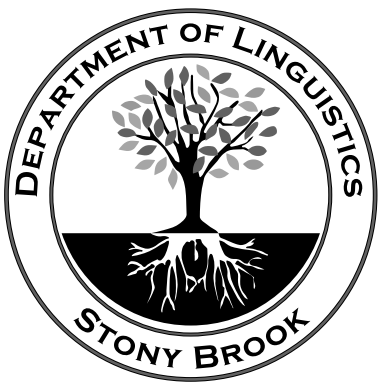
Syntactic tiers for movement and agreement

Day 2: A whirlwind tour of syntactic phenomena

Thomas Graf



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

KU Leuven Lecture Series
Decmber 5–7, 2023

A little less conversation, a little more action please

- We now have a good idea of how syntactic tiers work.
- Let's apply them to some phenomena!

Outline

Capturing variation

- Island effects vary in unexpected ways:
  - across languages
  - across lexical items
- Syntactic tiers can handle both because the tier projection can determine for each lexical item whether it projects.
- **Flip side**: need a story for why island constraints are still pretty uniform across languages

Accounting for gradience

- There's a **huge debate about gradience** in island effects.
  - **Option 1**: island constraints are purely a performance phenomenon
  - **Option 2**: islands are in the grammar, gradience is a performance phenomenon
  - **Option 3**: island are gradient constraints in the grammar
- Syntactic tiers are compatible with all three options.
- In particular, syntactic tiers make it easy to combine gradience with standard analyses.
- How? **Weighted/Probabilistic tier projection**

Gradience

- Tier projection can be made probabilistic to capture gradience (Mayer 2021; Torres et al. 2023)

Calculating gradience with probabilistic wh-tier

1. construct all possible versions of the wh-tier
2. filter out illicit tiers
3. sum up probabilities of remaining tiers

Example calculation

(1) * Who did Bill complain after having a meeting with ⟨who⟩ because he's always angry.

Item	Probability
any $\texttt{wh}^+$	1
any $\texttt{wh}^-$	1
because	.9
after	.8

- **Possible tiers**
  1. Project $\texttt{wh}^+$, $\texttt{wh}^-$, because, after: .72, illicit
  2. Project $\texttt{wh}^+$, $\texttt{wh}^-$, because: .18, licit
  3. Project $\texttt{wh}^+$, $\texttt{wh}^-$, after: .08, illicit
  4. Project $\texttt{wh}^+$, $\texttt{wh}^-$: .02, licit
- Overall probability: .2

Probe horizons VS the Ban on Improper Movement

- Tier-based story for islands also captures probe horizons (Keine 2016, 2019)

Example

- **Probe horizons**: A-movement features cannot probe into CPs
- **Syntactic tiers**: all C-heads project onto all A-movement tiers

- Syntactic tiers also show us that probe horizons are preferable to the Ban on Improper Movement (BoIM).

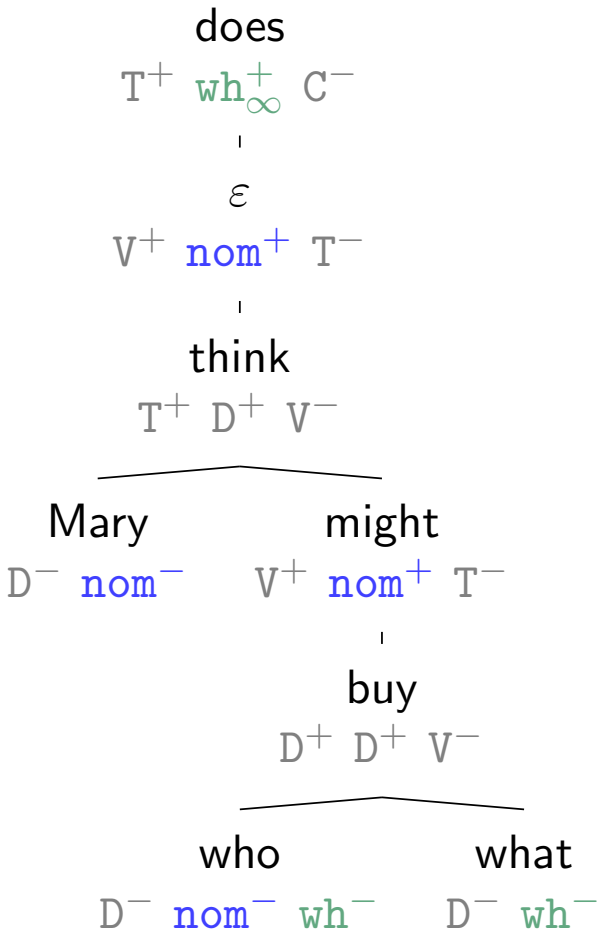
Ruling out improper movement via. . .

Probe horizons : no new tiers needed

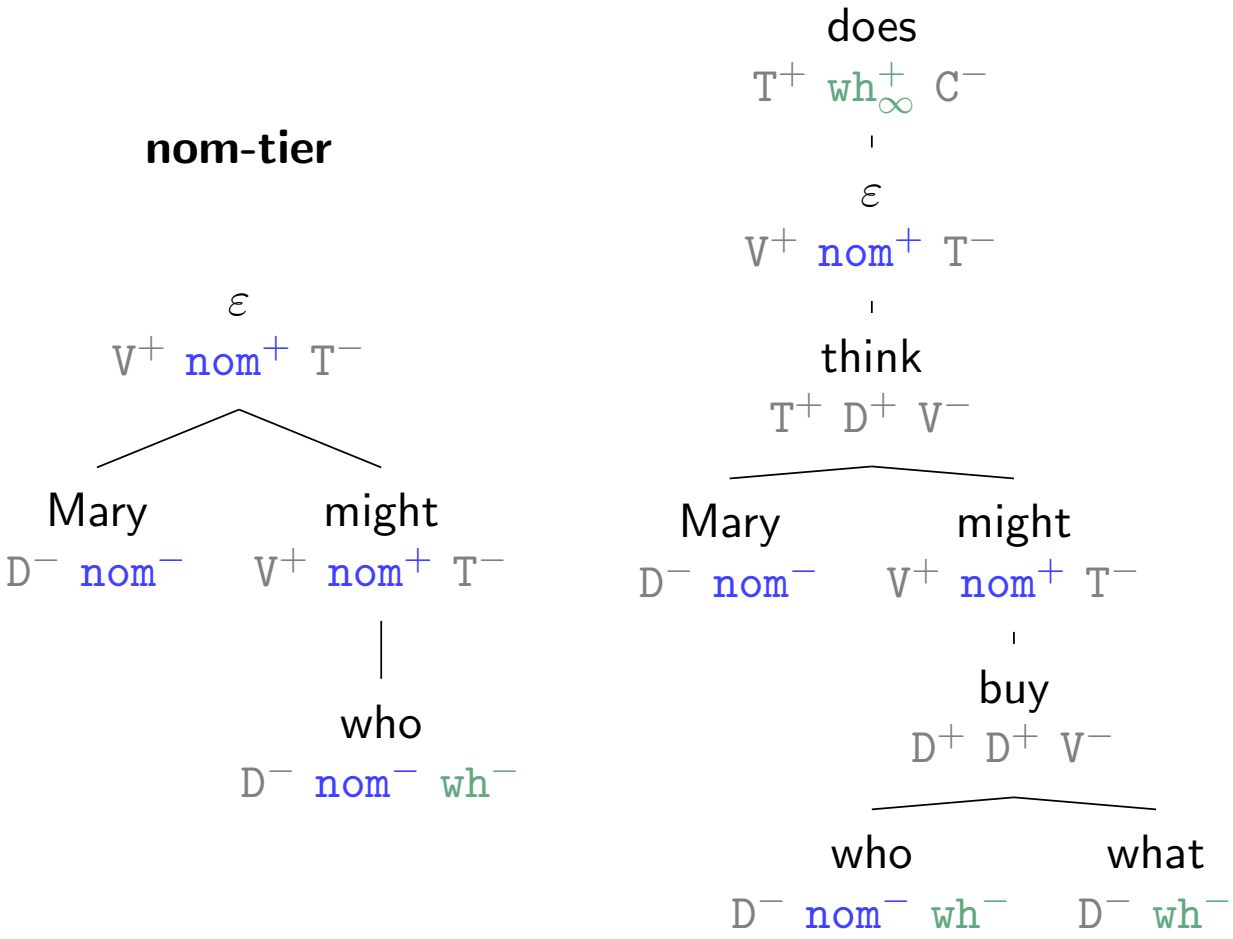
BoIM : up to $\frac{n^2-n}{2}$ new tiers

(where n is the number of movement types)

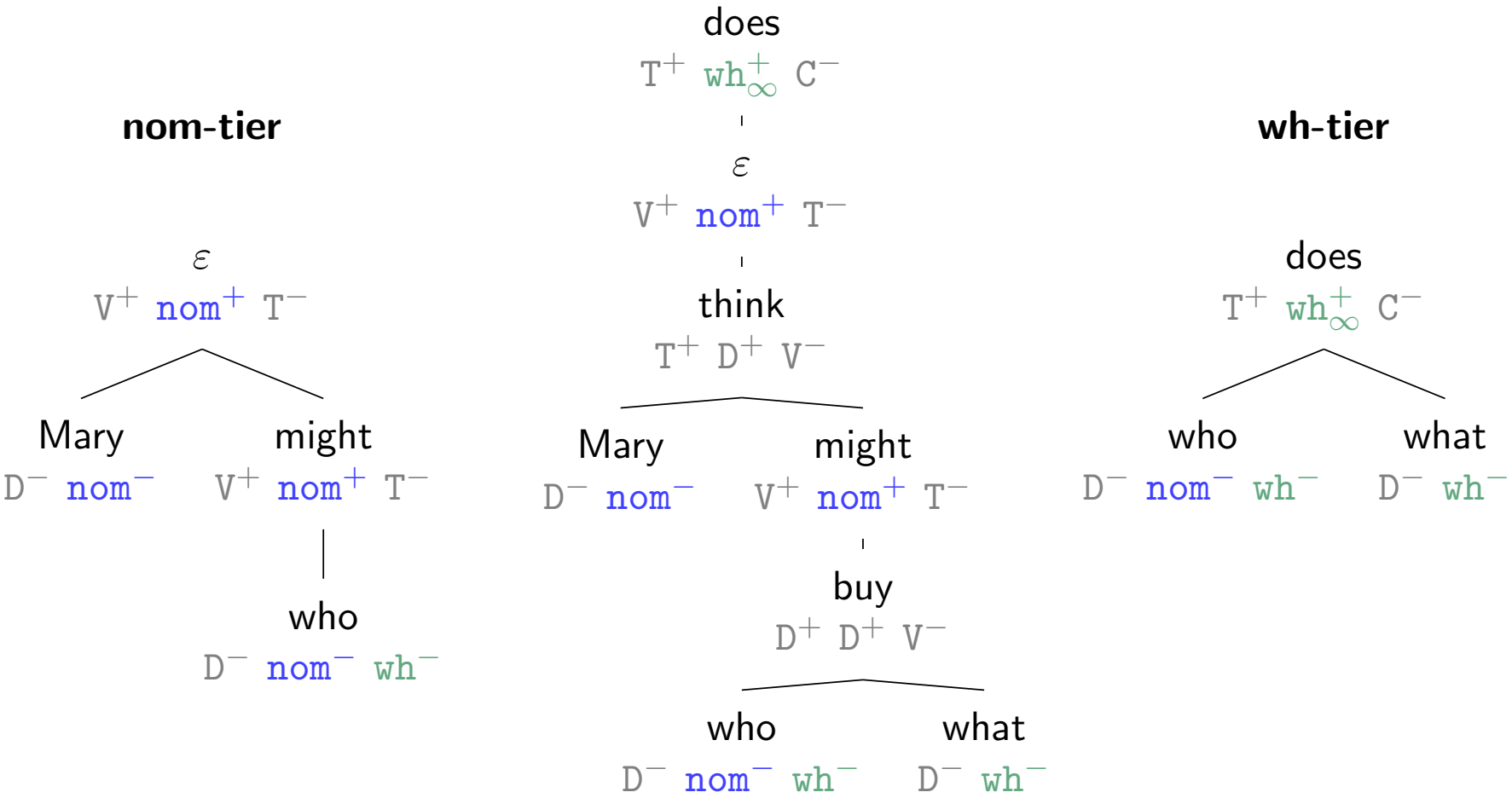
Multiple wh-movement is also TSL



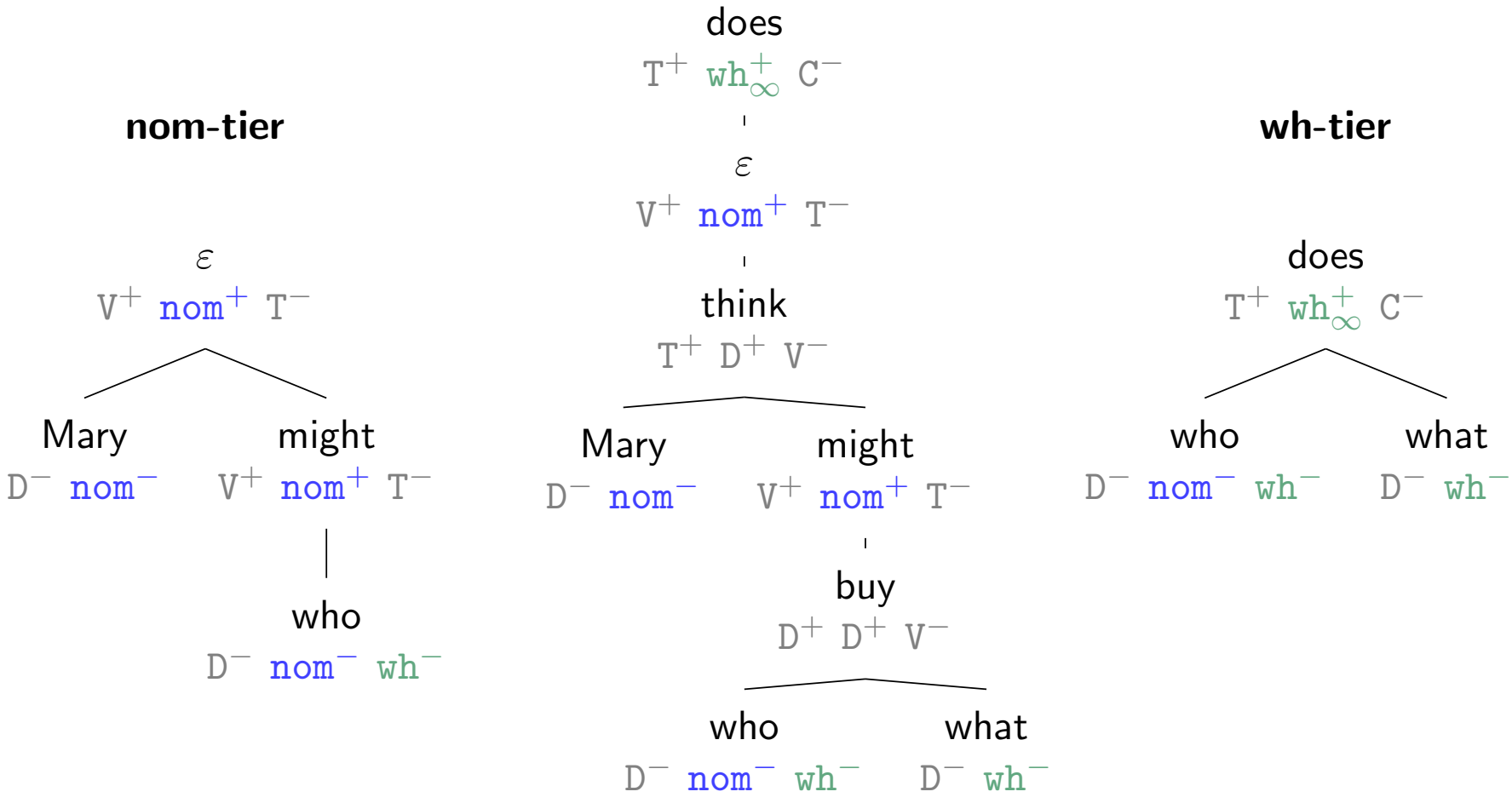
Multiple wh-movement is also TSL



Multiple wh-movement is also TSL



Multiple wh-movement is also TSL



Persistent feature checking as a constraint on tiers

On each f-tier, every $\mathtt{f}^{+}_{\infty}$ has **at least 1** $\mathtt{f}^{-}$ among its daughters.

Multiple wh-movement is not special

From the TSL-perspective, multiple wh-movement is already part of standard movement:

- Standard movement: every $\texttt{f}^{+}$ has exactly one $\texttt{f}^{-}$ daughter
- **TSL:** "exactly one" = "at least one" + "at most one"

multiple wh-movement standard movement - "at most one"

standard movement multiple wh-movement + "at most one"

Work to be done...

That's a nice starting point, but we need to account for...

- why there is no multiple nom-movement
- linearization,
- superiority effects/c-command,
- multiple movers with distinct targets,
- the 2-by-2 typology of wh-movement

Multiple wh	**Superiority**	**Example language**
-	-	Spanish
-	+	English
+	-	Russian
+	+	Bulgarian

Superiority effects

- Some languages have a limited superiority effect: the highest wh-mover must be first in the wh-cluster
- This is just a distributional constraint on $\texttt{wh}$ VS $\texttt{wh}_{\infty}$
- That would immediately tell us why superiority and multiple-wh are independent parameters

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each f-tier

1. Every $\texttt{f}^{+}$ has exactly 1 $\texttt{f}^{-}$ among its daughters.
2. Every $\texttt{f}^{-}$ has $\texttt{f}^{+}$ as its mother.

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each `f`-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each f-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

	Positive on f-tier	**Negative** on f-tier
$\mathtt{f}^{+}$	True	False
$\mathtt{f}^{-}$	False	True

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each f-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

	Positive on f-tier	**Negative** on f-tier
$\mathtt{f}^{+}$	True	False
$\mathtt{f}^{-}$	False	True
	False	False

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each f-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

	Positive on f-tier	**Negative** on f-tier
$\mathtt{f}^{+}$	True	False
$\mathtt{f}^{-}$	False	True
	False	False
	True	True

A 2-by-2 typology of tier nodes

Feature checking as a constraint on each f-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

	Positive on f-tier	**Negative** on f-tier
$\mathtt{f}^{+}$	True	False
$\mathtt{f}^{-}$	False	True
islands	False	False
	True	True

A 2-by-2 typology of tier nodes

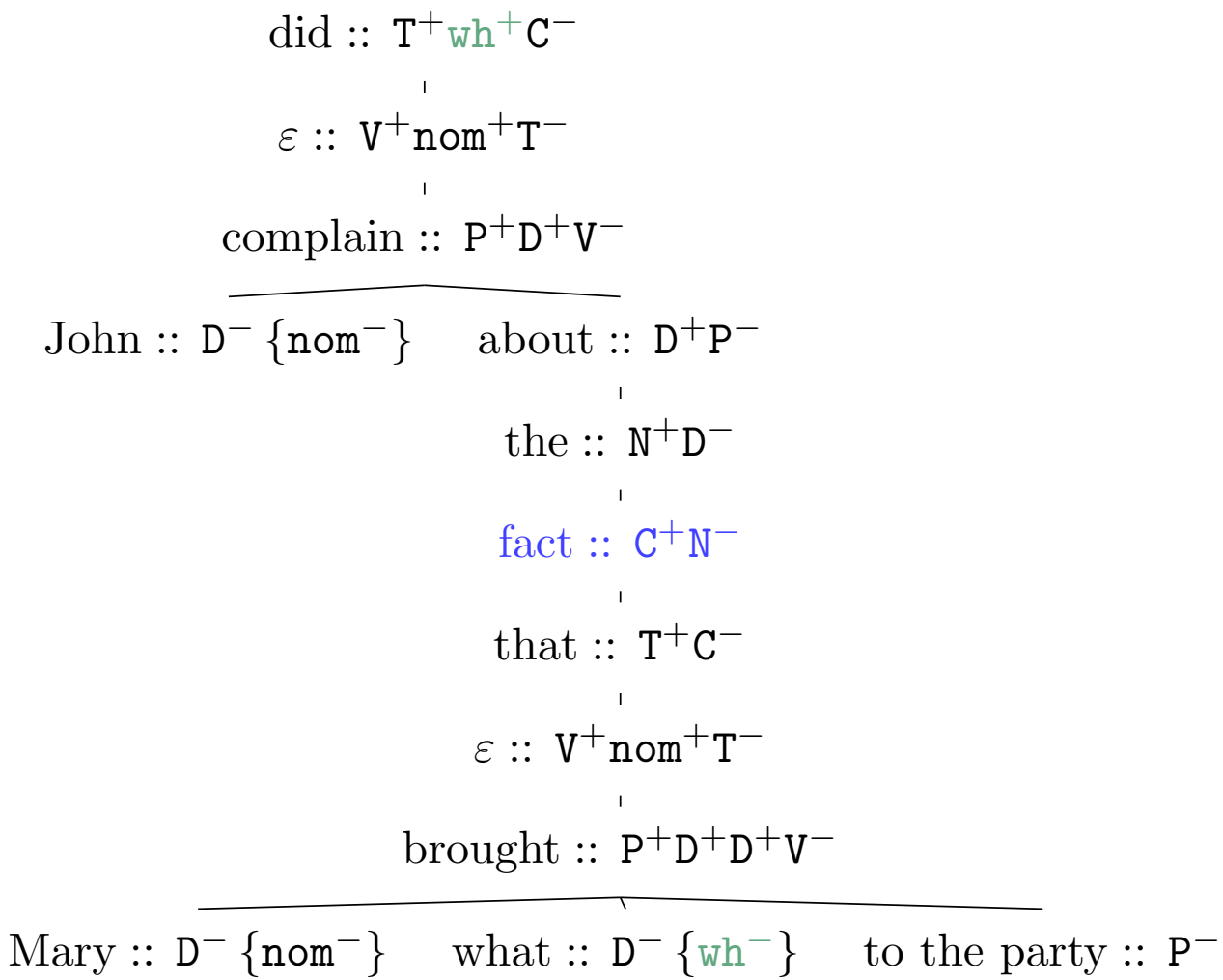
Feature checking as a constraint on each f-tier

1. Every **Positive** has exactly 1 **Negative** among its daughters.
2. Every **Negative** has **Positive** as its mother.

	Positive on f-tier	**Negative** on f-tier
$\mathtt{f}^{+}$	True	False
$\mathtt{f}^{-}$	False	True
islands	False	False
wh-agreement	True	True

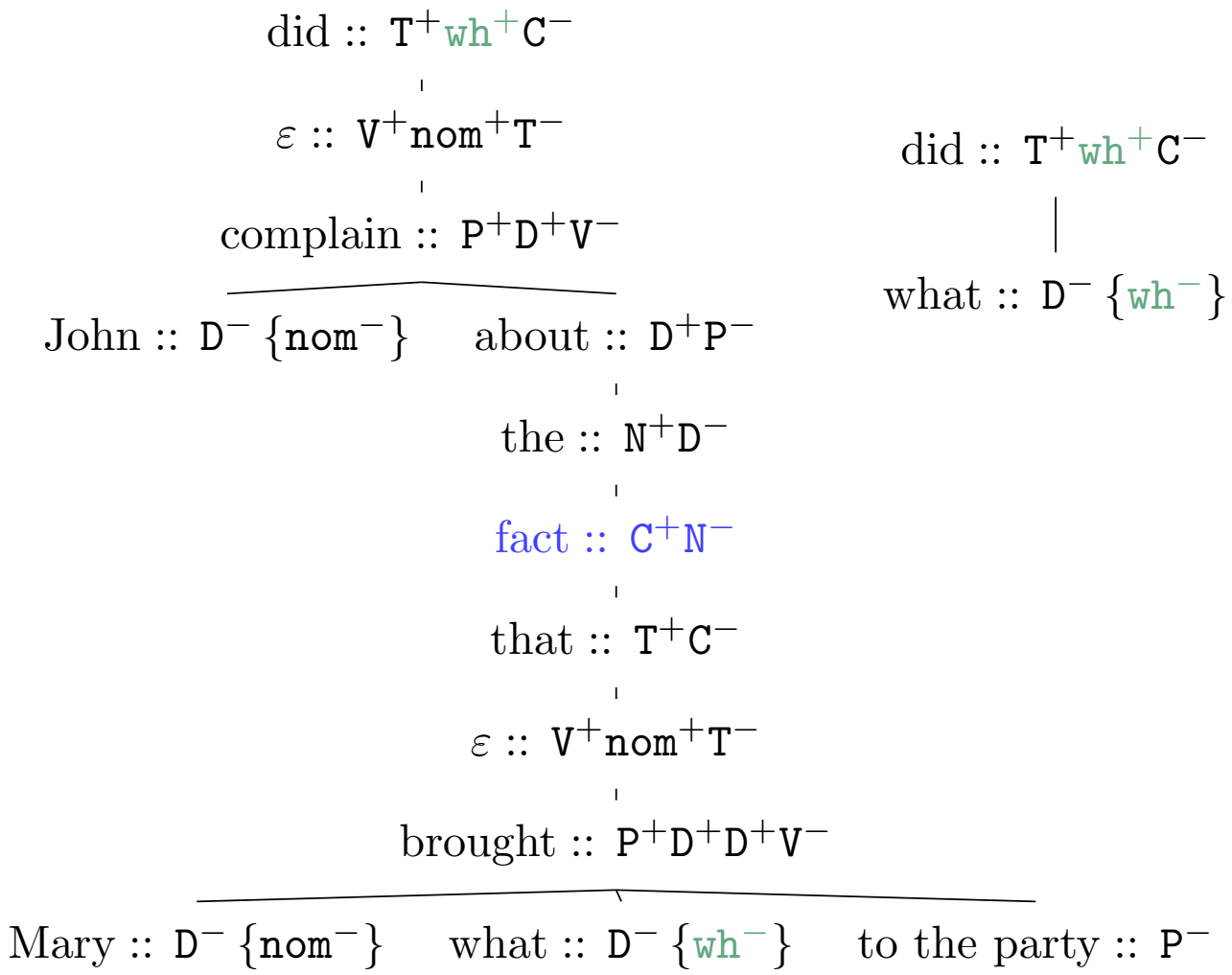
Islands as tier blockers (neither positive nor negative)

(2) * What did John complain about the fact that Mary brought ⟨what⟩ to the party?



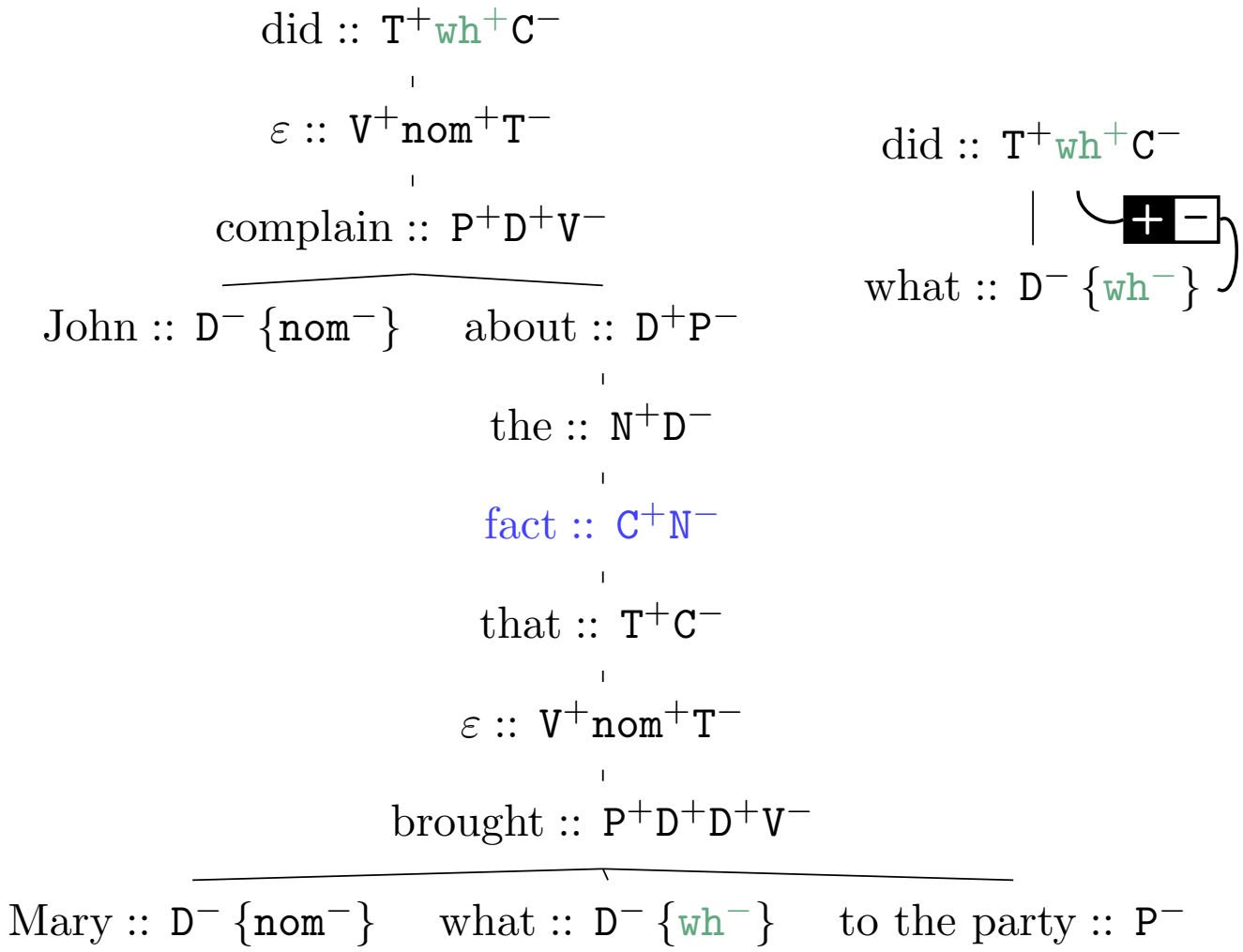
Islands as tier blockers (neither positive nor negative)

(2) * What did John complain about the fact that Mary brought ⟨what⟩ to the party?



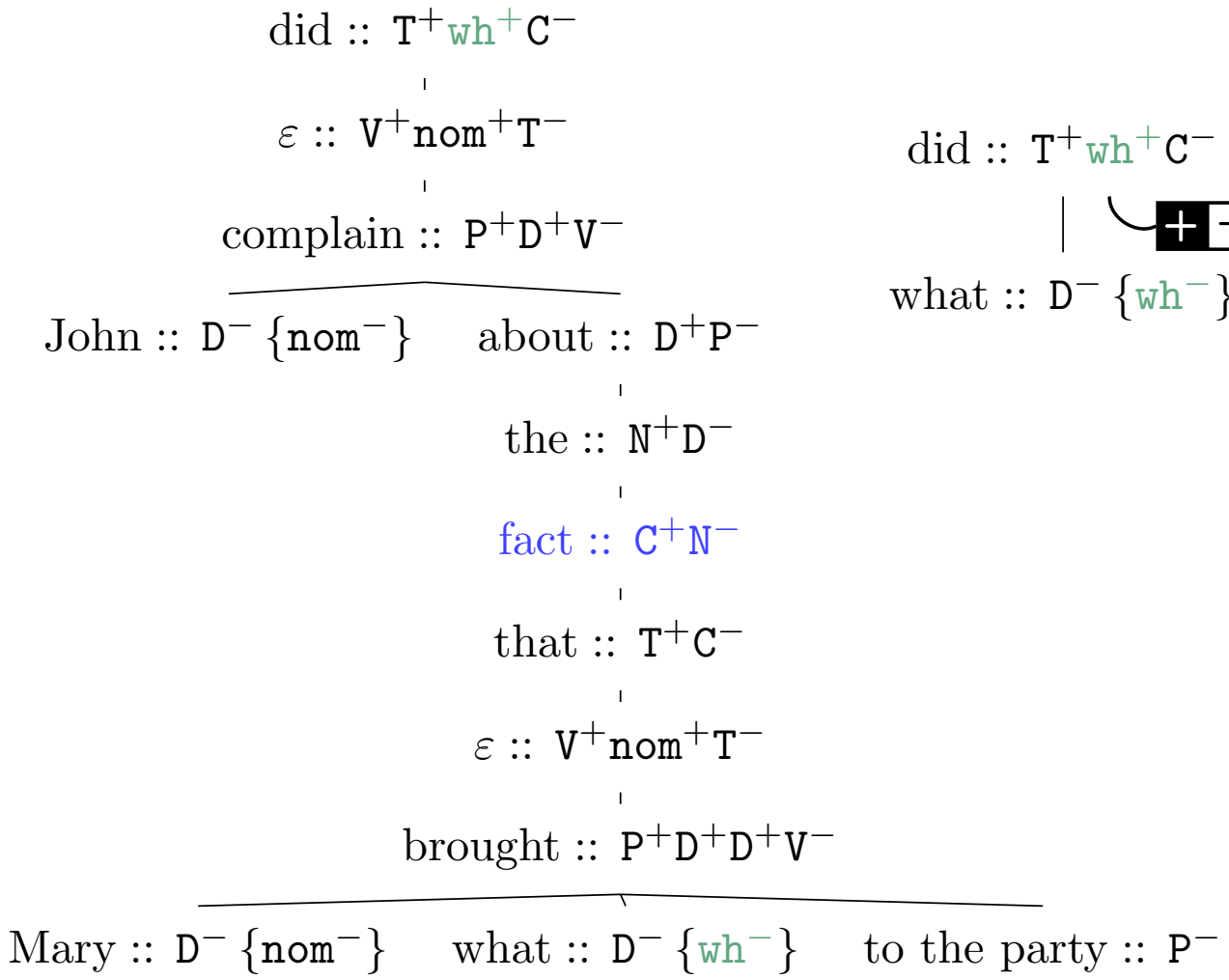
Islands as tier blockers (neither positive nor negative)

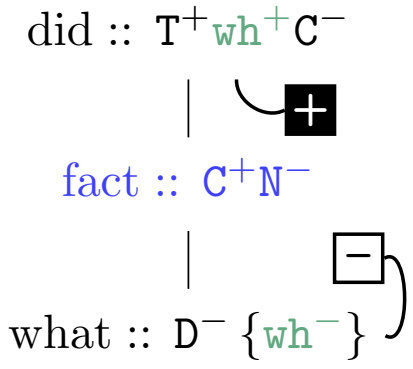
(2) * What did John complain about the fact that Mary brought ⟨what⟩ to the party?



Islands as tier blockers (neither positive nor negative)

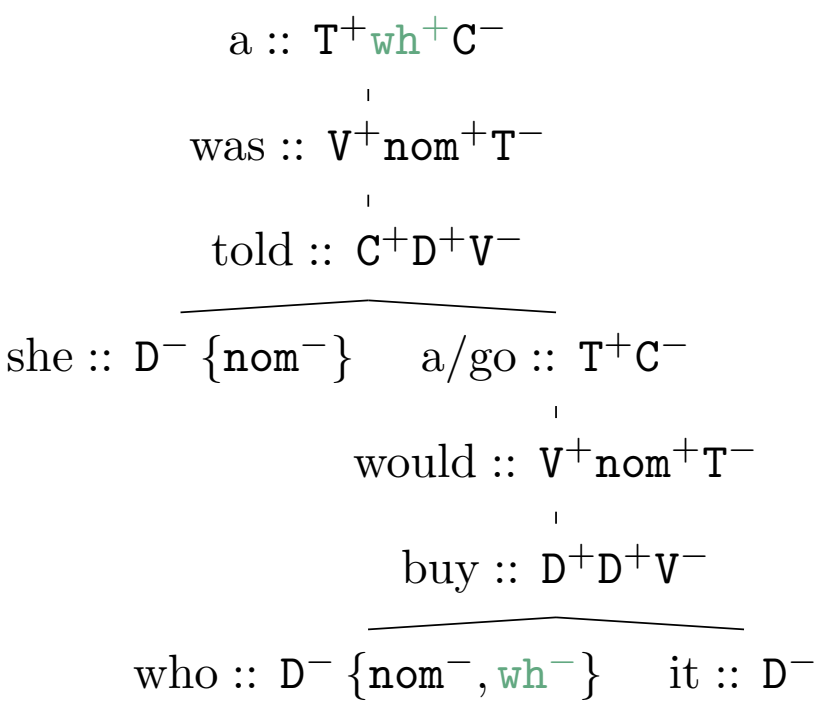
(2) * What did John complain about the fact that Mary brought ⟨what⟩ to the party?





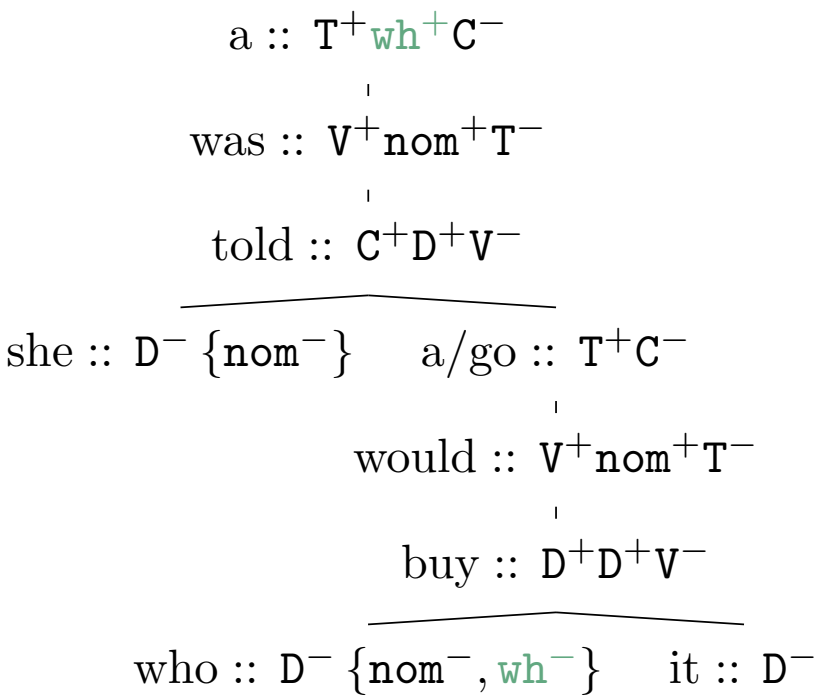
Wh-agreement via tier conduits (positive & negative)

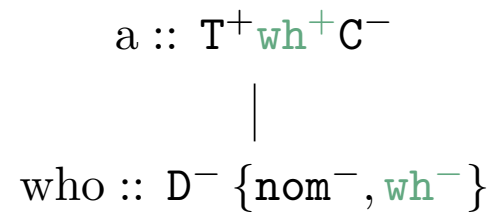
(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)



Wh-agreement via tier conduits (positive & negative)

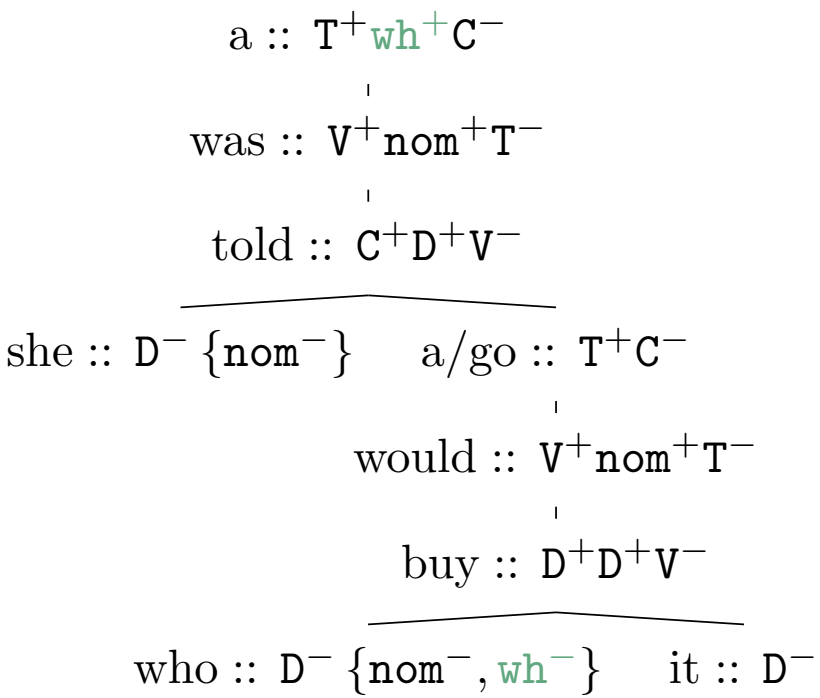
(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)

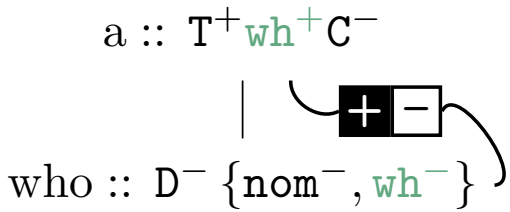




Wh-agreement via tier conduits (positive & negative)

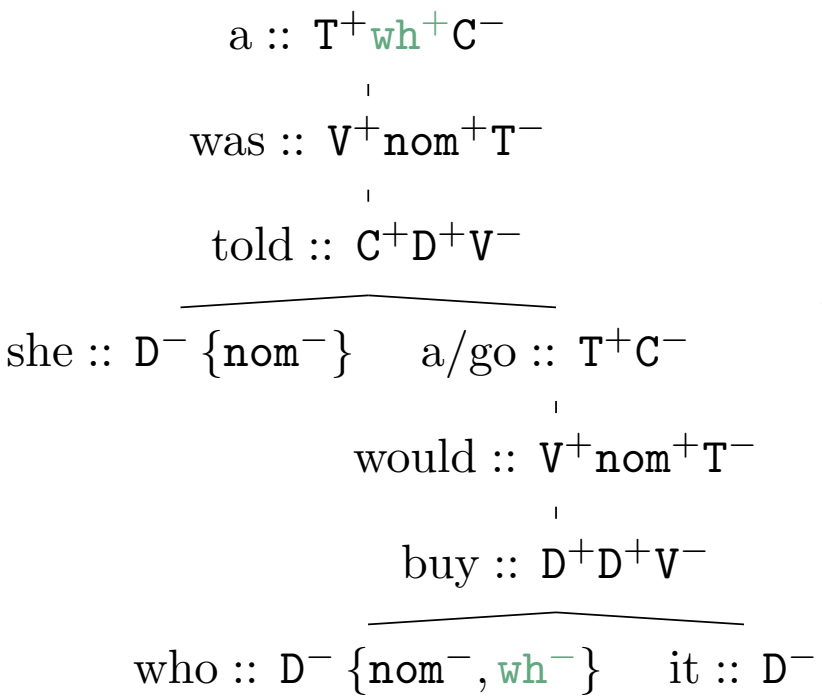
(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)

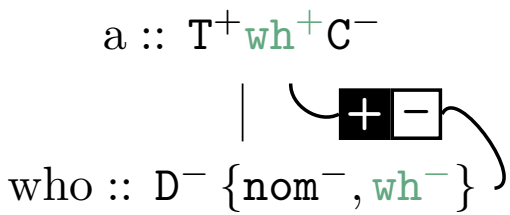


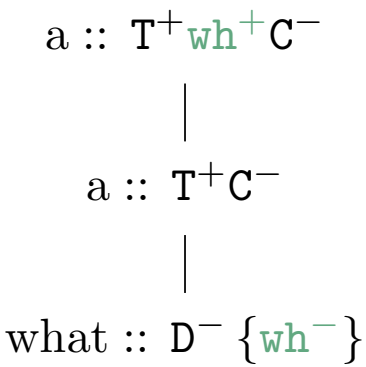


Wh-agreement via tier conduits (positive & negative)

(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)

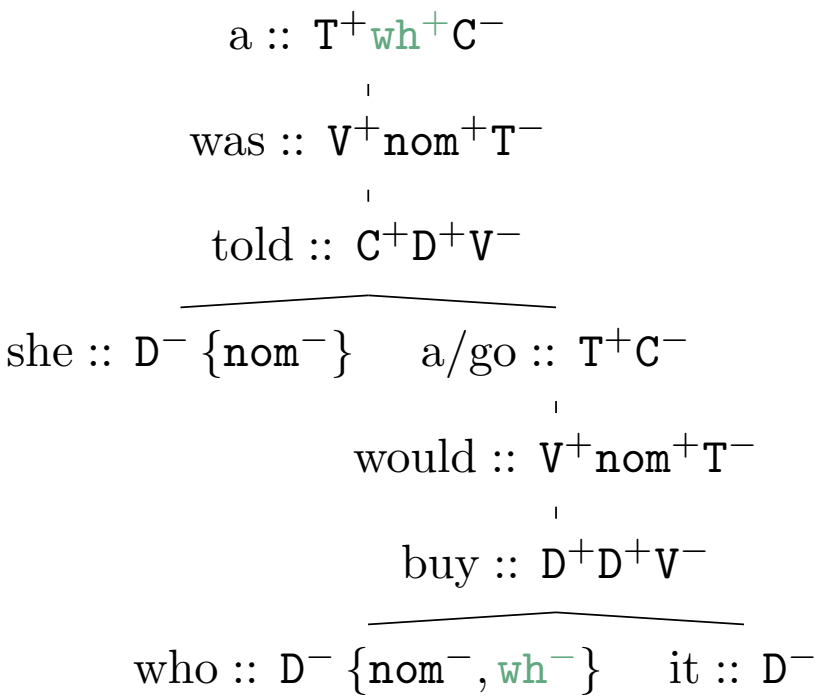


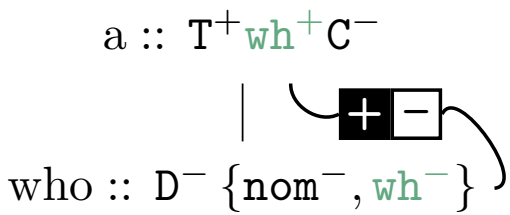


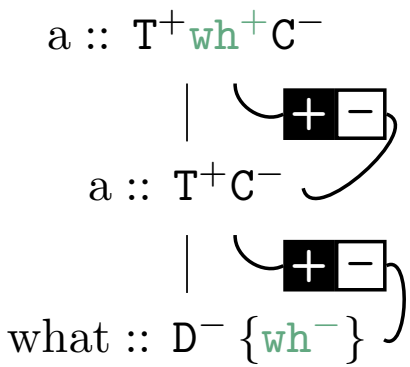


Wh-agreement via tier conduits (positive & negative)

(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)

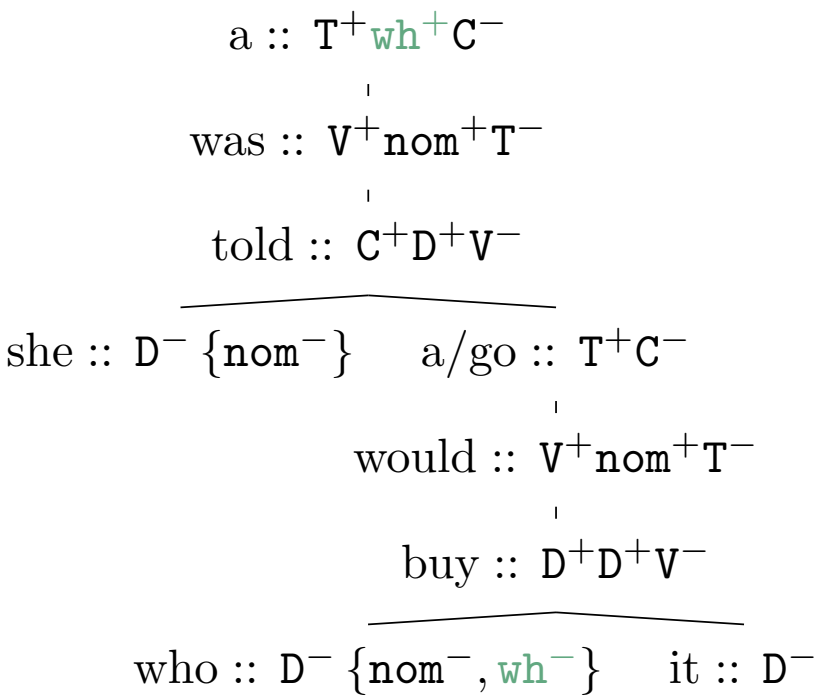


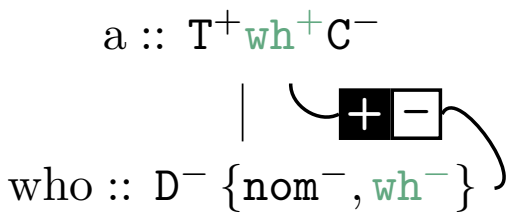


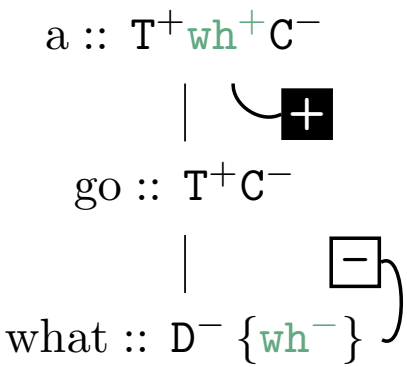


Wh-agreement via tier conduits (positive & negative)

(3) Cé **a**/*go dúradh léithi **a**/*go cheannódh é?
who C-wh/C was-said with-her C-wh/C would-buy it
'Who was she told would buy it?'
(Irish, McCloskey 2001:p.94)







Extraction morphology: An example from Wolof u-chains

Extraction morphology

morphology conditioned by the presence of movement

- **Wolof u-chains** (Niger-Congo)
  - (covert) wh-phrase moves to matrix Spec,CP
  - highest C-head *u* along the movement path must agree with wh-phrase in Class
  - intermediate C-heads may agree with wh-phrase in Class

(4) [ε **k**-u Kumba wax [ne **k**/**l**-u Isaa defe [
Q CL-u Kumba say FRC CL/EXPL-u Isaa think
ne **k**/**l**-u Maryam dóór t]]]
FRC CL/EXPL-u Maryam hit

'Who did Kumba say that Isaa thought that Maryam hit?'
(Torrence 2012:1171)

Why care about extraction morphology?

- **Highly relevant**
  poster child for **successive cyclic movement**
- **Good case study**
  data is robust, varied, typologically diverse

A typology of extraction morphology

based on Georgi (2017), a terrific paper

1. What is agreeing?
   - T **T**arget of movement (Wolof u-chains)
   - O some **O**ther head in the clause of the landing site (Duala *no*-marking on T)
2. What clauses display agreement?
   - F the clause of the **F**inal landing site (Chamorro C-agreement)
   - N clauses with **N**on-final landing sites (Kiitharaka focus marking on verb)
3. agreement is. . .
   - X mandatory for X
   - (X) optional for X

Example

Wolof u-chains are TF(N).

Extraction morphology and tiers

- Tiers work well for extraction morphology.
- The various options fall out nicely for T-patterns.

Central idea

- treat it as a distribution problem (not mapping/spell-out)
- use tiers to ensure correct distribution of agreeing forms

Two simplifications due to data gaps

- exactly one A′-mover per sentence
- only one type of A′-movement per sentence
  (no *which car that John bought did we trash*)

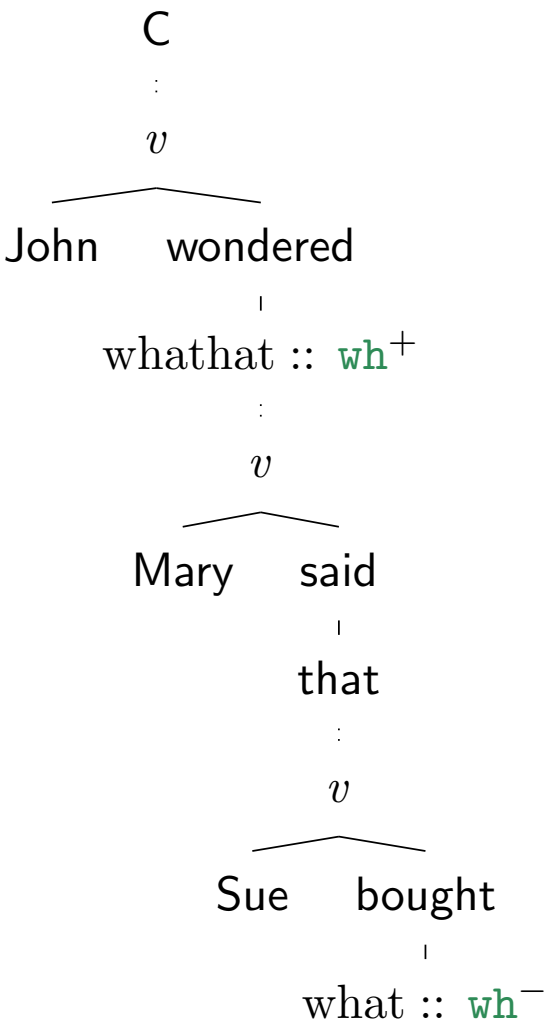
Toy example 1: TF English *whathat*

- Suppose English had a special C-head *whathat*, which has to be the final landing site of some wh-mover.

Analysis: Lexical accident

The lexicon happens to be such that

- *whathat* must carry $\texttt{wh}^{+}$,
- *that* must not carry $\texttt{wh}^{+}$,
- empty C must not carry $\texttt{wh}^{+}$.

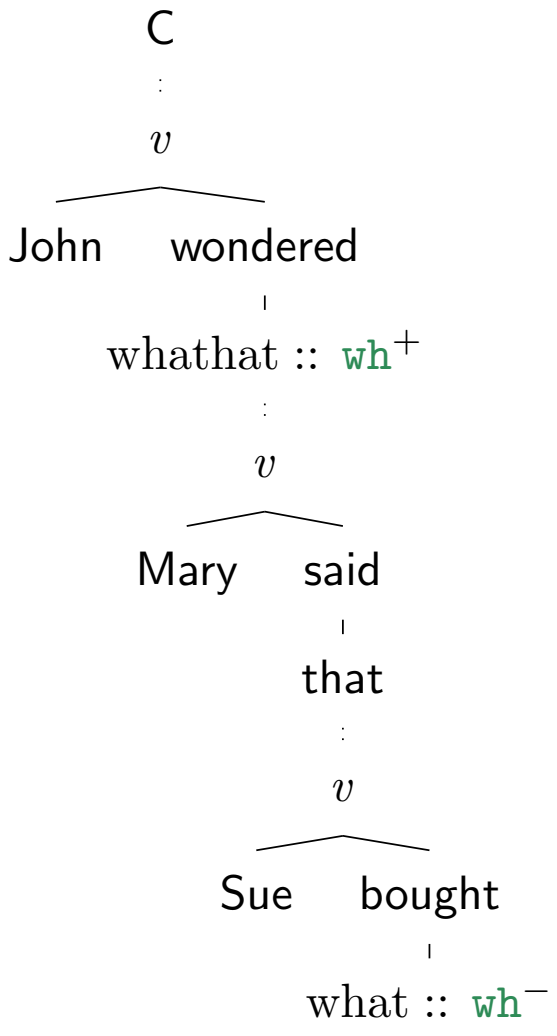


Toy example 2: TF English *whathat/whothat*

- Suppose English also had *whothat*,
  and targeted C must agree in animacy with the wh-mover.

Lexical accident + tier match

- $\texttt{wh}^{+}$ only on *whathat* and *whothat*
- on wh-tier, $\texttt{wh}^{+}$-head must match animacy of $\texttt{wh}^{-}$-daughter
  wh-tier: $\text{Xthat} :: \texttt{wh}^{+}\ \text{X} :: \texttt{wh}^{-}$



Toy example 2: TF English *whathat/whothat*

- Suppose English also had *whothat*,
  and targeted C must agree in animacy with the wh-mover.

wh-tier

whathat :: $\texttt{wh}^+$

|

what :: $\texttt{wh}^-$

Lexical accident + tier match

- $\texttt{wh}^+$ only on *whathat* and *whothat*
- on wh-tier, $\texttt{wh}^+$-head must match animacy of $\texttt{wh}^-$-daughter
  wh-tier: Xthat :: $\texttt{wh}^+$ X :: $\texttt{wh}^-$

Toy example 3: TFN English *whathat/whothat*

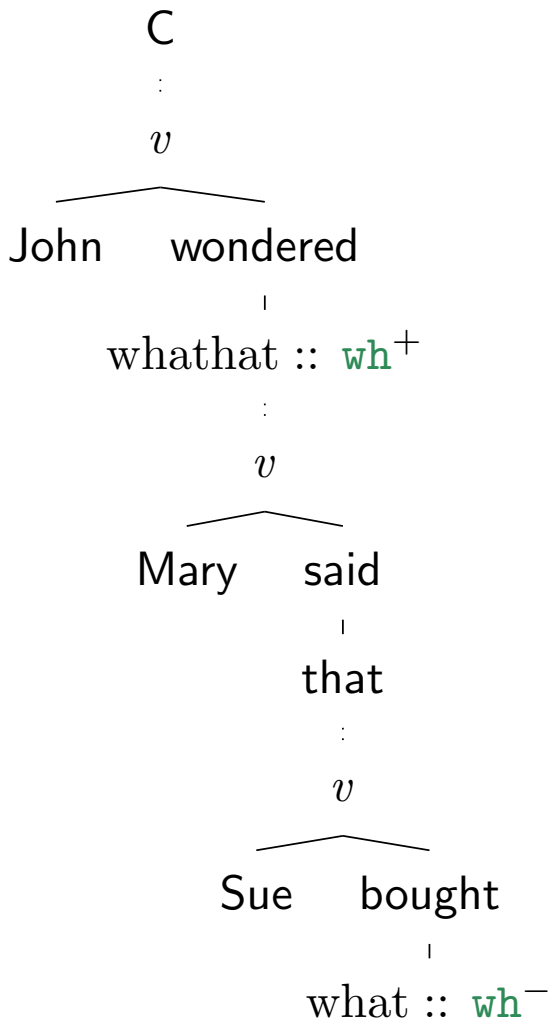
- Suppose every C-head along a wh-movement path has to be *whathat/whothat*

Lexical accident + EM-tier

- $\texttt{wh}^+$ only on *whathat* and *whothat*
- **EM-tier**
  - project all $\texttt{wh}^+$, all $\texttt{wh}^-$, and all C-heads
  - every *Xthat* must have daughter with *Xthat* (no $\texttt{wh}^+$) or matching $\texttt{wh}^-$
  - non-agreeing C above *Xthat*: *Xthat* must carry $\texttt{wh}^+$

EM-tier: Xthat :: $\texttt{wh}^+$ Xthat*

X :: $\texttt{wh}^-$



Toy example 3: TFN English *whathat/whothat*

- Suppose every C-head along a wh-movement path has to be *whathat/whothat*

wh-tier

whathat :: wh^+

|

what :: wh^-

Lexical accident + EM-tier

- wh^+ only on *whathat* and *whothat*
- **EM-tier**
  - project all wh^+, all wh^-, and all C-heads
  - every *Xthat* must have daughter with *Xthat* (no wh^+) or matching wh^-
  - non-agreeing C above *Xthat*: *Xthat* must carry wh^+

EM-tier: Xthat :: wh^+ Xthat*

X :: wh^-

Toy example 3: TFN English *whathat/whothat*

- Suppose every C-head along a wh-movement path has to be *whathat/whothat*

Lexical accident + EM-tier

- $\texttt{wh}^+$ only on *whathat* and *whothat*
- **EM-tier**
  - project all $\texttt{wh}^+$, all $\texttt{wh}^-$, and all C-heads
  - every *Xthat* must have daughter with *Xthat* (no $\texttt{wh}^+$) or matching $\texttt{wh}^-$
  - non-agreeing C above *Xthat*: *Xthat* must carry $\texttt{wh}^+$

EM-tier: $\text{Xthat} :: \texttt{wh}^+\ \text{Xthat}^*$
$\text{X} :: \texttt{wh}^-$

N-tier

C
|
whathat :: $\texttt{wh}^+$
|
that
|
what :: $\texttt{wh}^-$

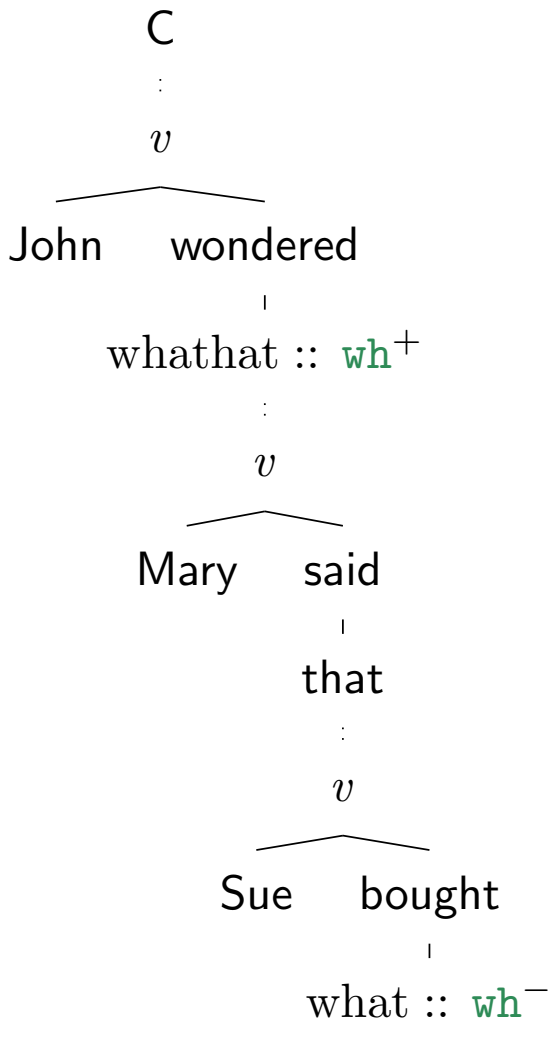
Toy example 4: T(F)(N) English *whathat/whothat*

- Suppose as before, except that *Xthat* is optional

No accident + EM-tier with Xthat only

- $\texttt{wh}^+$ on *Xthat* **or that**
- **EM-tier**
  - project all $\texttt{wh}^+$, all $\texttt{wh}^-$, and all C-heads **that are Xthat**
  - every *Xthat* must have daughter with *Xthat* (no $\texttt{wh}^+$) or matching $\texttt{wh}^-$
  - non-agreeing C above *Xthat*: *Xthat* must carry $\texttt{wh}^+$

EM-tier: $(\text{X})\text{that} :: \texttt{wh}^+ \ \text{Xthat}^*$
$\text{X} :: \texttt{wh}^-$



Toy example 4: T(F)(N) English *whathat/whothat*

- Suppose as before, except that *Xthat* is optional

N-tier

whathat :: $\texttt{wh}^+$

|

what :: $\texttt{wh}^-$

No accident + EM-tier with Xthat only

- $\texttt{wh}^+$ on *Xthat* **or that**
- **EM-tier**
  - project all $\texttt{wh}^+$, all $\texttt{wh}^-$, and all C-heads **that are Xthat**
  - every *Xthat* must have daughter with *Xthat* (no $\texttt{wh}^+$) or matching $\texttt{wh}^-$
  - non-agreeing C above *Xthat*: *Xthat* must carry $\texttt{wh}^+$

EM-tier: (X)that :: $\texttt{wh}^+$ Xthat*

X :: $\texttt{wh}^-$

Summary of analytical tricks for T-patterns

- **F-patterns arise from lexicon**
  - no F: $\texttt{wh}^+$ only on default C
  - F: $\texttt{wh}^+$ only on agreeing C
  - (F): $\texttt{wh}^+$ on either one
- **N-patterns captured via tiers**
  - construct **daisy chain** of agreeing mother-daughter configurations
  - top of daisy chain: $\texttt{f}^+$ (don't agree beyond the landing site)
  - bottom of daisy chain: $\texttt{f}^-$ (don't agree below the base position)
  - optionality = not projecting uninflected heads
- If we need agreement with some feature of the mover, find a suitable tier to enforce the match condition on.

A brief look at an O pattern

- **Ewe (Niger-Congo)** is OF(N)
- 3rd person subject pronouns **é** and **wò** alternate based on wh-movement
- **é** is default
- if some XP wh-moves across Spec,TP, then
  - **é must** become **wò** in the clause with the final landing site,
  - **é may** become **wò** in a clause with a non-final landing site.

(5) [$_{\text{CP}}$ Meka-e$_k$ **wò**/*__é__ gblɔ [$_{\text{CP}}$ be **wò**/__é__-bu [$_{\text{CP}}$ be
who-Foc he say that he-think that
wò/__é__-fò t_k]]]?
he-hit

'Who did he$_i$ say that he$_j$ thinks that he$_m$ hit?'

Reanalyzing Ewe as a TF(N) pattern

- Successive-cyclic movement cannot explain Ewe alternations, but it's **very natural with tiers**
- Suppose Ewe has two T-heads:
  default T and **wh-agreeing $\mathrm{T}_{\mathtt{wh}}$**
- Additional constraint on nom-tier:
  $\mathrm{T}_{\mathtt{wh}}$ :: $\mathtt{nom}^{+}$ must not have é :: $\mathtt{nom}^{-}$ as a daughter
- Distribution of $\mathrm{T}_{\mathtt{wh}}$ almost the same as *Xthat* in toy example 4
  EM-tier: $\mathrm{C} :: \mathtt{wh}^{+}\ \mathrm{T}^{*}_{wh}\ \mathrm{X} :: \mathtt{wh}^{-}$
- We need a second tier to enforce the F pattern:
  1. Project all $\mathtt{wh}^{+}$, $\mathtt{wh}^{-}$, and all T heads
  2. $\mathtt{wh}^{+}$ must not have default T as daughter

  F-tier: $\mathrm{C} :: \mathtt{wh}^{+}\ \mathrm{T}_{wh}\ \{\mathrm{T},\ \mathrm{T}_{wh}\}^{*}\ \mathrm{X} :: \mathtt{wh}^{-}$

Another case: **no**-marking in Duala

- **Duala (Niger-Congo)** is OF
- A′-movement of object or adjunct triggers insertion of **no** after finite verb in T
- same analysis as Ewe, except that:
  - the mover must not carry $\texttt{nom}^-$, and
  - we do not allow $\mathrm{T}_{\texttt{wh}}$ as a daughter of $\mathrm{T}_{\texttt{wh}}$

EM-tier: $\mathrm{C} :: \texttt{wh}^+\ \mathrm{T}_{wh}\ \mathrm{X} :: \texttt{wh}^-$

F-tier: $\mathrm{C} :: \texttt{wh}^+\ \mathrm{T}_{wh}\ \mathrm{T}^*\ \mathrm{X} :: \texttt{wh}^-$

An extraction morphology surprise: floating quantifiers!

(6) These kids may (all) have (all) been (all) believed to (all) t like candy.

- Languages vary greatly in
  - what quantifiers may float
  - what movements license quantifier float
  - in what positions quantifiers may float
  - whether floating quantifiers exhibit agreement
- **Tiers capture all that variation!**

Sketch of analysis

- Lexicon contains $\mathbf{QX}_{\mathtt{f}} :: \mathbf{P}^{+}\mathbf{P}^{-}$ for every
  - floating quantifier **Q** with agreement **X**
  - movement type $\mathtt{f}$ that licenses quantifier float of **Q**
  - position **P** where **Q** may float
- Distribution of floating quantifiers regulated via EM-tier
  EM-tier: $\mathtt{f}^{+}\ \mathrm{QX}^{*}_{\mathtt{f}}\ \mathrm{X} :: \mathtt{f}^{-}$

An extraction morphology surprise: floating quantifiers!

(6) These kids may (all) have (all) been (all) believed to (all) t like candy.

- Languages vary greatly in
  - what quantifiers may float
  - what movements license quantifier float
  - in what positions quantifiers may float
  - whether floating quantifiers exhibit agreement
- **Tiers capture all that variation!**

Sketch of analysis

- Lexicon contains $\mathbf{QX}_{\mathtt{f}} :: \mathbf{P}^{+}\mathbf{P}^{-}$ for every
  - floating quantifier **Q** with agreement **X**
  - movement type $\mathtt{f}$ that licenses quantifier float of **Q**
  - position **P** where **Q** may float
- Distribution of floating quantifiers regulated via EM-tier

  EM-tier: $\mathtt{f}^{+}\ \mathrm{QX}^{*}_{\mathtt{f}}\ \mathrm{X} :: \mathtt{f}^{-}$

Acknowledgments

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

References I

Georgi, Doreen. 2017. Patterns of movement reflexes as the result of the order of merge and agree. *Linguistic Inquiry* 48:585–626.

Keine, Stefan. 2016. *Probes and their horizons*. Doctoral Dissertation, University of Massachusetts, Amherst.

Keine, Stefan. 2019. Selective opacity. *Linguistic Inquiry* 50:13–62.

Mayer, Connor. 2021. Capturing gradience in long-distance phonology using probabilistic tier-based strictly local grammars. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2021*, 39–50.

McCloskey, James. 2001. The morphosyntax of wh-extraction in Irish. *Journal of Linguistics* 37:67–100.

Torrence, Harold. 2012. The morpho-syntax of silent *wh*-expressions in Wolof. *Natural Language and Linguistic Theory* 30:1147–1184.

Torres, Charles, Kenneth Hanson, Thomas Graf, and Connor Mayer. 2023. Modeling island effects with probabilistic tier-based strictly local grammars over trees. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2023*, pp–pp.