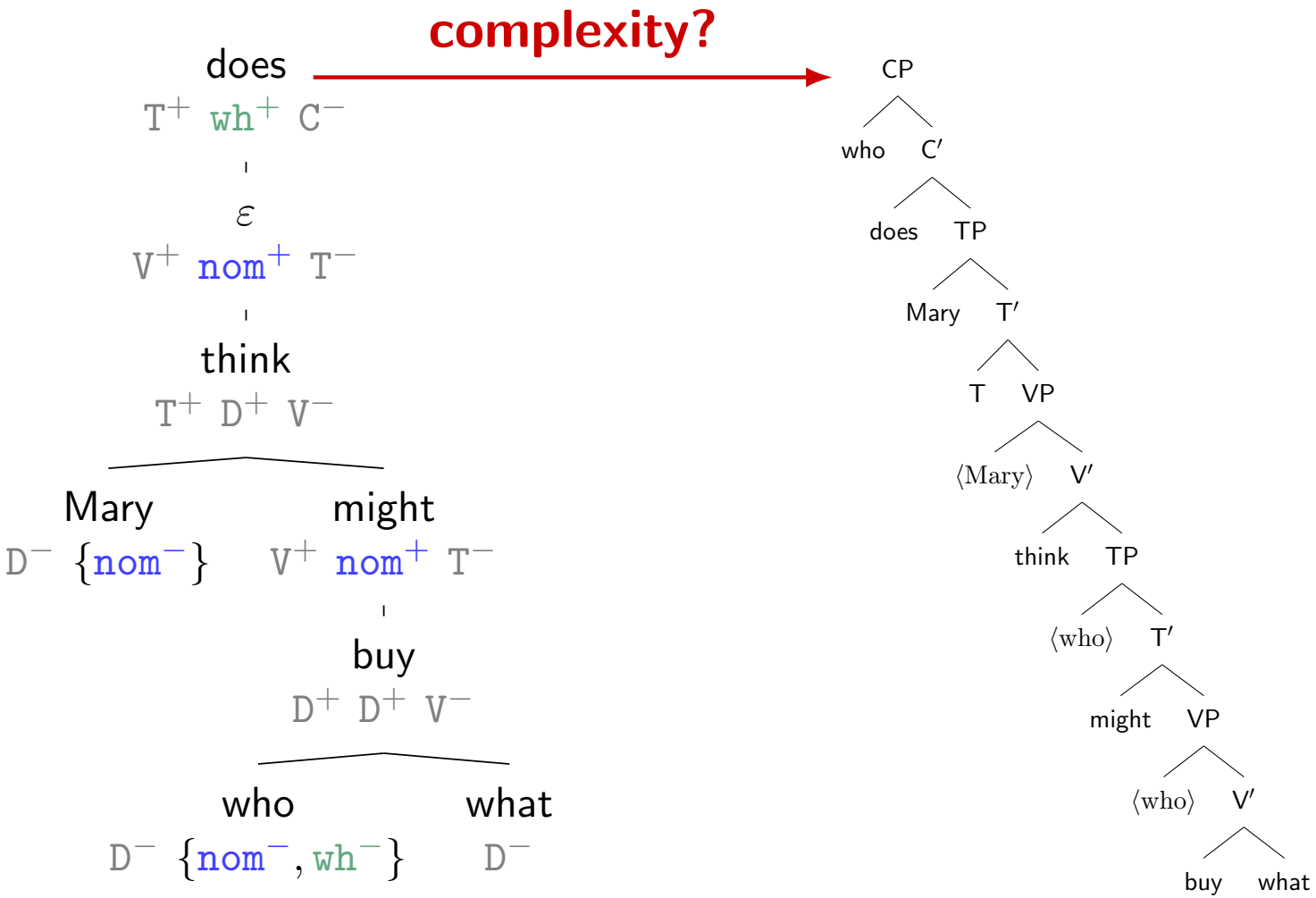
An incomplete picture

Movement isn't just a syntactic dependency, it **affects the output**.



An incomplete picture

Movement isn't just a syntactic dependency, it **affects the output**.



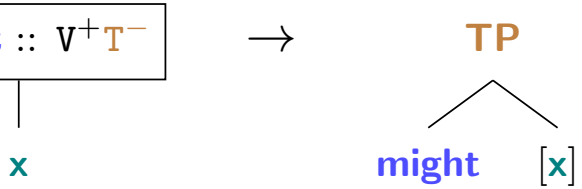
Phrase structure trees without movement

no movement $\Rightarrow$ easy translation to phrase structure trees

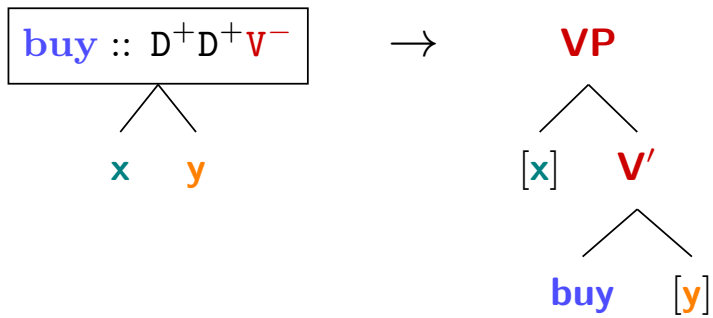
who :: $\mathtt{D}^-$ $\rightarrow$ who

what :: $\mathtt{D}^-$ $\rightarrow$ what

might :: $\mathtt{V}^+\mathtt{T}^-$ $\rightarrow$



buy :: $\mathtt{D}^+\mathtt{D}^+\mathtt{V}^-$ $\rightarrow$



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D⁻ → who

what :: D⁻ → what

might :: V⁺T⁻ → TP

x

might [x]

buy :: D⁺D⁺V⁻ → VP

x y

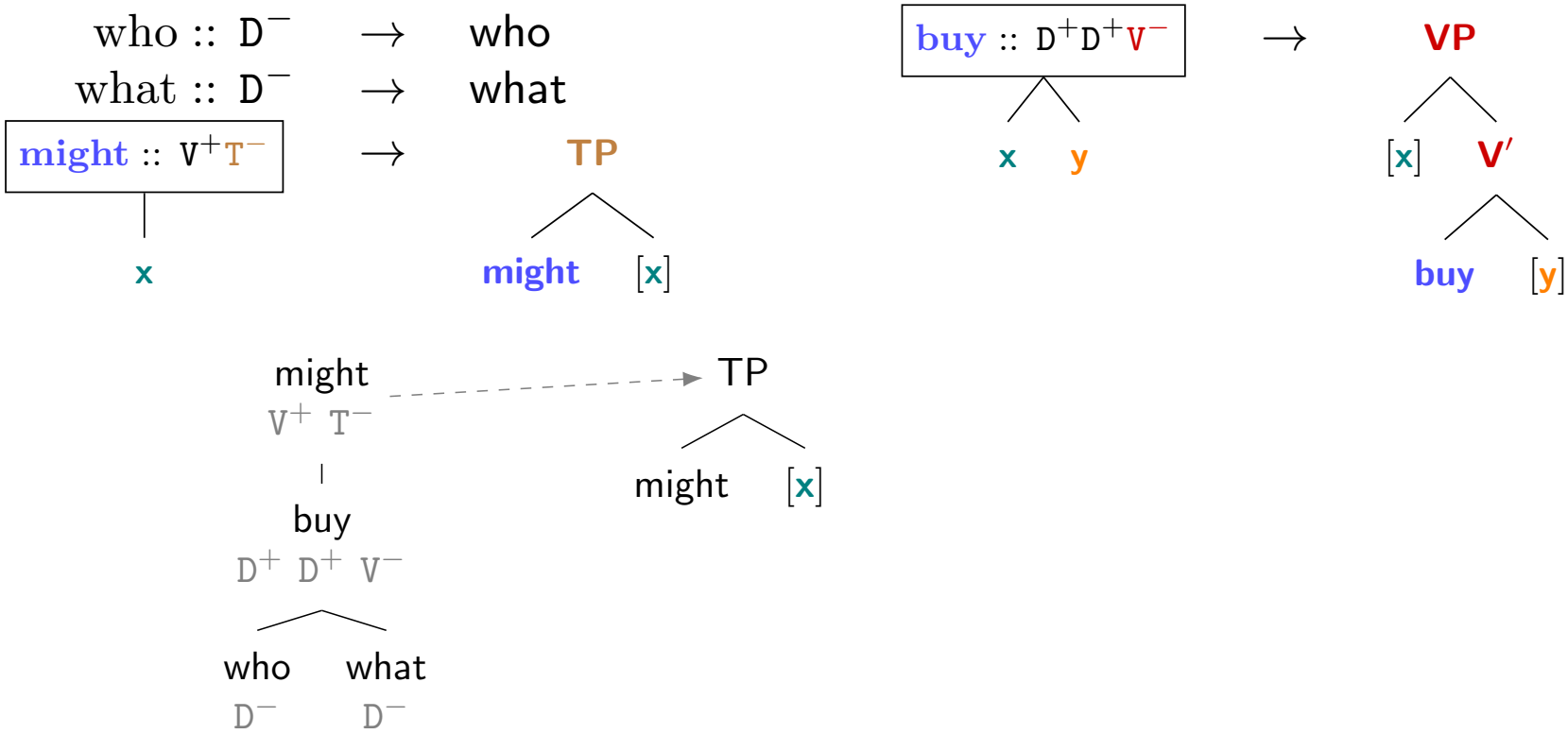
[x] V′

buy [y]

might
V⁺ T⁻
|
buy
D⁺ D⁺ V⁻
who what
D⁻ D⁻

Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D^- → who

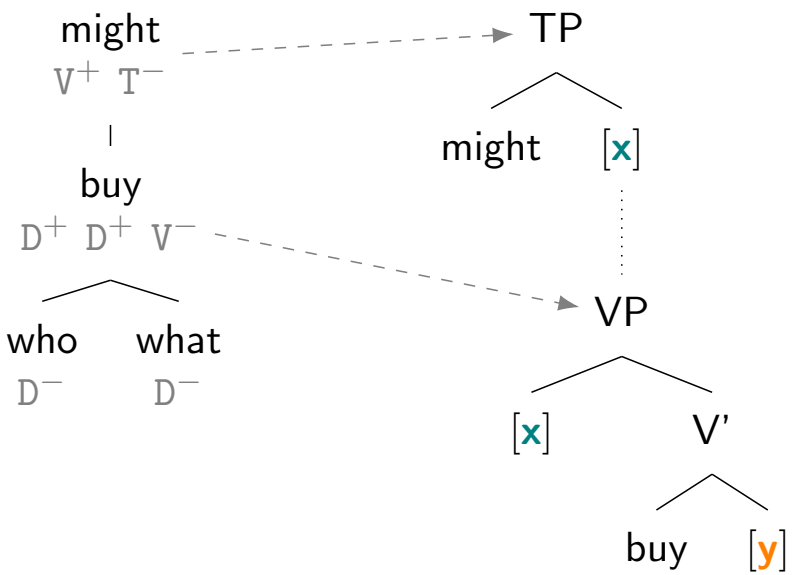
what :: D^- → what

might :: V^+T^- → TP

x — might [x]

buy :: $D^+D^+V^-$ → VP

x y — [x] V′ — buy [y]



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D⁻ → who

what :: D⁻ → what

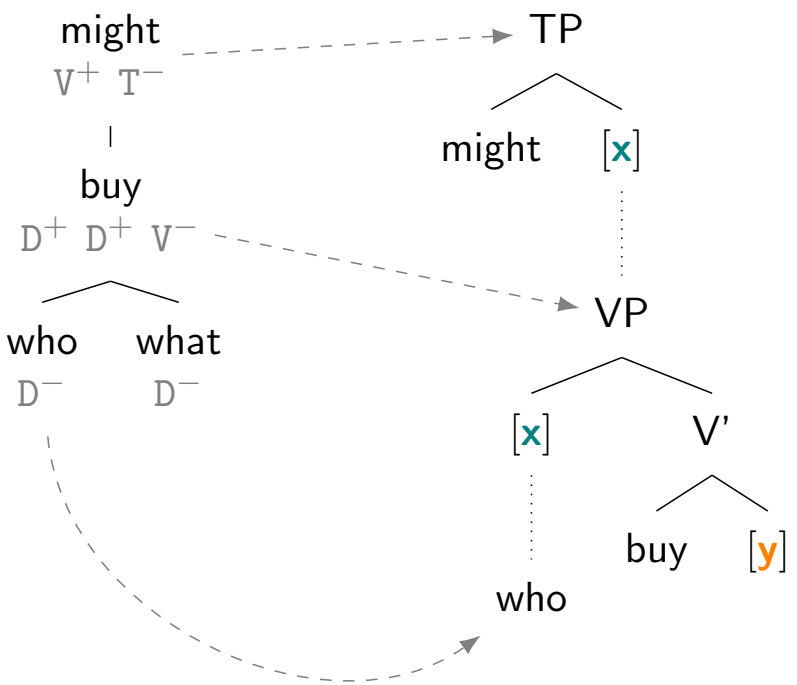
might :: V⁺T⁻ → TP

x → might [x]

buy :: D⁺D⁺V⁻ → VP

x y → [x] V′

buy [y]



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D⁻ → who

what :: D⁻ → what

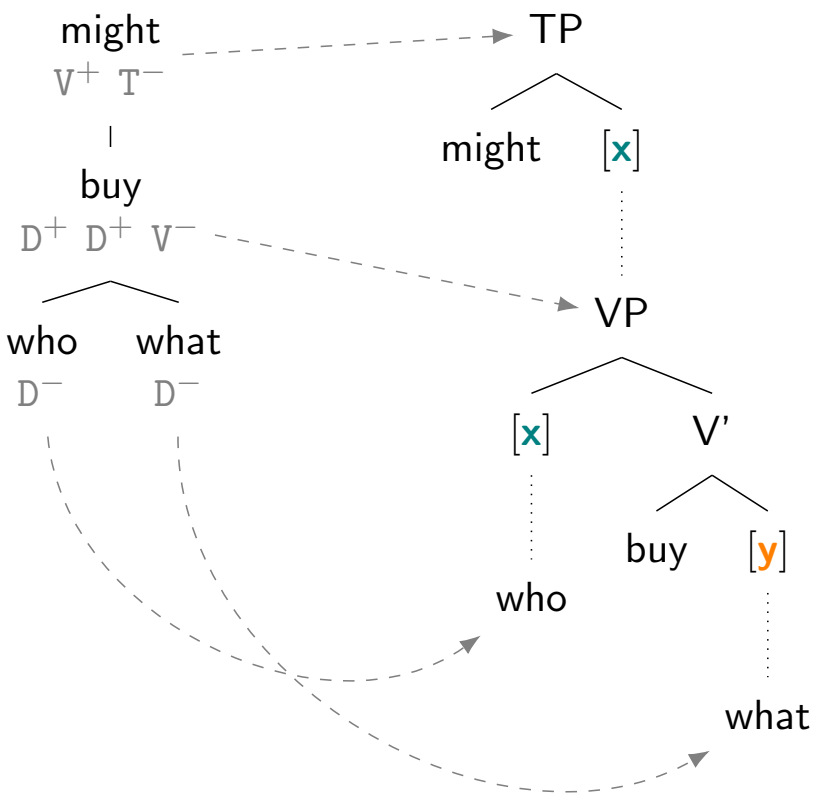
might :: V⁺T⁻ → TP

x → might [x]

buy :: D⁺D⁺V⁻ → VP

x y → [x] V′

buy [y]



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D⁻ → who

what :: D⁻ → what

might :: V⁺T⁻ → TP

x

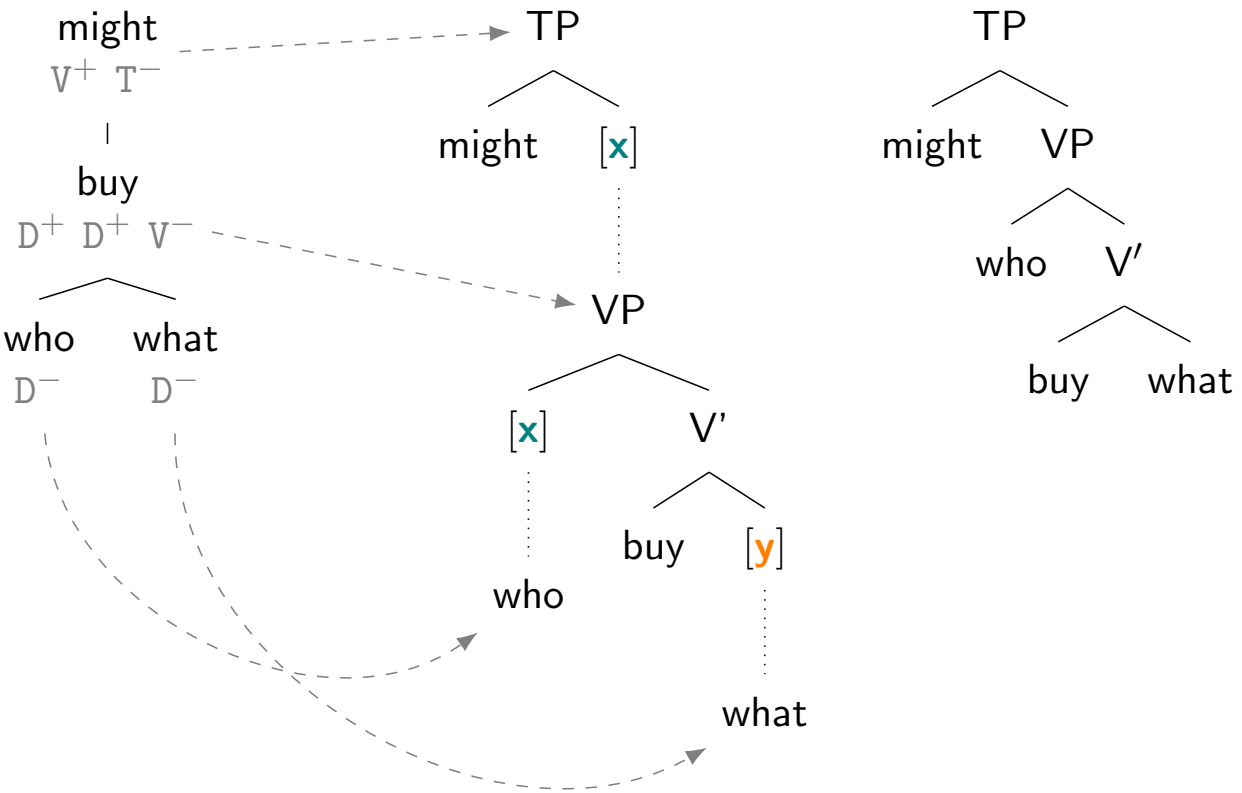
might [x]

buy :: D⁺D⁺V⁻ → VP

x y

[x] V′

buy [y]



Phrase structure trees without movement

no movement ⇒ easy translation to phrase structure trees

who :: D⁻ → who

what :: D⁻ → what

might :: V⁺T⁻ → TP

x

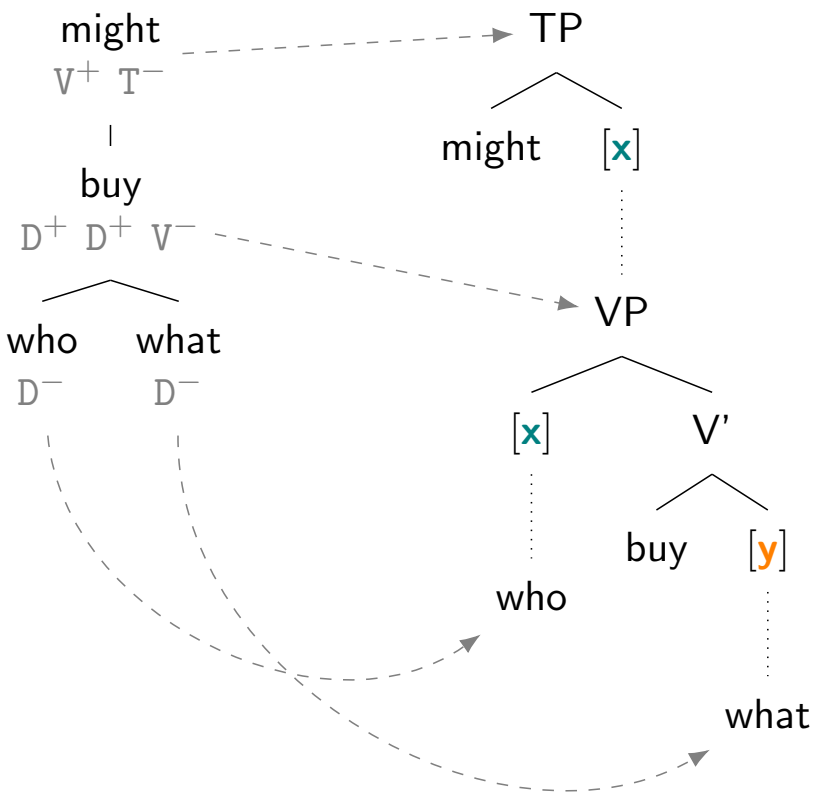
might [x]

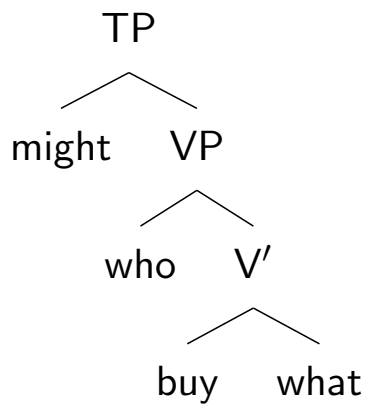
buy :: D⁺D⁺V⁻ → VP

x y

[x] V′

buy [y]





≈ **Tree Substitution Grammar**

ISL transductions

Movement-free translation is **input strictly local (ISL)**.

The intuition behind ISL

Each node **n** in input **i** is rewritten as some finite structure based only on **n** itself and its local context in **i**.

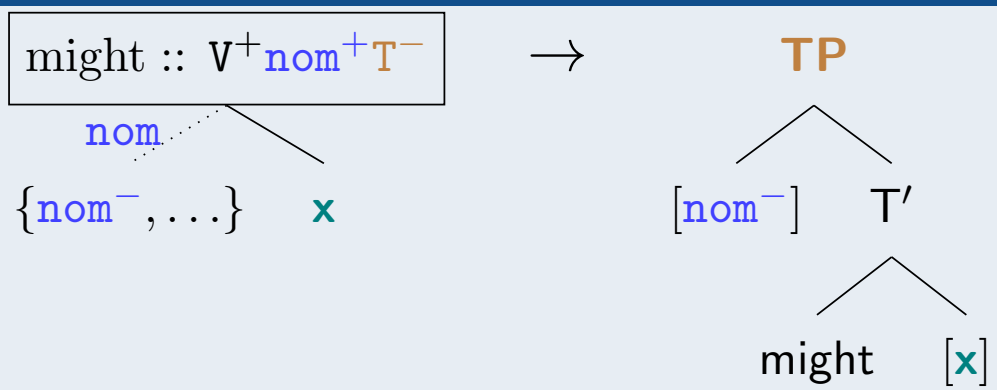
- ubiquitous in phonology, morphology (Chandlee 2014; Chandlee and Heinz 2018)
- tree-version for syntax originally proposed for very different purposes (Graf 2020; see also Ji and Heinz 2020 and Ikawa et al. 2020)

Jane Chandlee

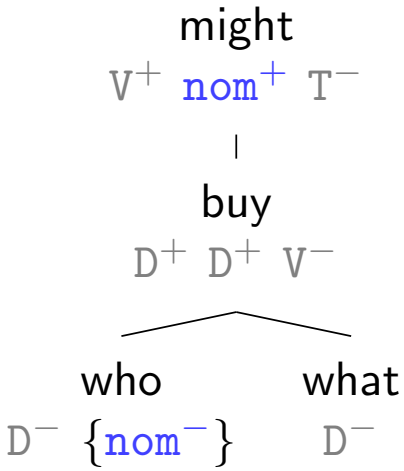
Adding tier relations

- ISL limited to local contexts $\Leftrightarrow$ unbounded movement
- **But:** movement is **local over tiers**
- suffices to enrich ISL rewrite rules with tier-daughter relation

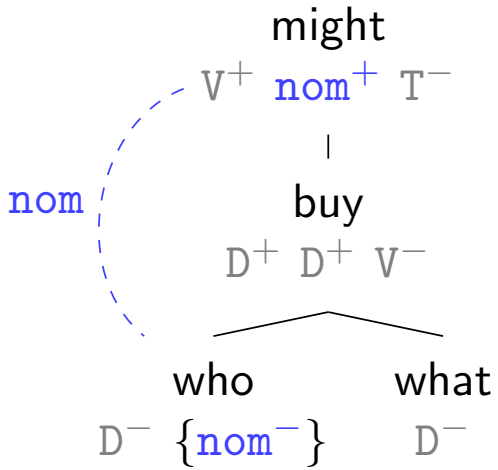
ISL rewrite rule with nom-daughter relation



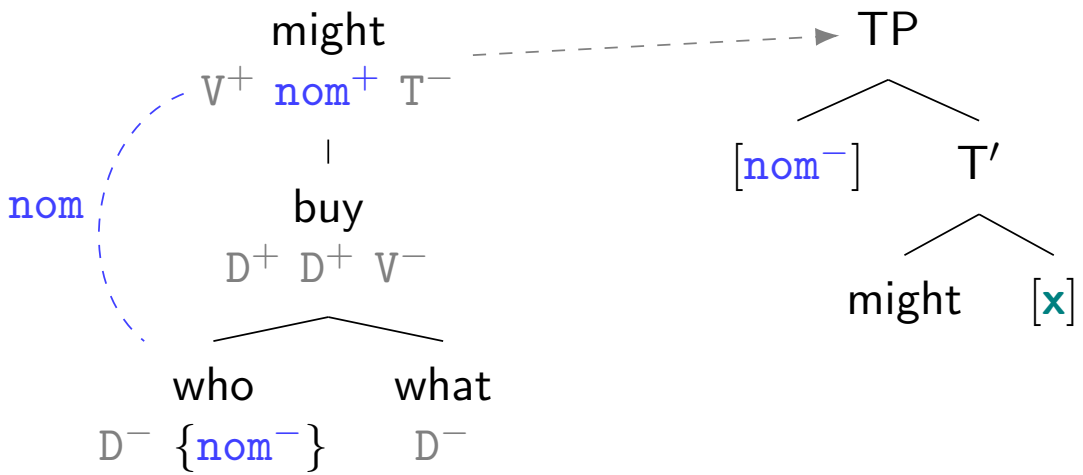
Example with subject movement



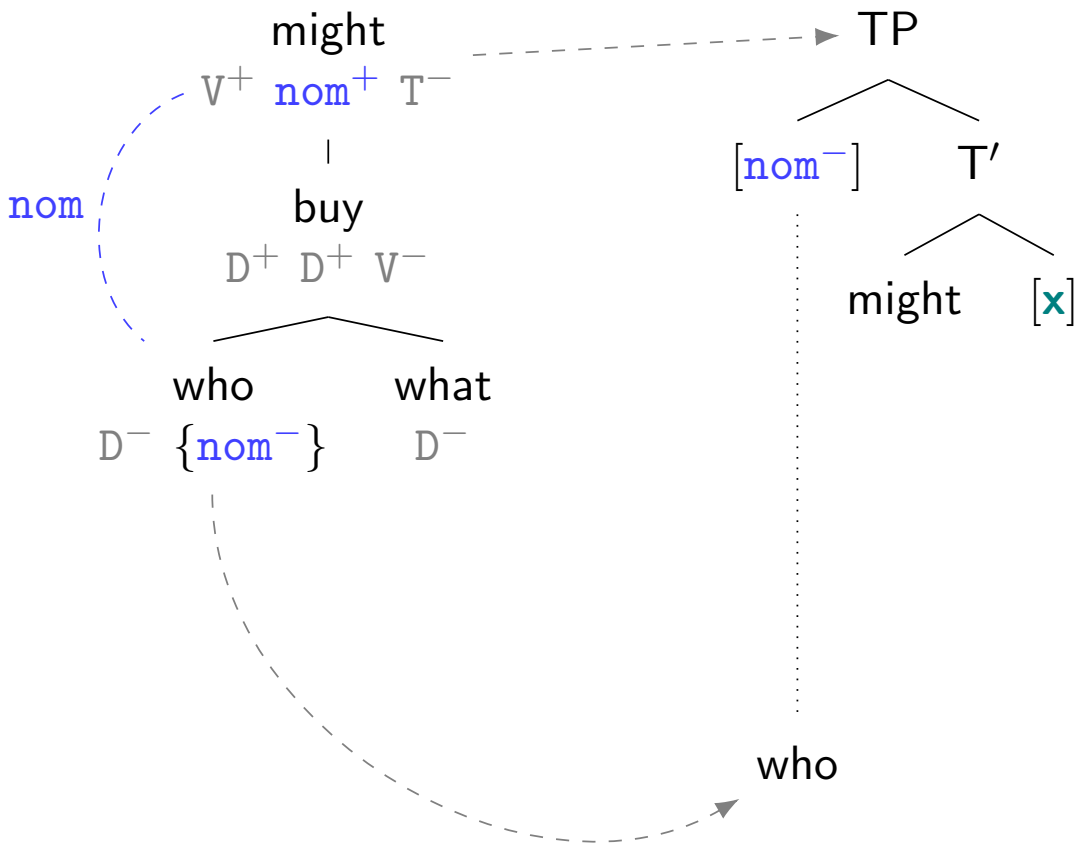
Example with subject movement



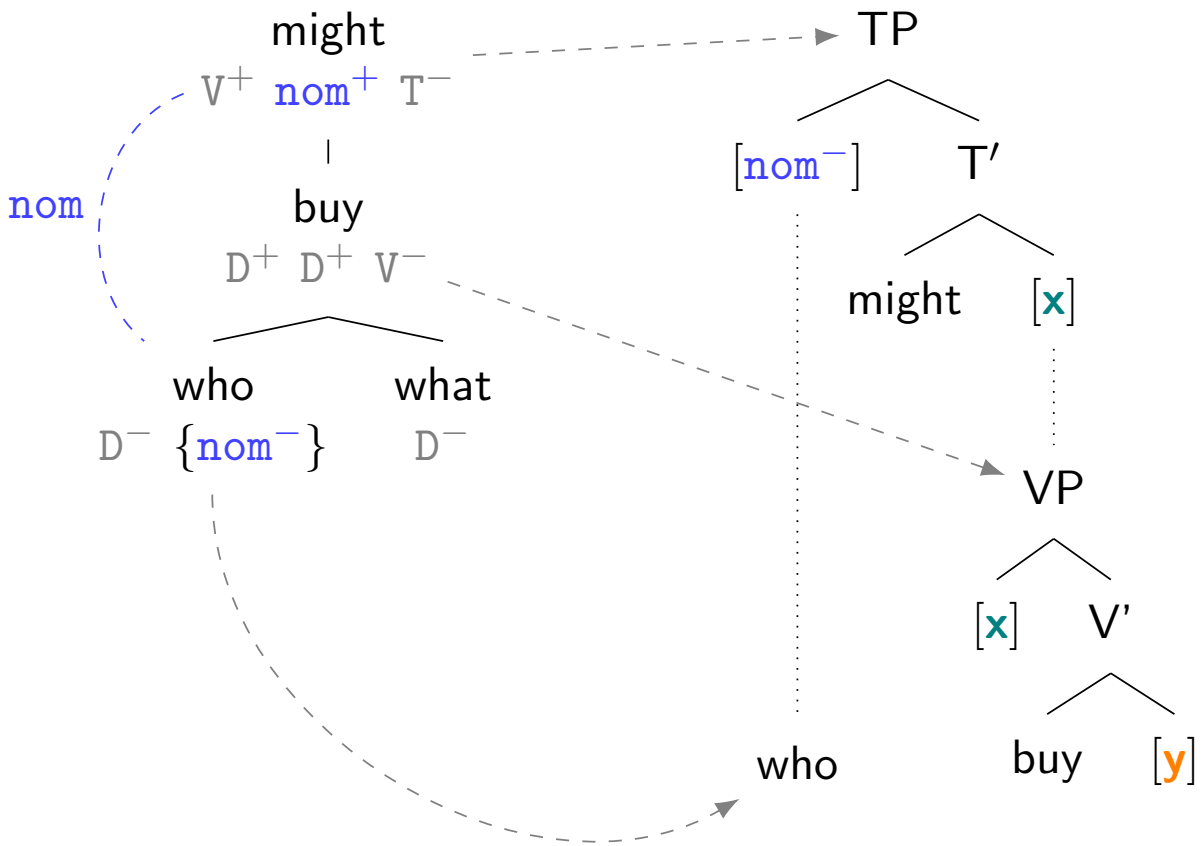
Example with subject movement



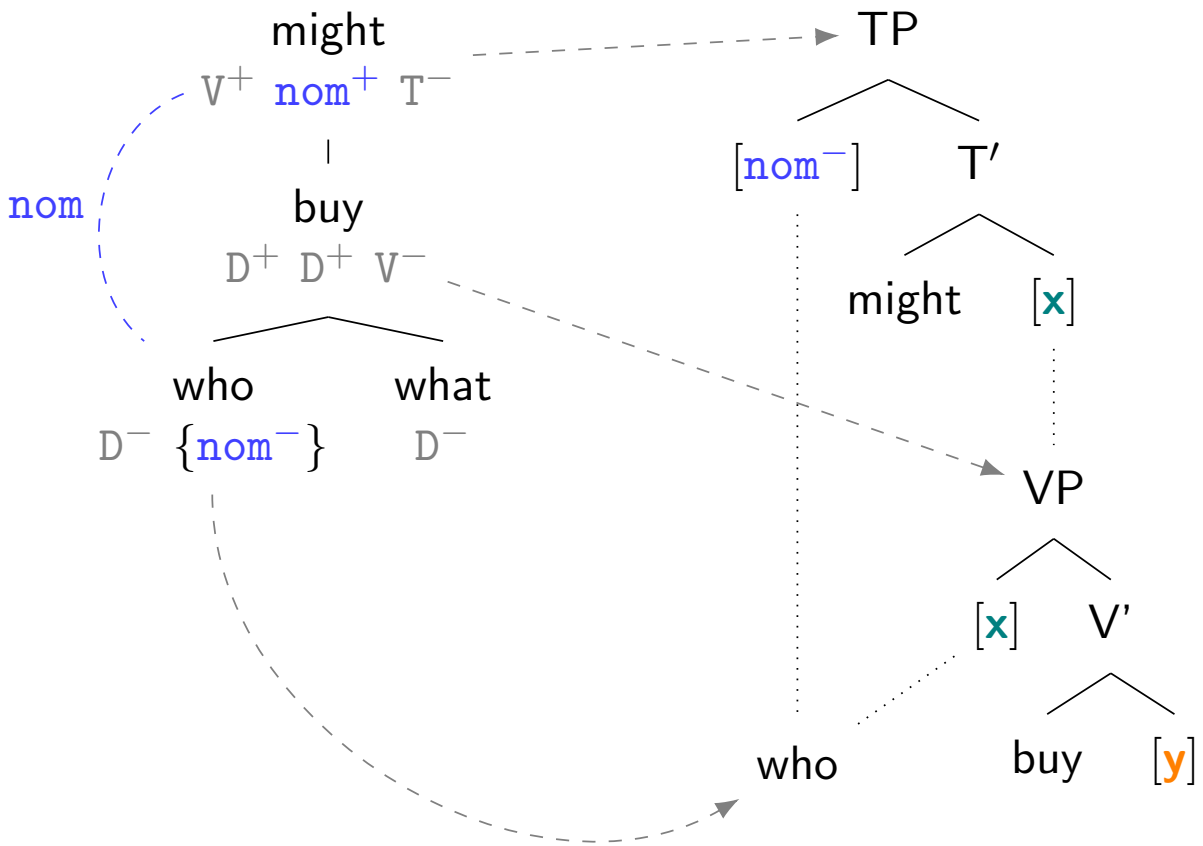
Example with subject movement



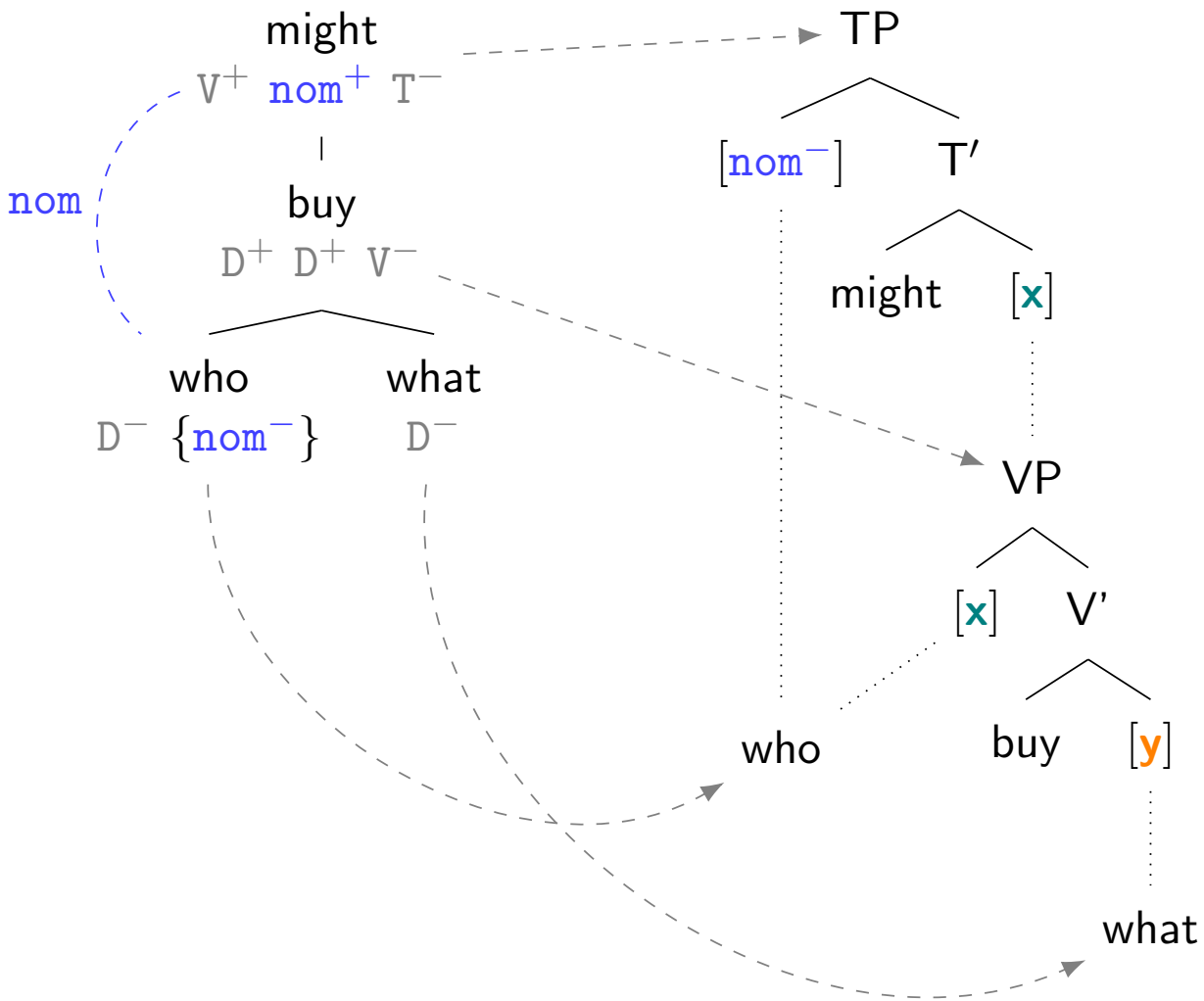
Example with subject movement



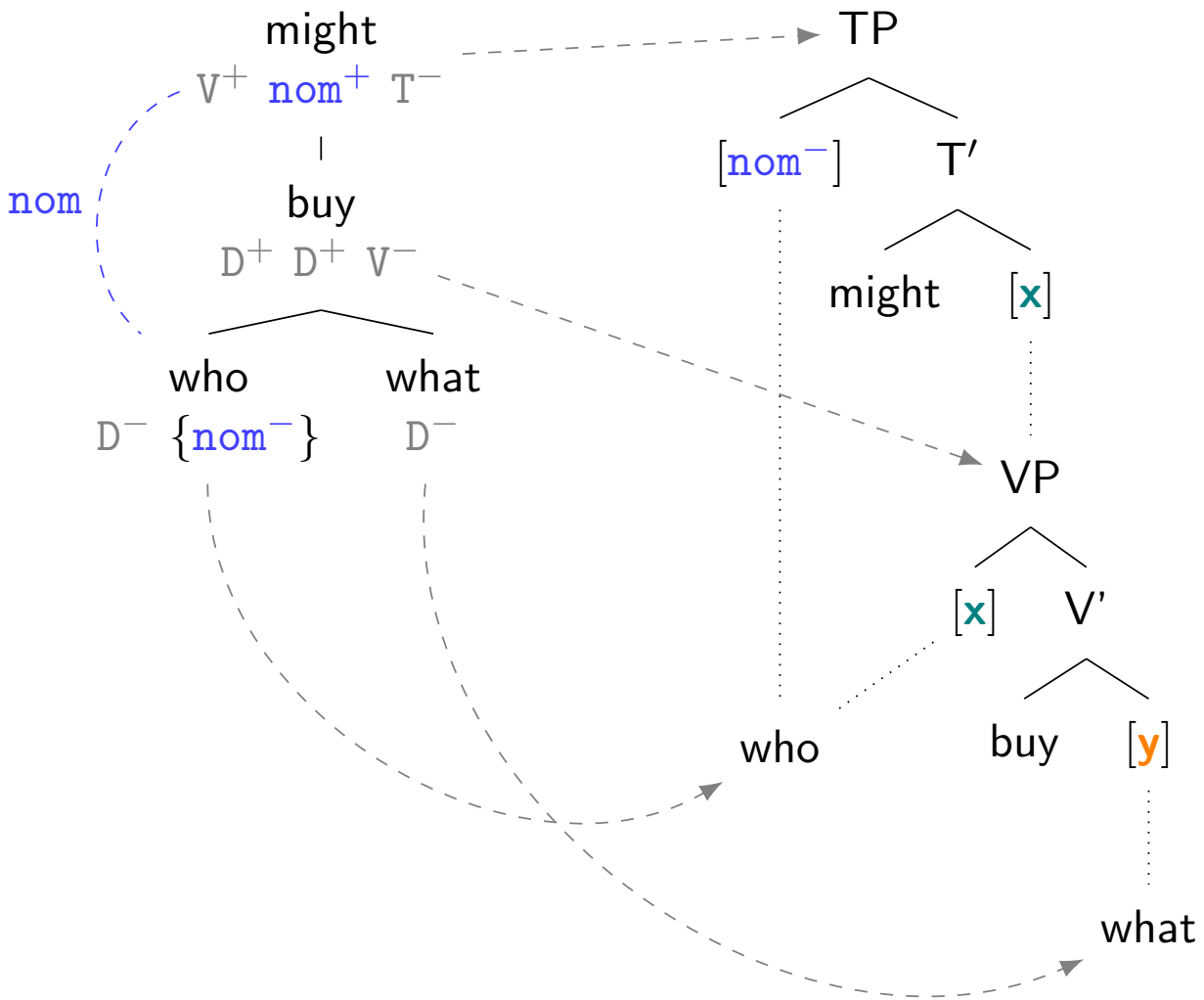
Example with subject movement

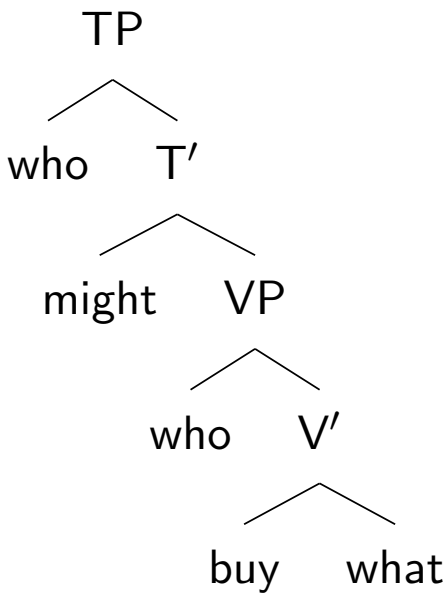


Example with subject movement

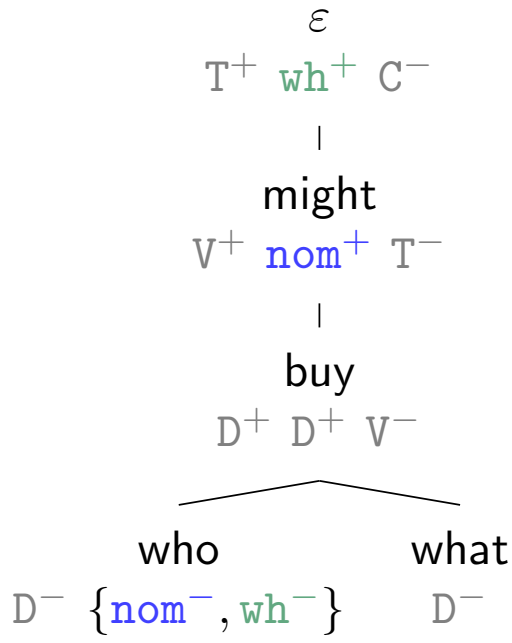


Example with subject movement

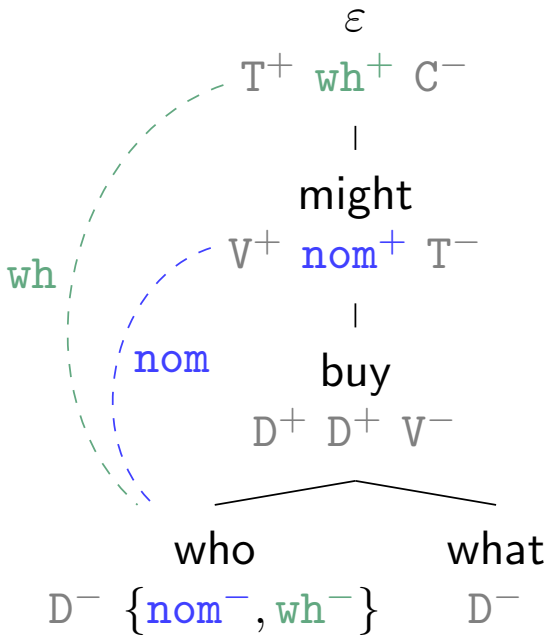




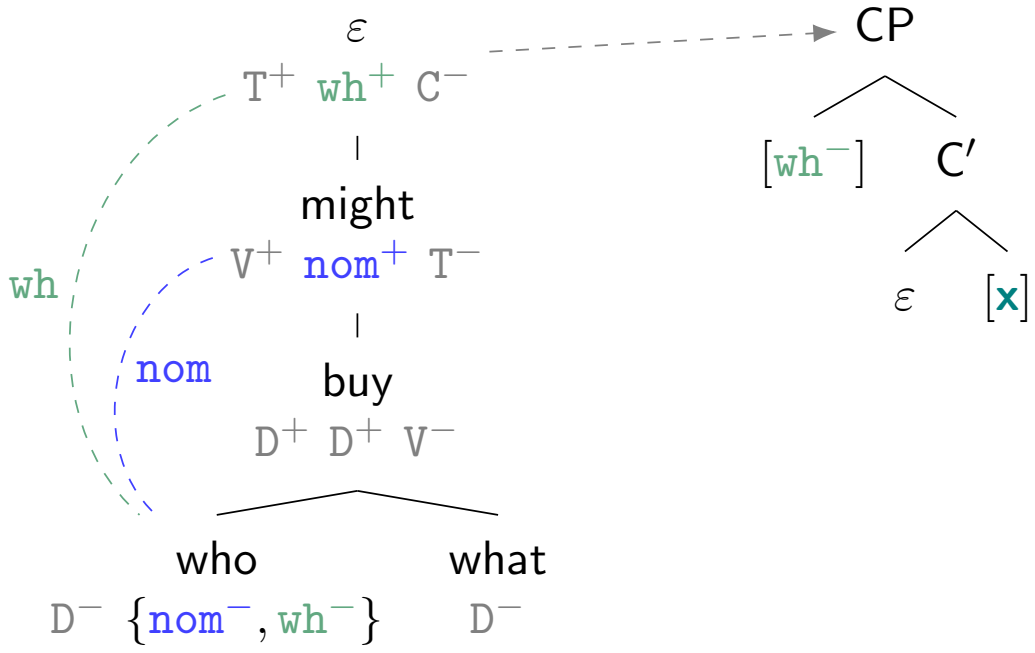
Multiple copies are straightforward



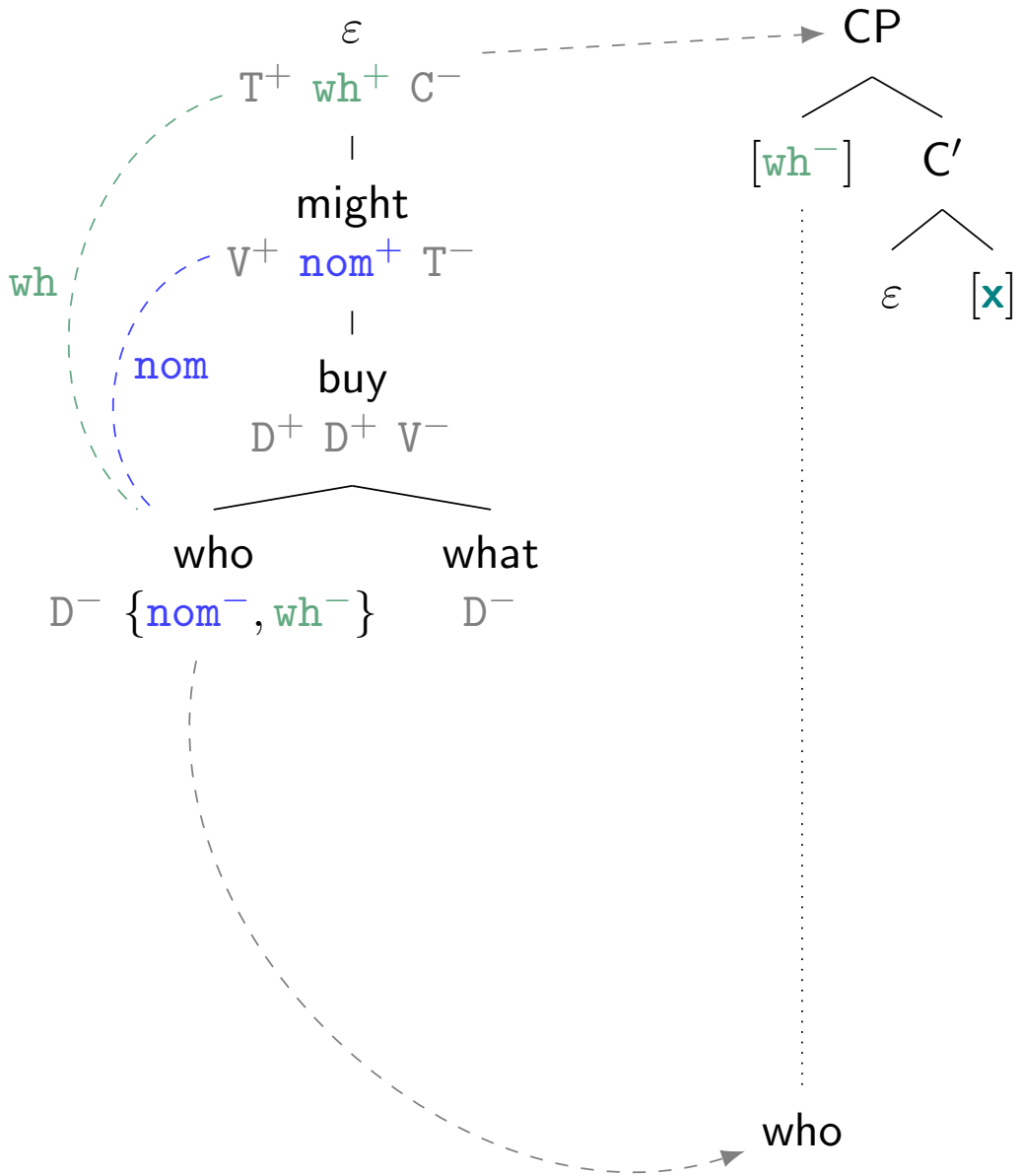
Multiple copies are straightforward



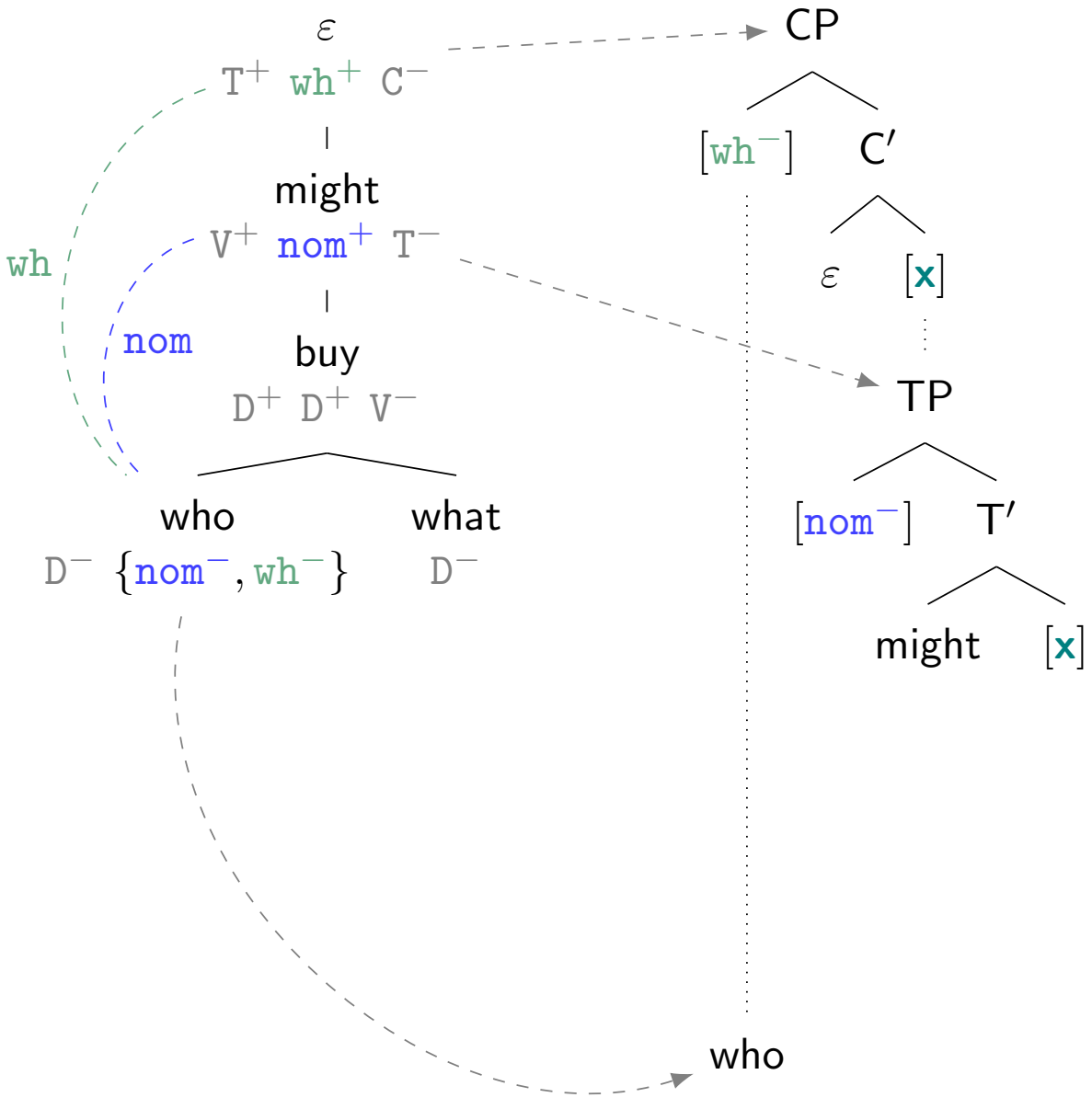
Multiple copies are straightforward



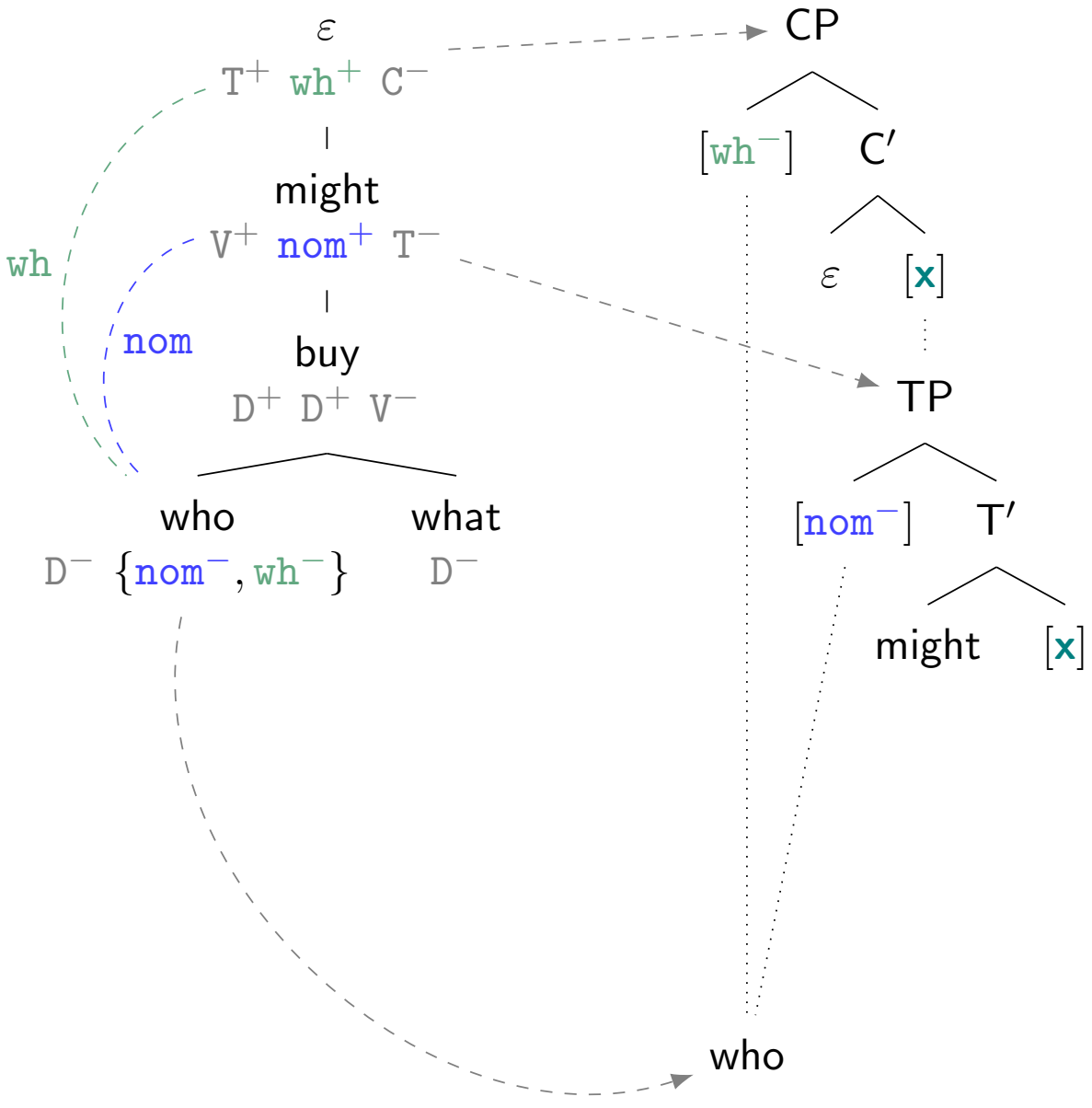
Multiple copies are straightforward



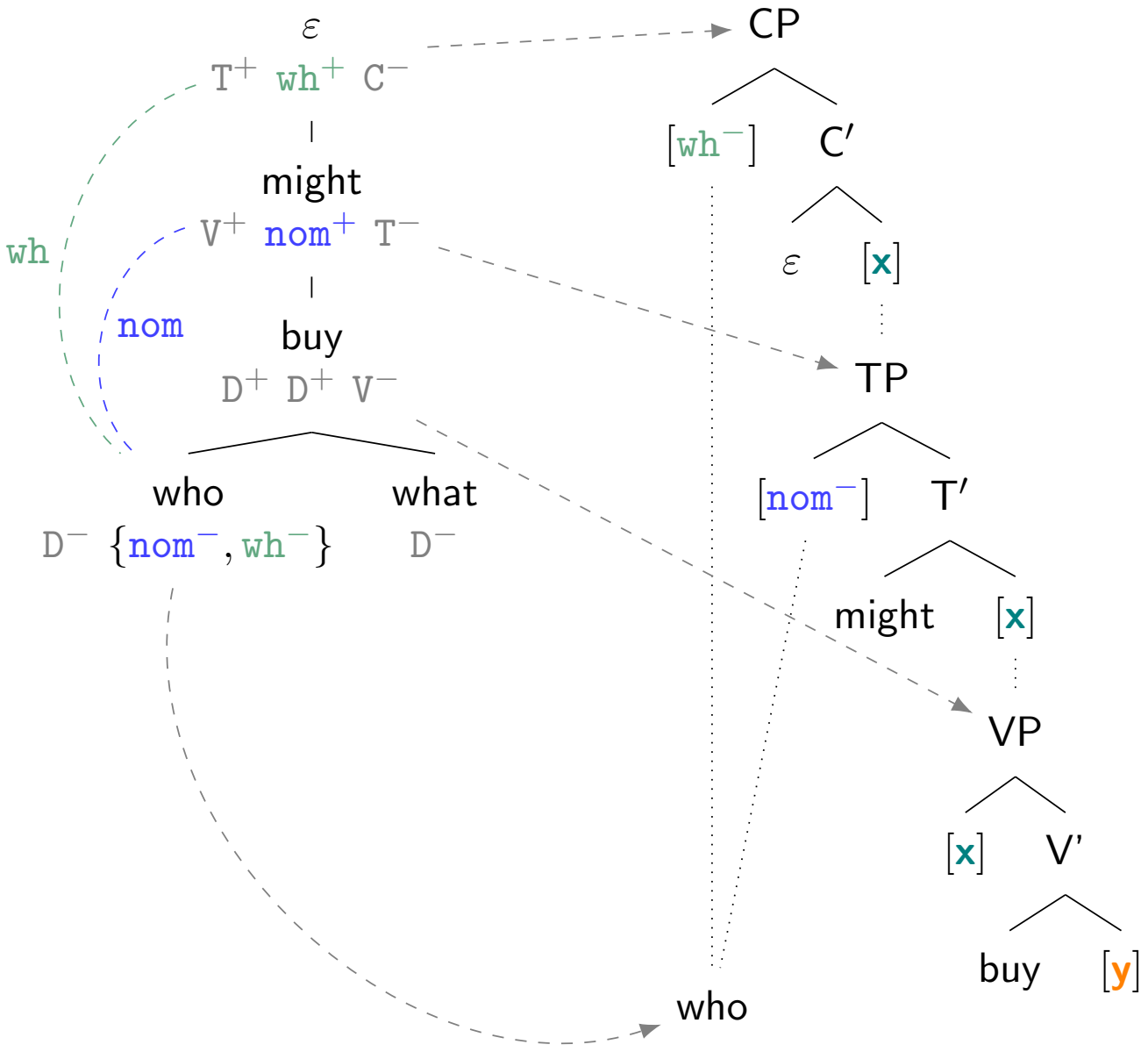
Multiple copies are straightforward



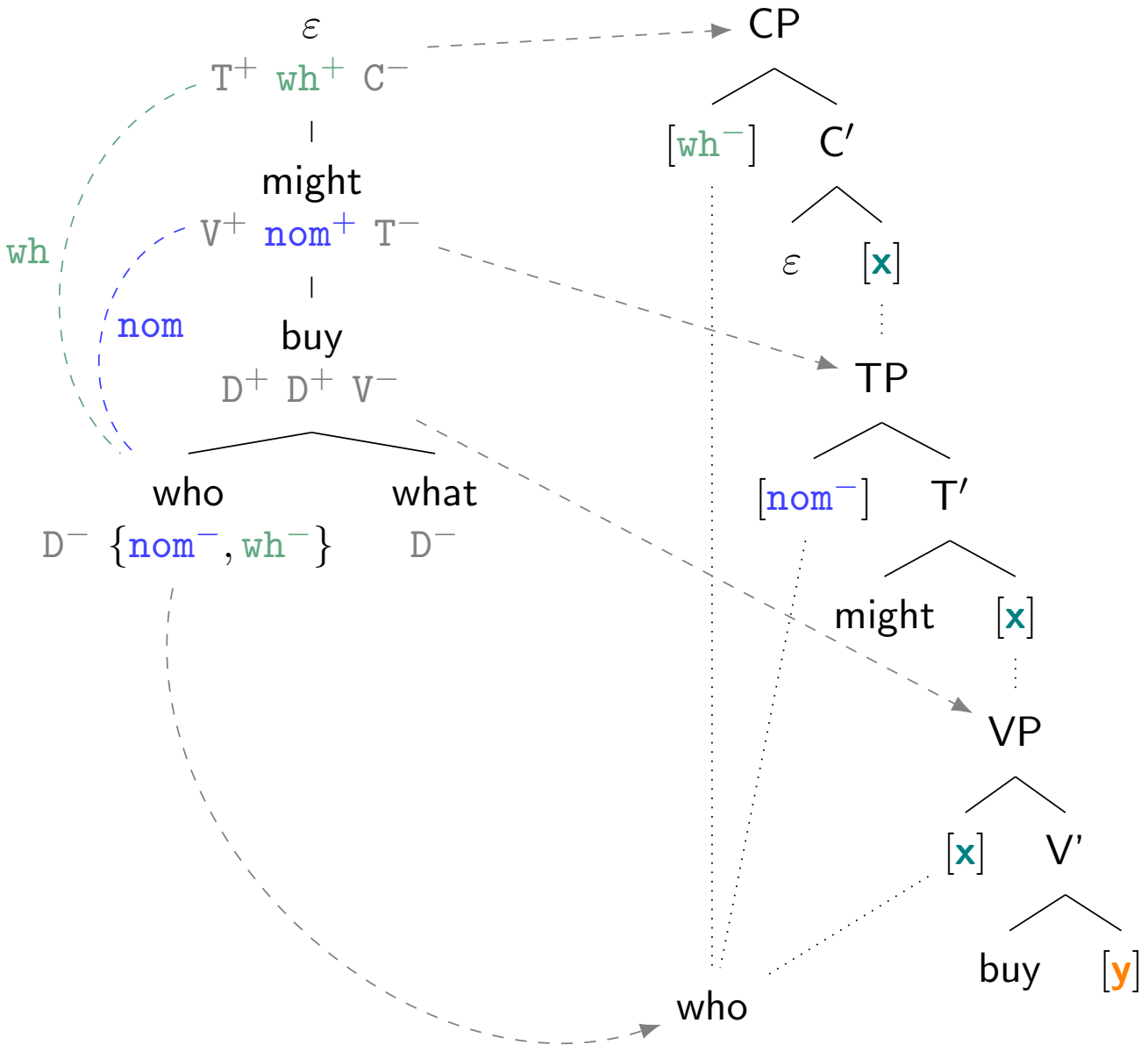
Multiple copies are straightforward



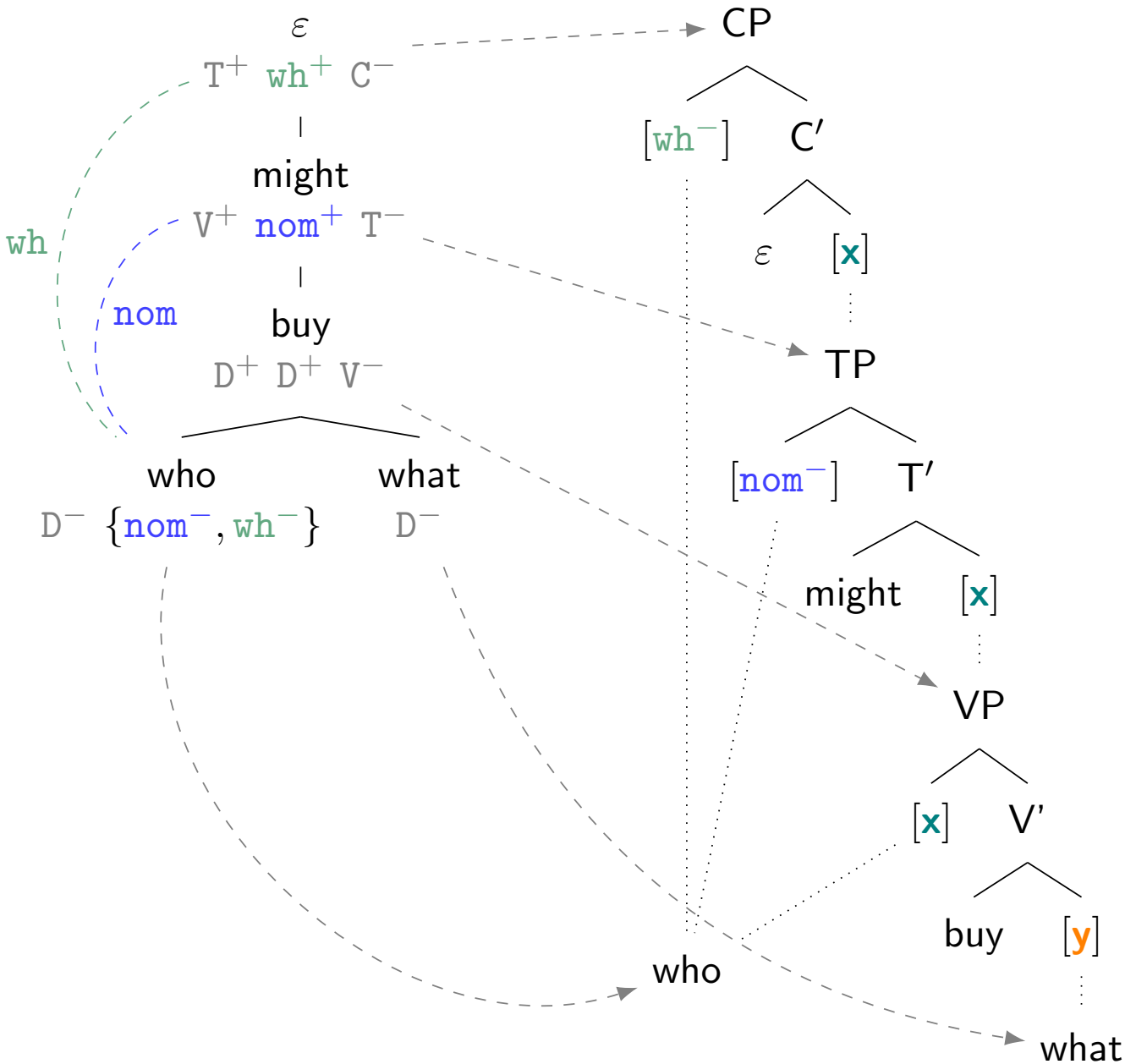
Multiple copies are straightforward



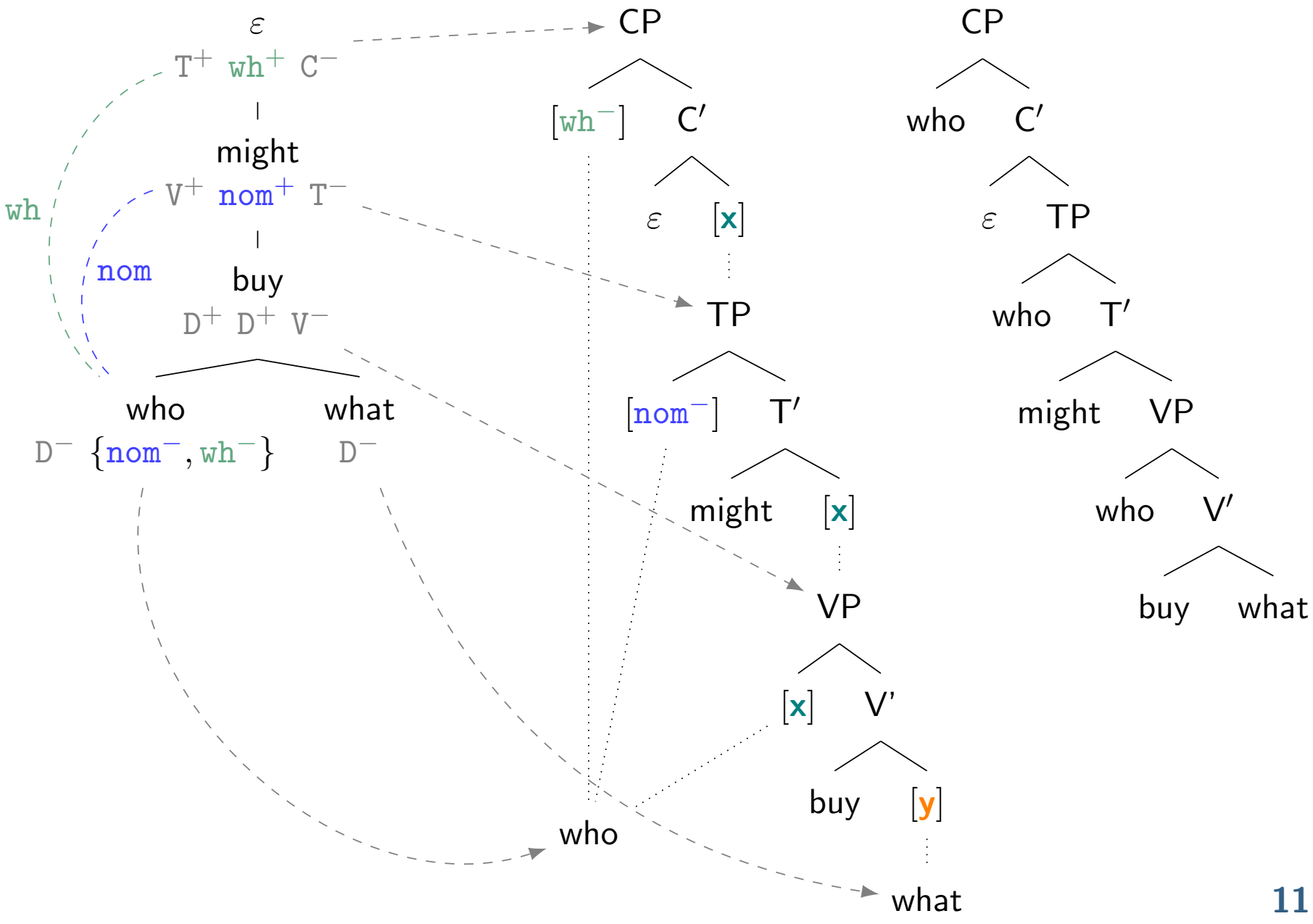
Multiple copies are straightforward



Multiple copies are straightforward



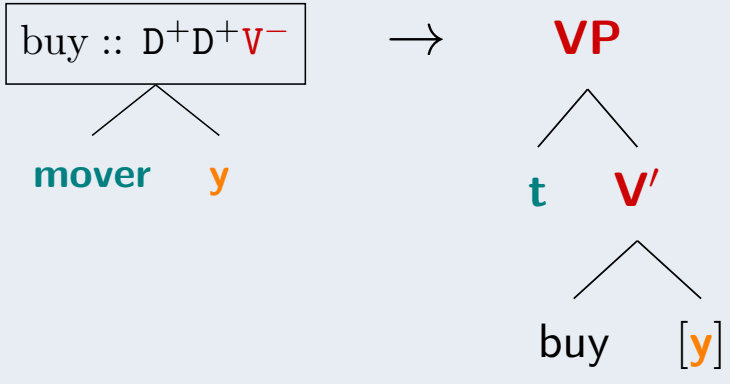
Multiple copies are straightforward



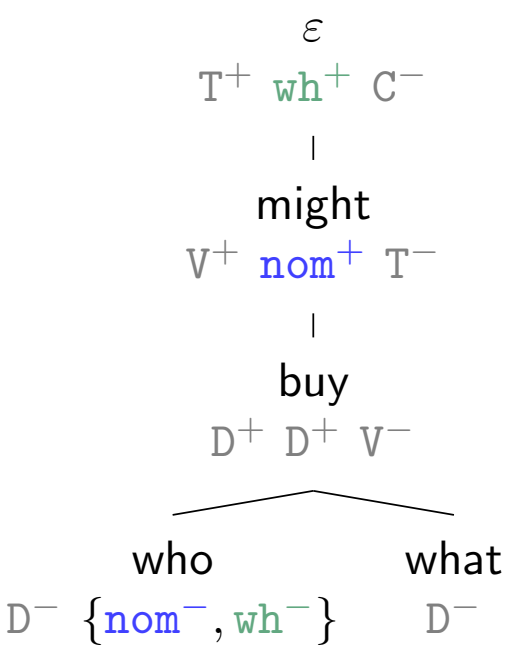
Linearization/traces: tricky

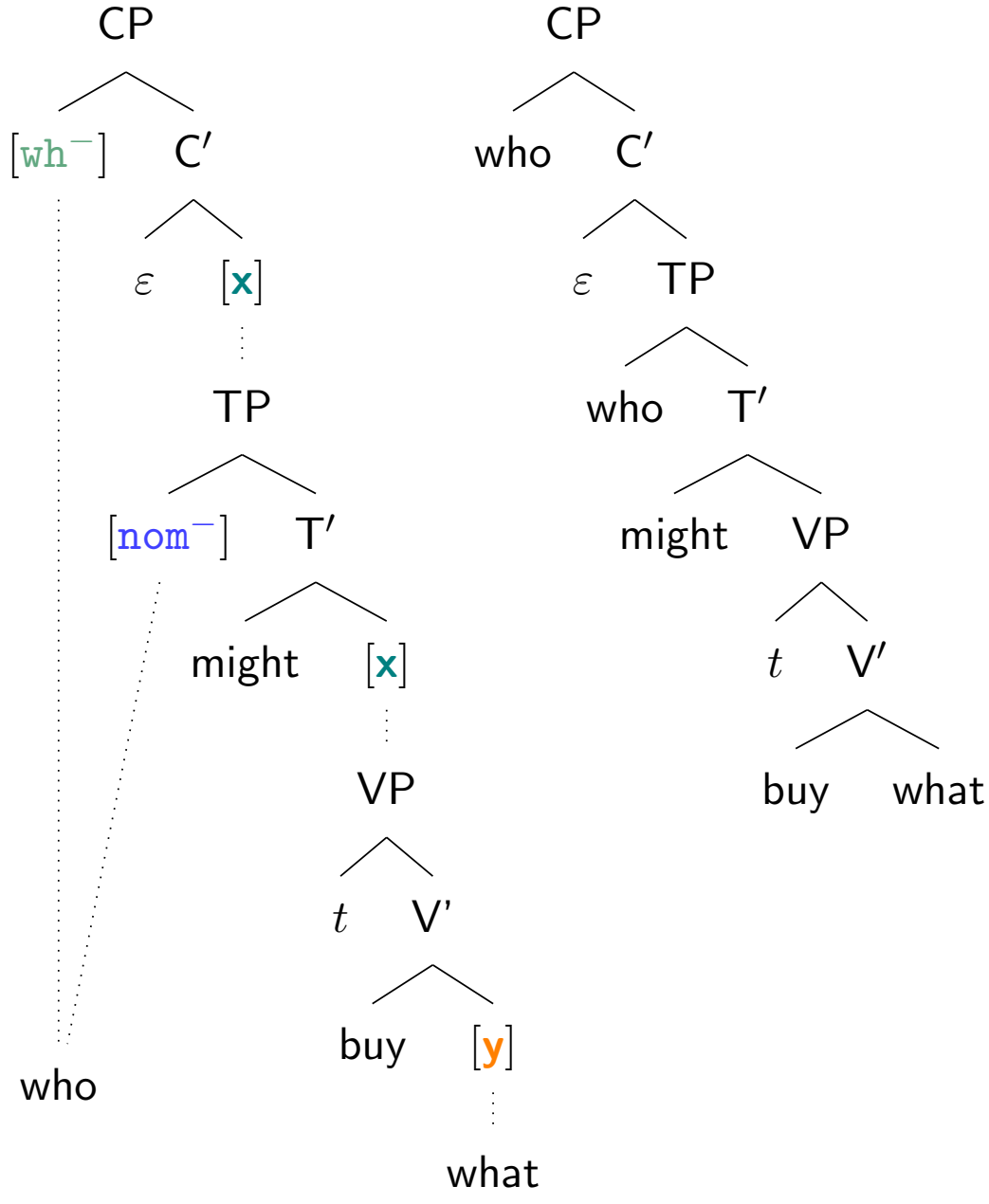
- Copies don't tell us how to pronounce the tree.
- **Traces**: unpronounced landing sites of movers

A delinking rule for base positions



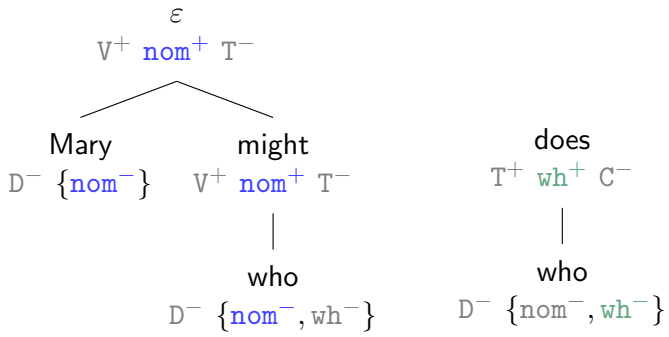
Base delinking example





Lexical predictability

- Given two landing sites **x** and **y** on different tiers, one cannot tell from the tiers whether **x** or **y** is higher.
- We **cannot distinguish final from non-final landing sites**.



Lexical predictability requirement of delinking

Delinking works only if one knows whether to

1. insert a copy, or
2. insert a trace.

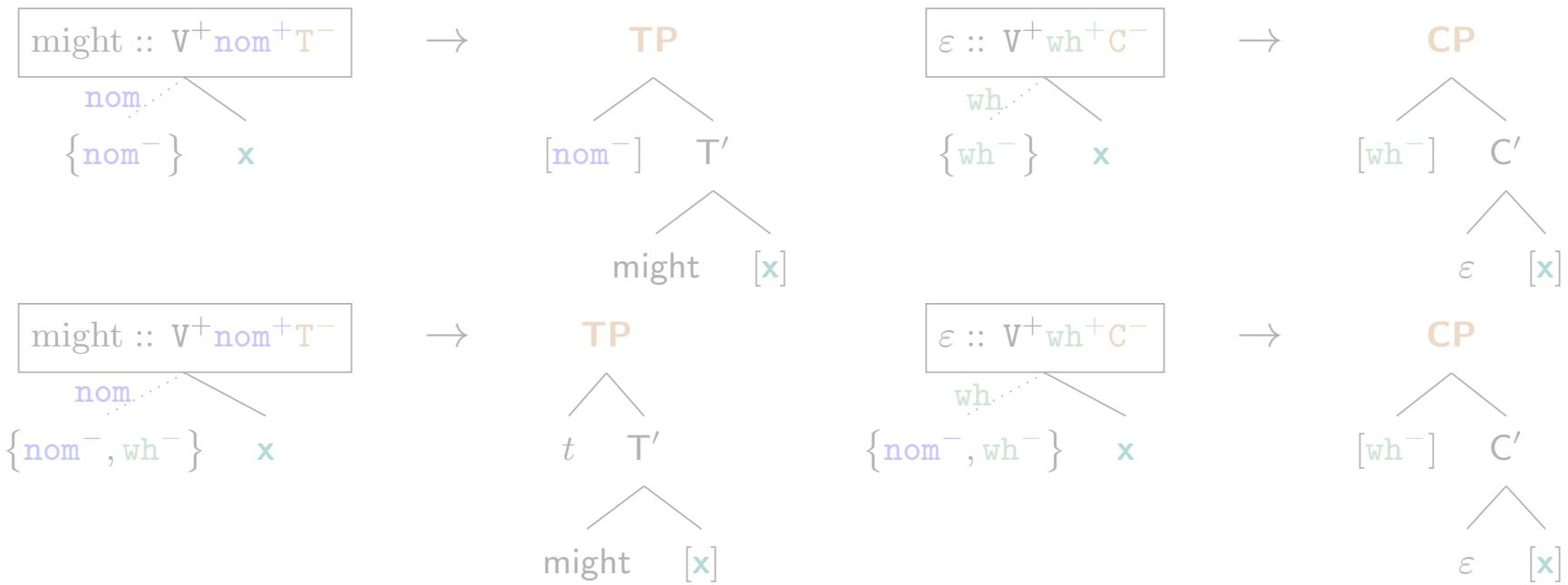
Due to the limitations of tiers,
this must be **inferable directly from the mover**.

Empirical support

- Lexical predictability holds for nom and wh movement.

Ban on Improper Movement (BoIM)

If a mover undergoes both nom and wh movement, nom movement derivationally precedes wh movement.

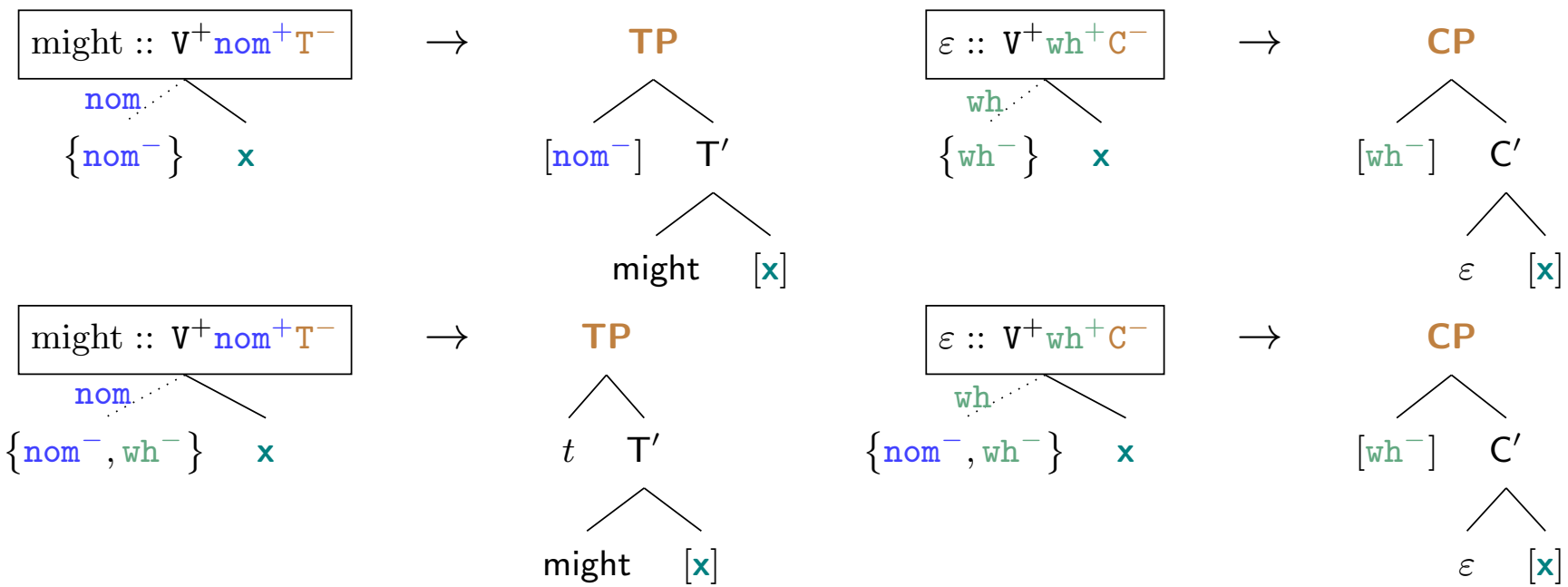


Empirical support

- Lexical predictability holds for nom and wh movement.

Ban on Improper Movement (BoIM)

If a mover undergoes both nom and wh movement, nom movement derivationally precedes wh movement.



Output-oriented Ban on Improper Movement

- BoIM is a particular instance of a more general requirement.

Output-oriented Ban on Improper Movement (OOBoIM)

- Let l be an arbitrary lexical item with $\{\mathtt{f}^-, \mathrm{g}_0^-, \ldots, \mathrm{g}_n^-\}$.
- If l's final movement step is f-movement in **some** derivation, then l's final movement step is f-movement in **all** derivations.

- Kenneth Hanson's analysis of MG corpus supports **even stronger version**:
  if $\mathtt{f} \prec \mathrm{g}$ for l in some derivation,
  then $\mathtt{f} \prec \mathrm{g}$ for l in all derivations.
- OOBoIM permits BoIM violations
  hyperraising

Kenneth Hanson

More complex movement

- **Overt/covert**
  - MGs allow for covert movement, which does not affect linearization.
  - This is fine as long as lexical predictability holds, e.g. because overt/covert distinction is on licensee features
- **Sideward movement**
  - Pretty much the same, but uses tier sibling relation instead of tier daughter relation
- **Persistent features**
  - Persistent $\mathtt{f}^-$ can check multiple $\mathtt{f}^+$
  - used for successive cycle movement
  - Persistent movement is still a TSL dependency
  - But with transductions, we get a complexity swap: we can handle traces, but not copies
  - My hunch: that's a good thing

Example with persistent features

Persistence seems to be likely to be discussed

$$\varepsilon :: \mathtt{V}^{+}\mathtt{nom}^{+}\mathtt{T}^{-}$$
$$|$$
$$\text{seems} :: \mathtt{T}^{+}\mathtt{V}^{-}$$
$$|$$
$$\text{to} :: \mathtt{V}^{+}\mathtt{snom}^{+}\mathtt{T}^{-}$$
$$|$$
$$\text{be likely} :: \mathtt{T}^{+}\mathtt{V}^{-}$$
$$|$$
$$\text{to} :: \mathtt{V}^{+}\mathtt{snom}^{+}\mathtt{T}^{-}$$
$$|$$
$$\text{be discussed} :: \mathtt{D}^{+}\mathtt{V}^{-}$$
$$|$$
$$\text{persistence} :: \mathtt{D}^{-}\,\{\mathtt{nom}^{-}\}$$

Example with persistent features

Persistence seems to be likely to be discussed

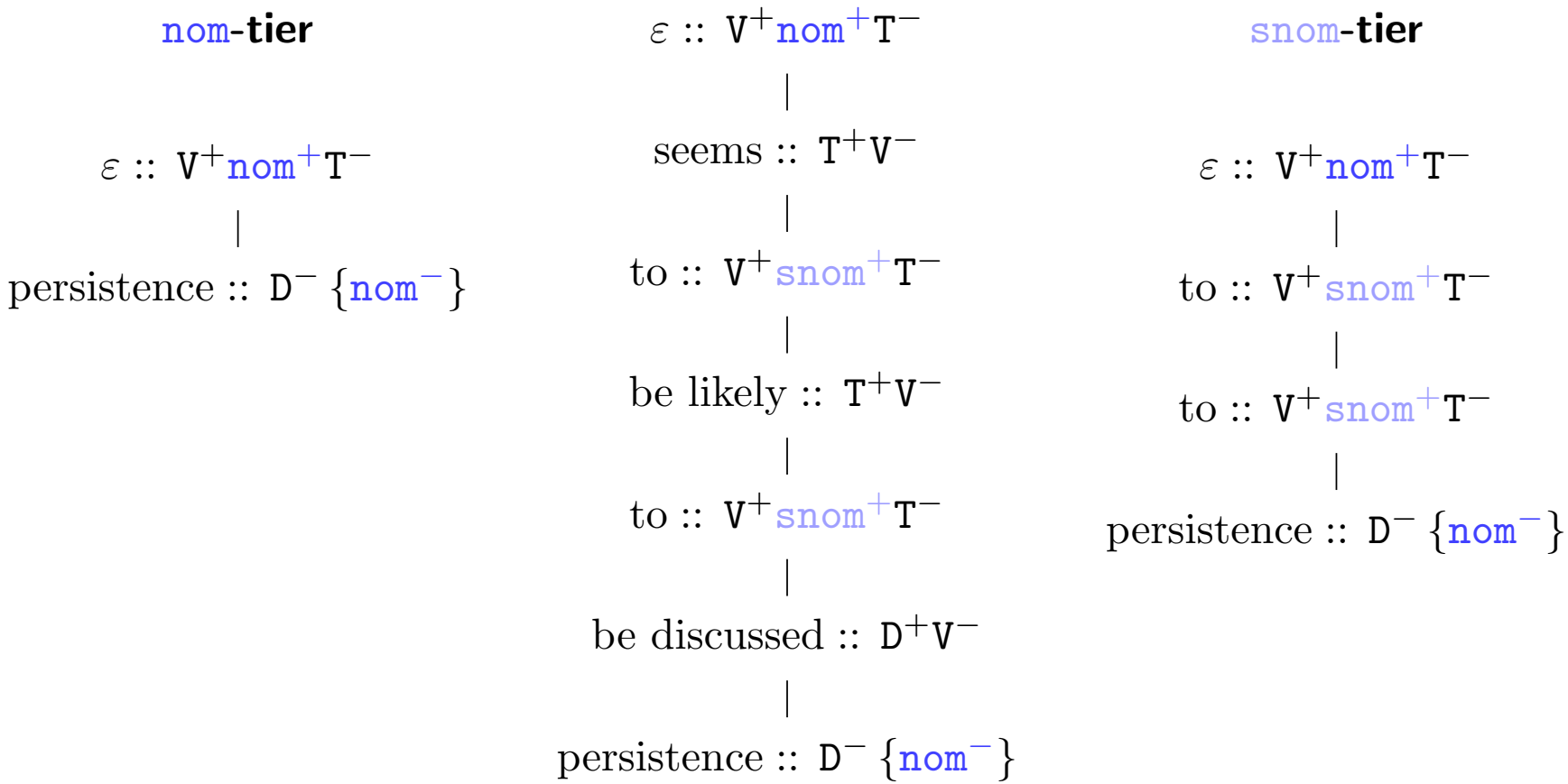
nom-tier

$\varepsilon :: \mathtt{V}^{+}\mathtt{nom}^{+}\mathtt{T}^{-}$
|
persistence :: $\mathtt{D}^{-}\,\{\mathtt{nom}^{-}\}$

$\varepsilon :: \mathtt{V}^{+}\mathtt{nom}^{+}\mathtt{T}^{-}$
|
seems :: $\mathtt{T}^{+}\mathtt{V}^{-}$
|
to :: $\mathtt{V}^{+}\mathtt{snom}^{+}\mathtt{T}^{-}$
|
be likely :: $\mathtt{T}^{+}\mathtt{V}^{-}$
|
to :: $\mathtt{V}^{+}\mathtt{snom}^{+}\mathtt{T}^{-}$
|
be discussed :: $\mathtt{D}^{+}\mathtt{V}^{-}$
|
persistence :: $\mathtt{D}^{-}\,\{\mathtt{nom}^{-}\}$

Example with persistent features

Persistence seems to be likely to be discussed



Conclusion

- Movement is both a syntactic dependency and an operation.
- In both cases the core of movement is **local over tiers**.
- Identifying mover with all landing sites (copies)
  easier than identifying output-relevant landing sites (traces)

Outlook: Bringing it all together

- TSL perspectives of all movement types
  covert, successive-cyclic, sidewards, multiple wh
- Tiers for islands, extraction morphology and path conditions
  Wolof u-chains, floating quantifiers, Germanic wh-copying
- Learning
  SL learning of Merge features, ??? for Move features

Acknowledgments

This work was supported by the National Science Foundation under Grant No. BCS-1845344.

References I

Chandlee, Jane. 2014. *Strictly local phonological processes*. Doctoral Dissertation, University of Delaware. URL `http://udspace.udel.edu/handle/19716/13374`.

Chandlee, Jane. 2022. Less is more: Reexamining assumptions through the narrow focus of subregularity. *Theoretical Linguistics* 48:205–218.

Chandlee, Jane, and Jeffrey Heinz. 2018. Strict locality and phonological maps. *Linguistic Inquiry* 49:23–60.

Graf, Thomas. 2018. Why movement comes for free once you have adjunction. In *Proceedings of CLS 53*, ed. Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit, and Miachel Tabatowski, 117–136.

Graf, Thomas. 2020. Curbing feature coding: Strictly local feature assignment. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2020*, 362–371.

Graf, Thomas. 2022a. Diving deeper into subregular syntax. *Theoretical Linguistics* 48:245–278.

Graf, Thomas. 2022b. Subregular linguistics: Bridging theoretical linguistics and formal grammar. *Theoretical Linguistics* 48:145–184.

Heinz, Jeffrey. 2018. The computational nature of phonological generalizations. In *Phonological typology*, ed. Larry Hyman and Frank Plank, Phonetics and Phonology, chapter 5, 126–195. Mouton De Gruyter.

References II

Ikawa, Shiori, Akane Ohtaka, and Adam Jardine. 2020. Quantifier-free tree transductions. In *Proceedings of the Society for Computation in Linguistics (SCiL)*, volume 3, 455–458.

Ji, Jing, and Jeffrey Heinz. 2020. Input strictly local tree transducers. In *Language and Automata Theory and Applications: 14th International Conference, LATA 2020, Milan, Italy*, volume 12038 of *LNCS*, 369–381.

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.