Optimizing for parsing: modeling word-order preferences

Lei Liu

Leipzig University

Cyclop retreat 2024 Winter

slides updated on November 28, 2024

The presentation in bullet points:

The presentation in bullet points:

- A left-corner Minimalist Grammar parsing model captures mixed word-order preferences.

The presentation in bullet points:

- A left-corner Minimalist Grammar parsing model captures mixed word-order preferences.
- Syntactic operations allow optimization for parsing?

Outline

Outline

1. Introduction

2. Modeling mix word-order preferences
   - English heavy NP shift
   - Japanese transitive

3. Next steps and open questions

Introduction

Previously in my parsing corner...

- Complexity metric: tenure

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing
- Empirical phenomena: multiple embedding

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing
- Empirical phenomena: multiple embedding
  - multiple left-, right-embedding: ☺

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing
- Empirical phenomena: multiple embedding
  - multiple left-, right-embedding: 😊
  - multiple center-embedding: 😳

Previously in my parsing corner...

- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing
- Empirical phenomena: multiple embedding
  - multiple left-, right-embedding: 😊
  - multiple center-embedding: 😳

Previously in my parsing corner...

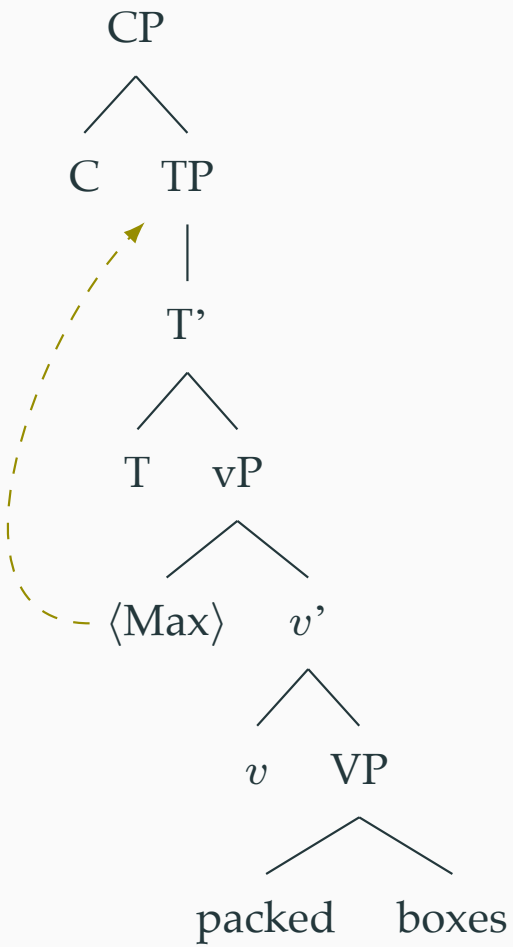
- Complexity metric: tenure
  - how long nodes/parse items are retained in memory
  - reliable metric for processing difficulties
    - in top-down MG parsing
    - in left-corner MG parsing
- Empirical phenomena: multiple embedding
  - multiple left-, right-embedding: 😊
  - multiple center-embedding: 😳

In this episode...

left-corner MG parsing model and word-order preferences.

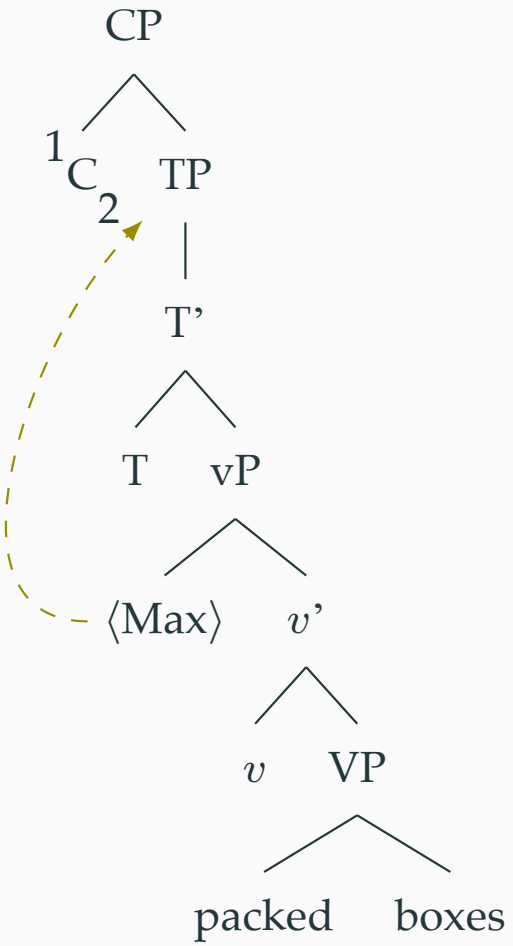
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



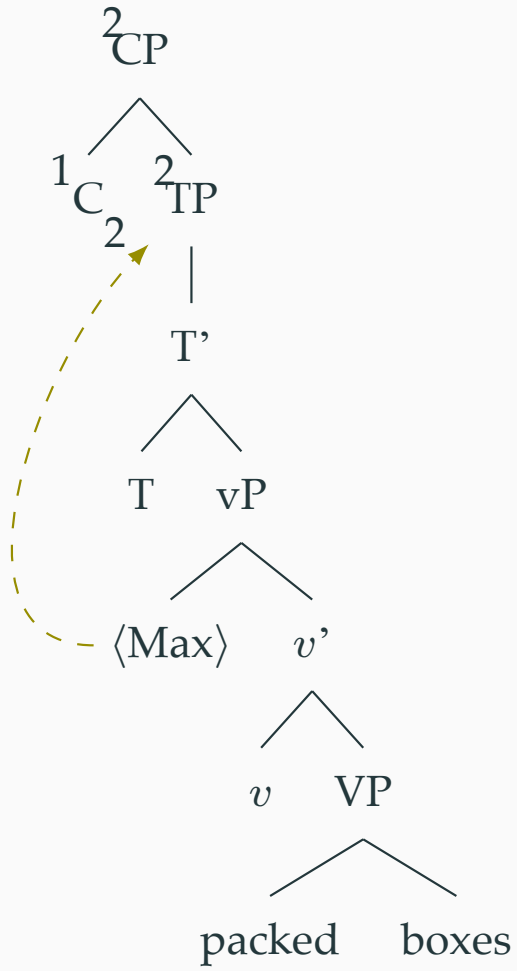
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



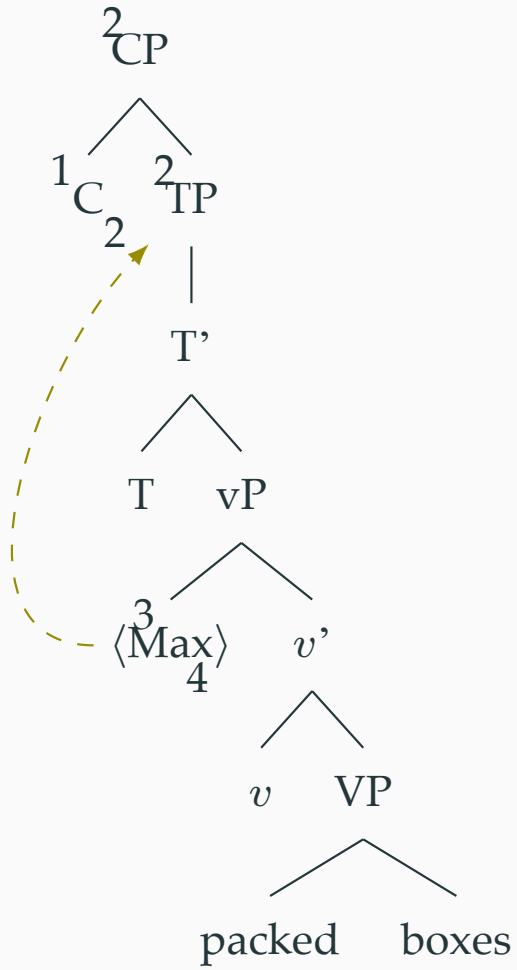
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



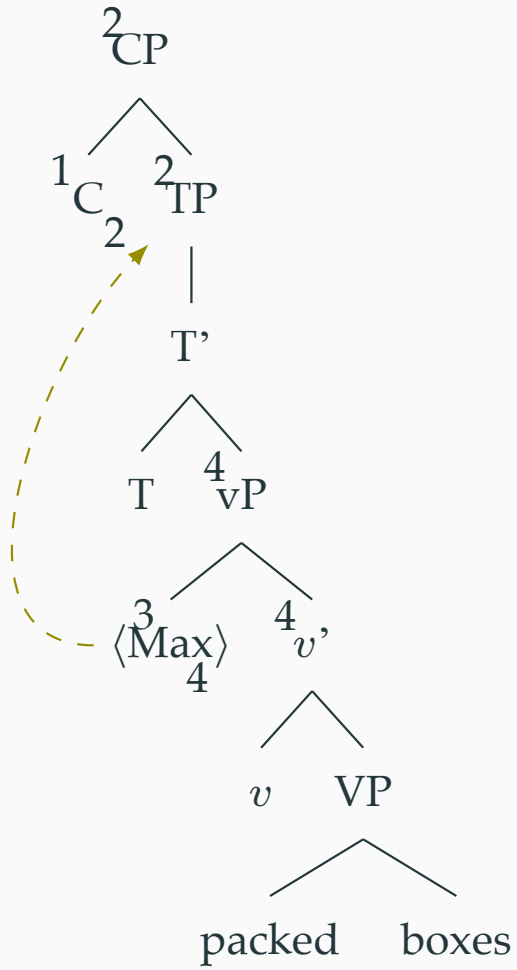
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



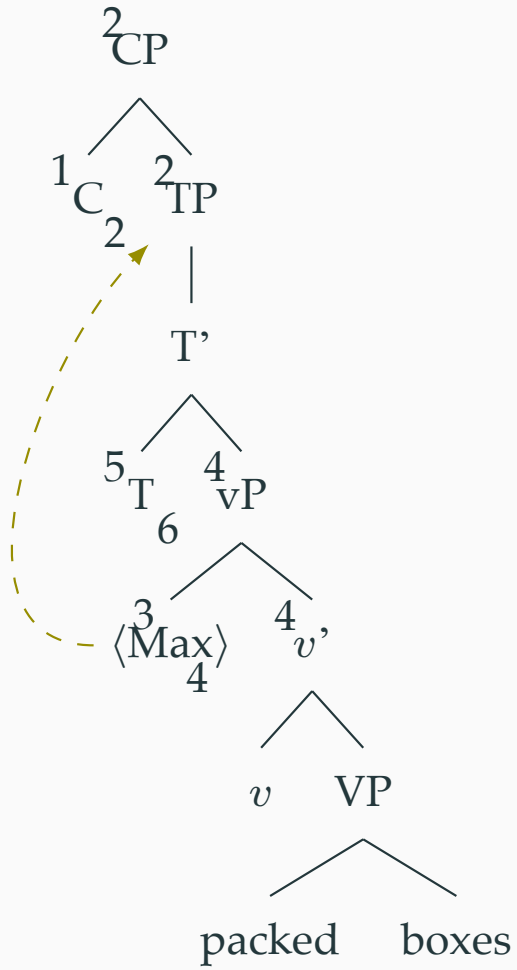
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



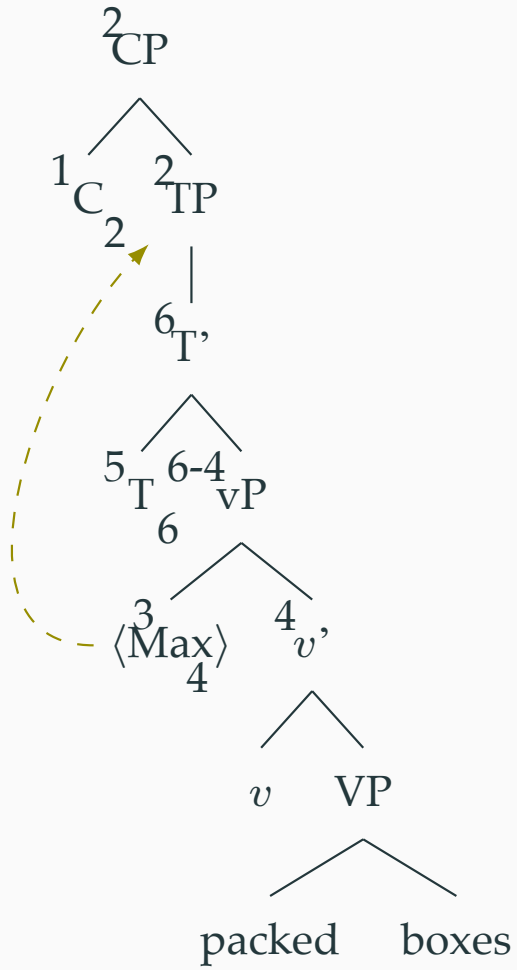
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



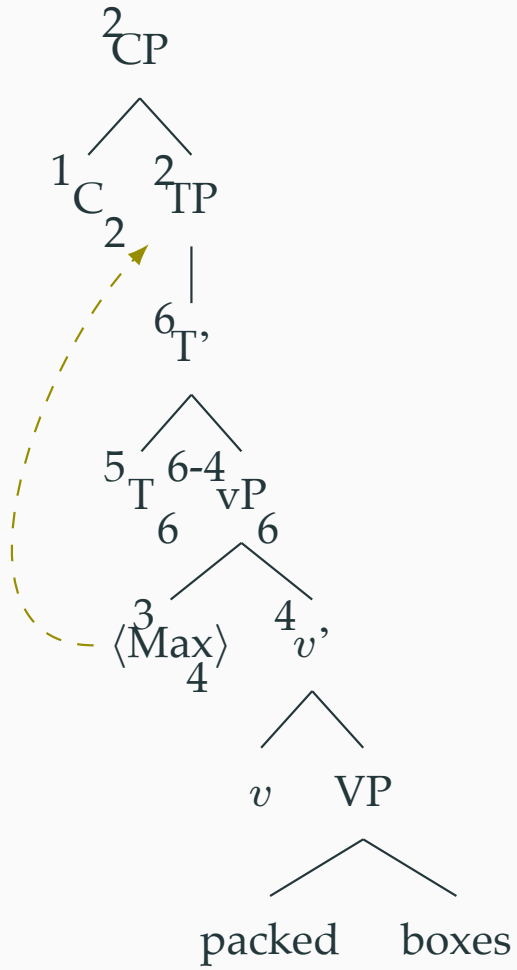
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



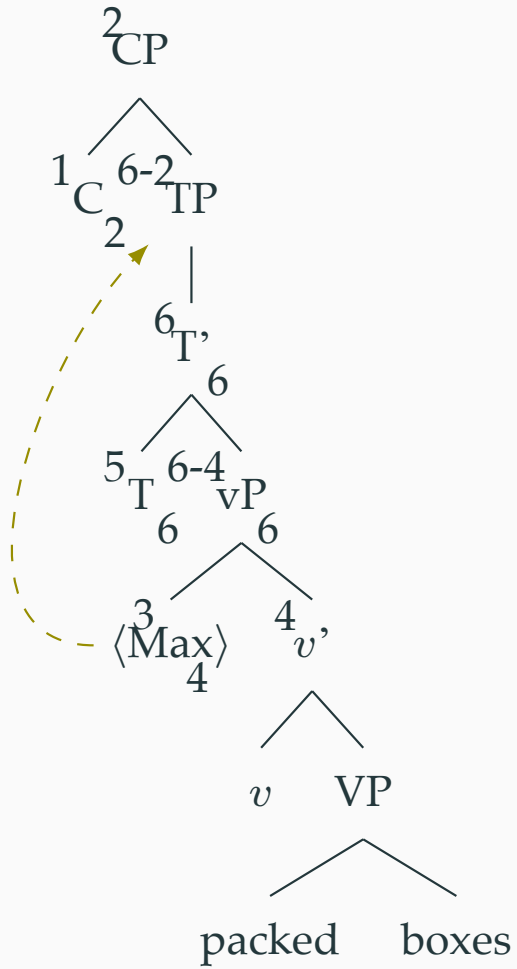
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



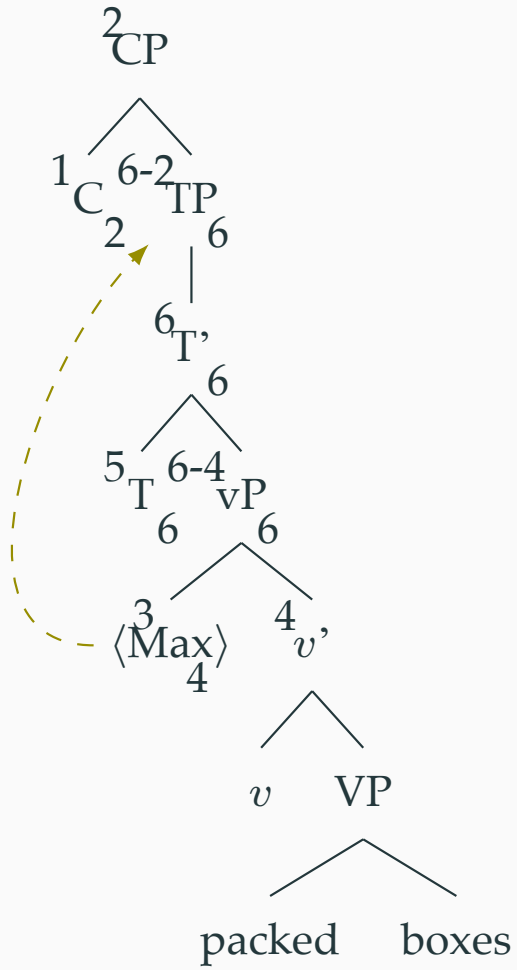
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



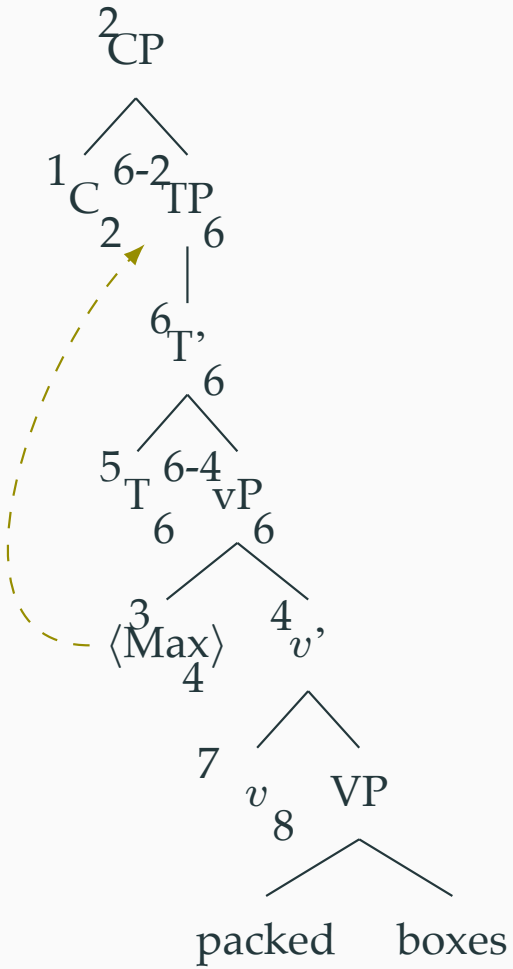
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



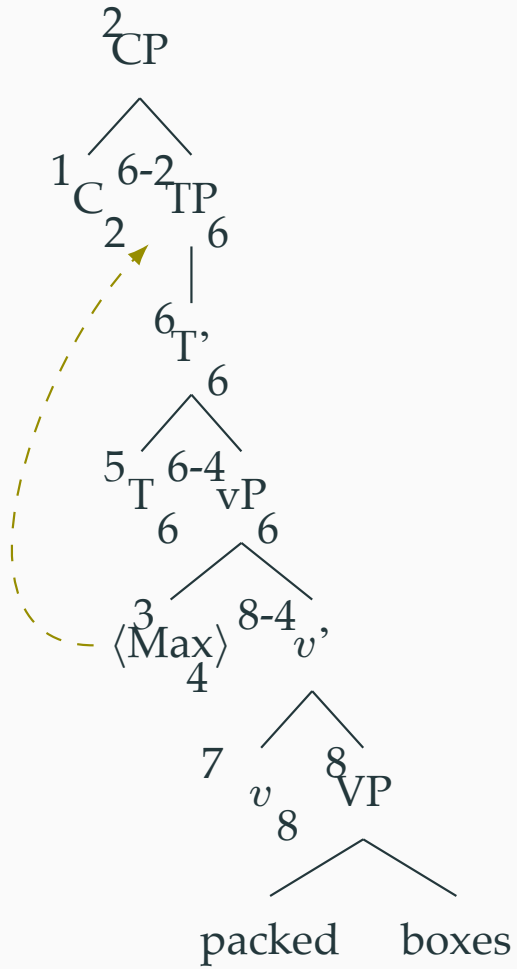
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



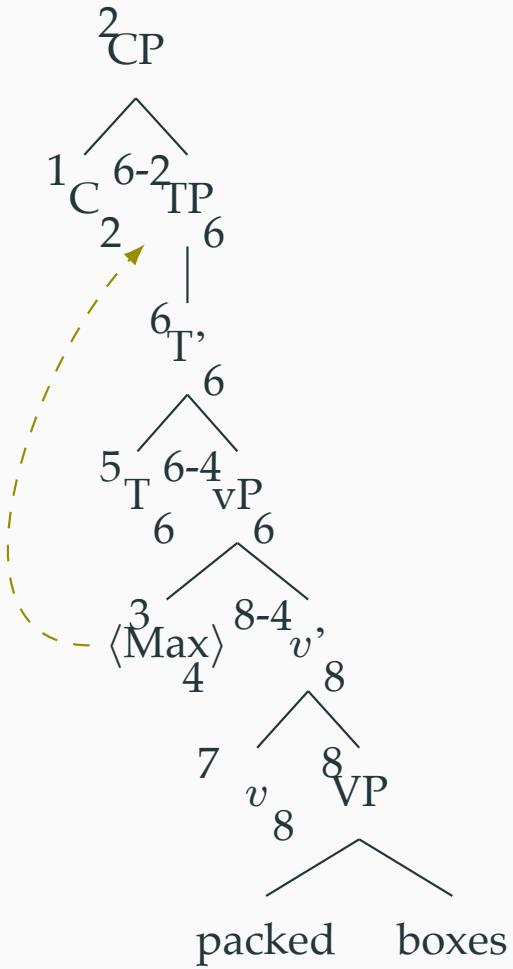
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



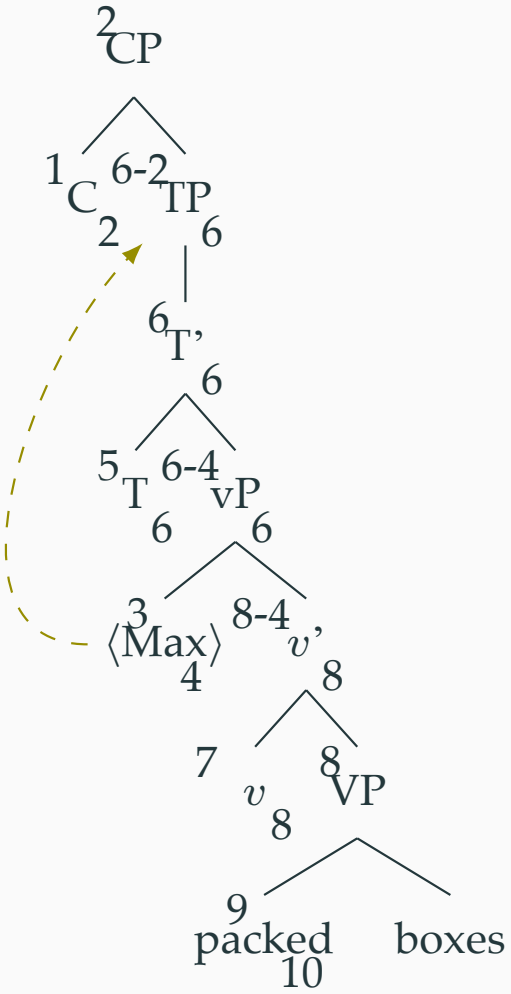
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



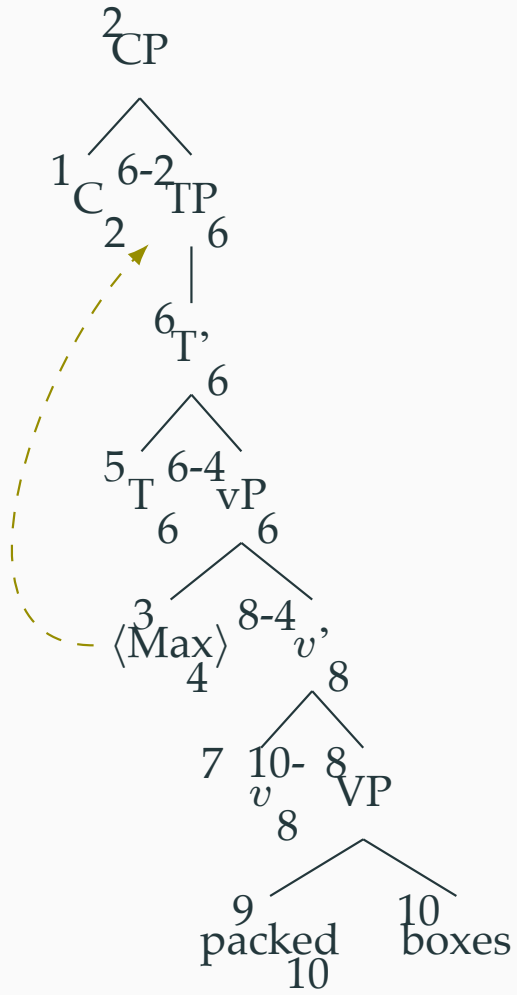
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



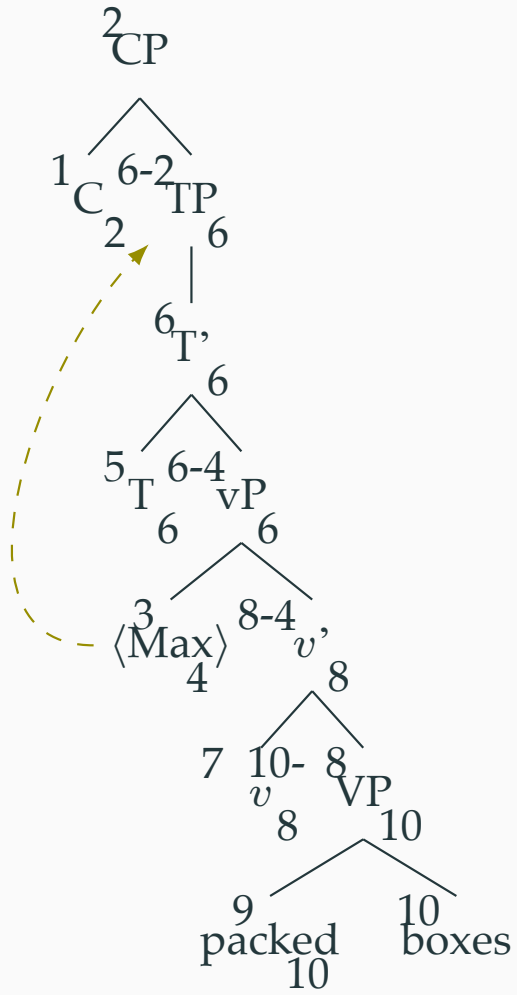
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



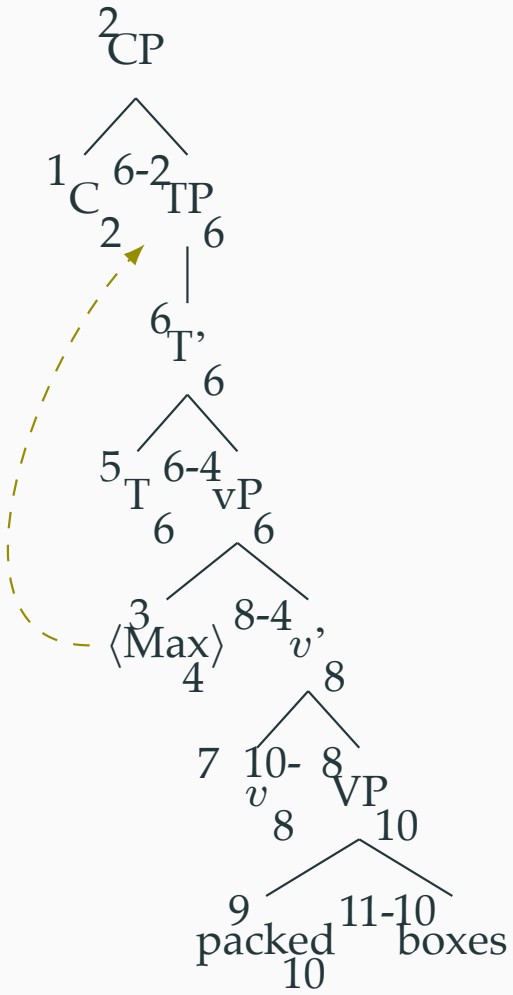
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



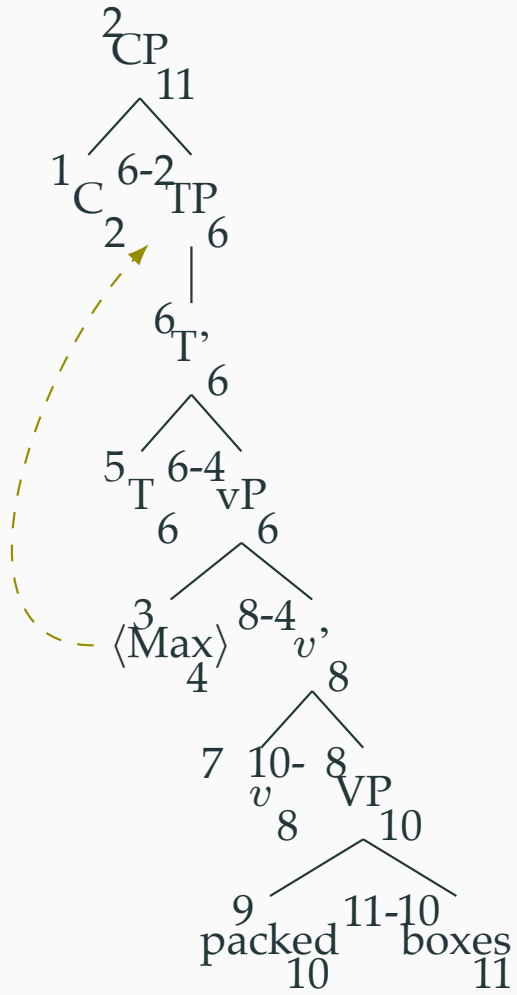
LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



LC MG parser

Step 1 read C
Step 2 predict CP, TP
Step 3 read Max
Step 4 predict *vP*, *v*'
Step 5 read T
Step 6 predict T', *vP*
Step 6' connect down
Step 6" un-move
Step 6"' connect up
Step 7 read *v*
Step 8 predict *v*′, VP
Step 8' connect down
Step 9 read packed
Step 10 predict VP, boxes
Step 10' connect down
Step 11 read boxes and complete



Modeling mix word-order preferences

Overview: mixed word-order preferences

- Short-before-long preference

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) arm-chair judgement

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)
- No preference

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)
- No preference
  - Mandarin preverbal PPs — UD (Liu 2020)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)
- No preference
  - Mandarin preverbal PPs — UD (Liu 2020)
  - Japanese preverbal PPs — UD (Liu 2020)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)
- No preference
  - Mandarin preverbal PPs — UD (Liu 2020)
  - Japanese preverbal PPs — UD (Liu 2020)

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS) — arm-chair judgement
  - English post-verbal PP adjuncts — Universal Dependencies (UD) (Liu 2020)
  - Mandarin *ba*-construction and its alternative — Penntree bank (Liu 2022)
- Long-before-short preference
  - Japanese di/transitive — Self-paced reading (Yamashita and Chang 2001)
- No preference
  - Mandarin preverbal PPs — UD (Liu 2020)
  - Japanese preverbal PPs — UD (Liu 2020)

LC MG parsing captures all of the above!

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS)
  - English post-verbal PP adjuncts
  - Mandarin *ba*-construction and its alternative
- Long-before-short preference
  - Japanese transitive
- No preference
  - Mandarin preverbal PPs
  - Japanese preverbal PPs

Overview: mixed word-order preferences

- Short-before-long preference
  - English heavy NP shift (HNPS)
    - how does the model work
- Long-before-short preference
  - Japanese transitive
    - syntactic analyses
- No preference
  - Mandarin preverbal PPs
    - Other metrics?

English HNPS: Data

(1) Max put [$_{PP}$ in his car] [$_{DP}$ all the boxes of home furnishings].

(2) ?Max put [$_{DP}$ all the boxes of home furnishings] [$_{PP}$ in his car].

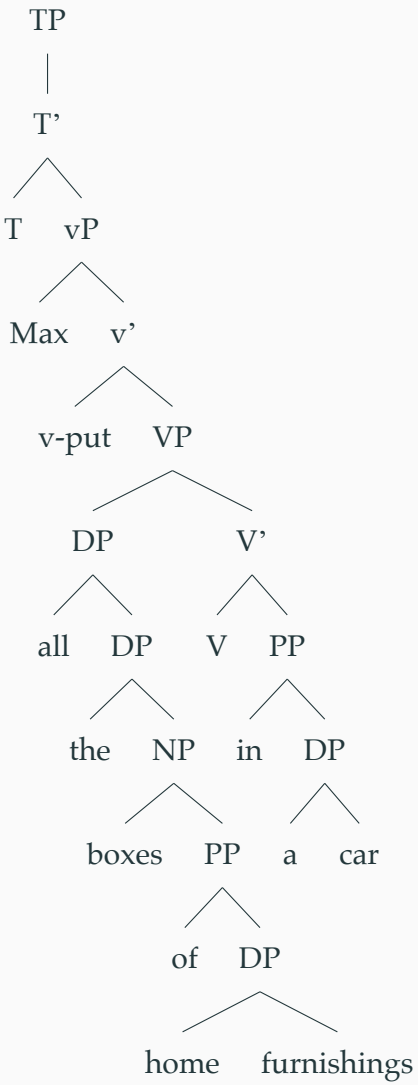
English HNPS: Data

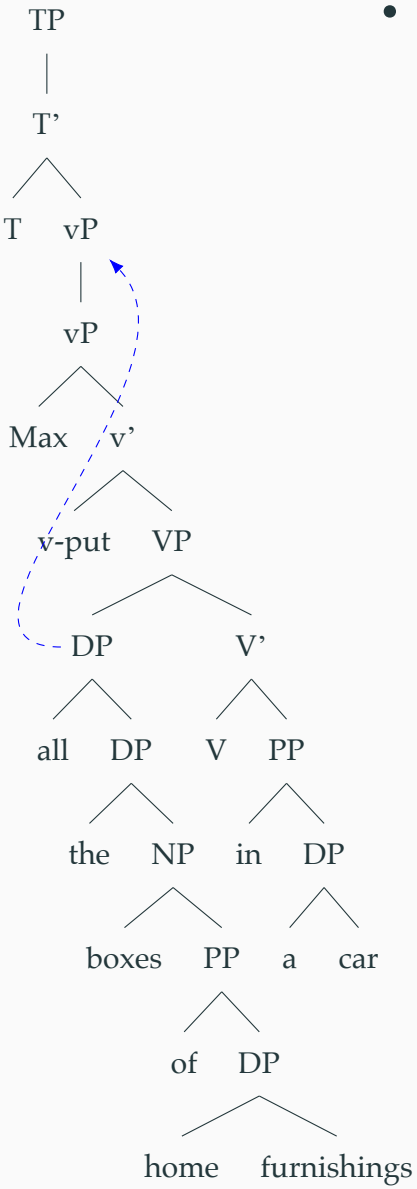
(1) Max put [$_{PP}$ in his car] [$_{DP}$ all the boxes of home furnishings].

(2) ?Max put [$_{DP}$ all the boxes of home furnishings] [$_{PP}$ in his car].

A shifted order (PP-DP) is preferred when the DP is heavy (e.g. Brown corpus Wasow 1997).

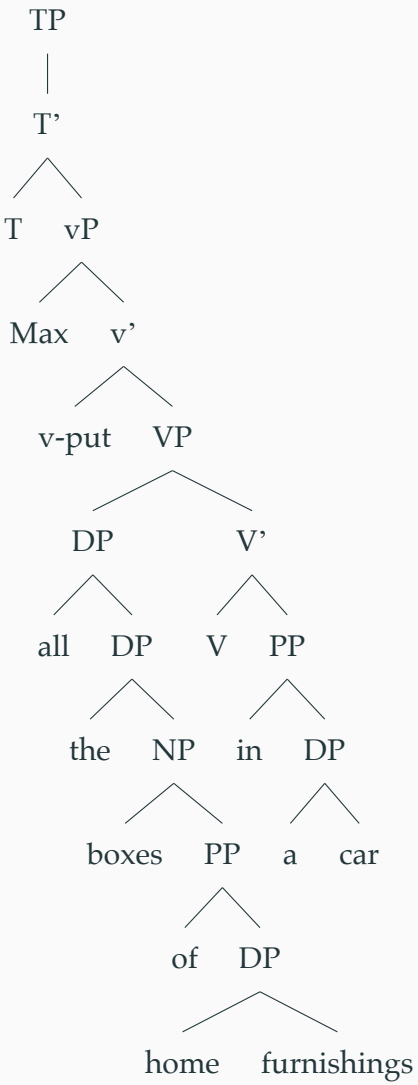
English HNPS: Syntax and modeling

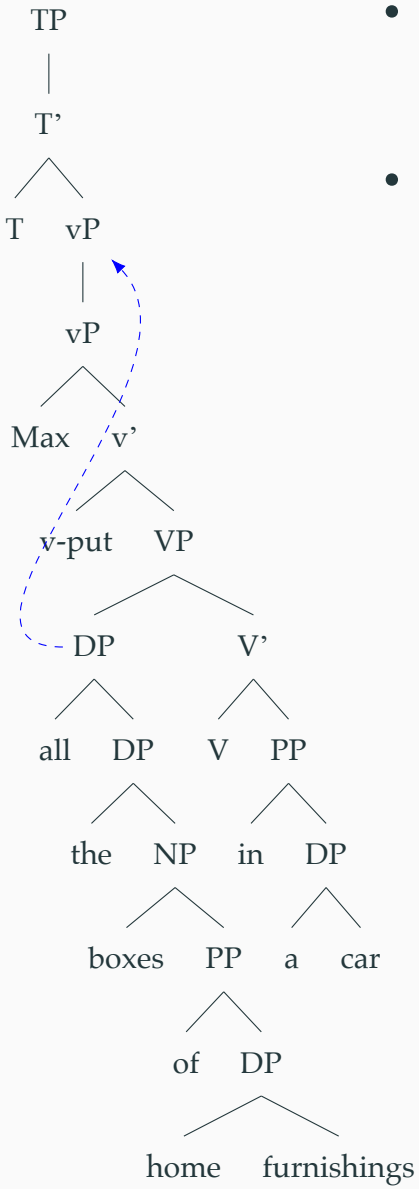




- Shifted order: rightward movement (Ross 1986)

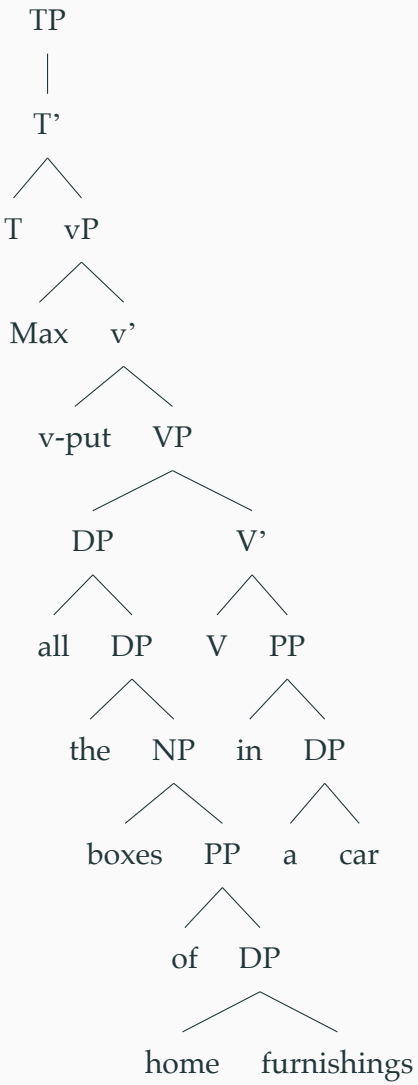
English HNPS: Syntax and modeling

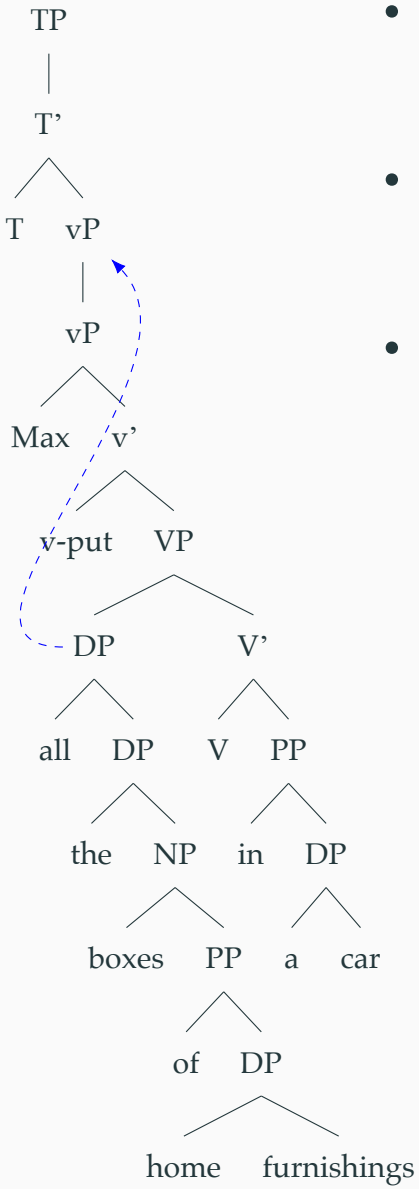




- Shifted order: rightward movement (Ross 1986)
- Ignored: V-to-v movement, ArgO-movement

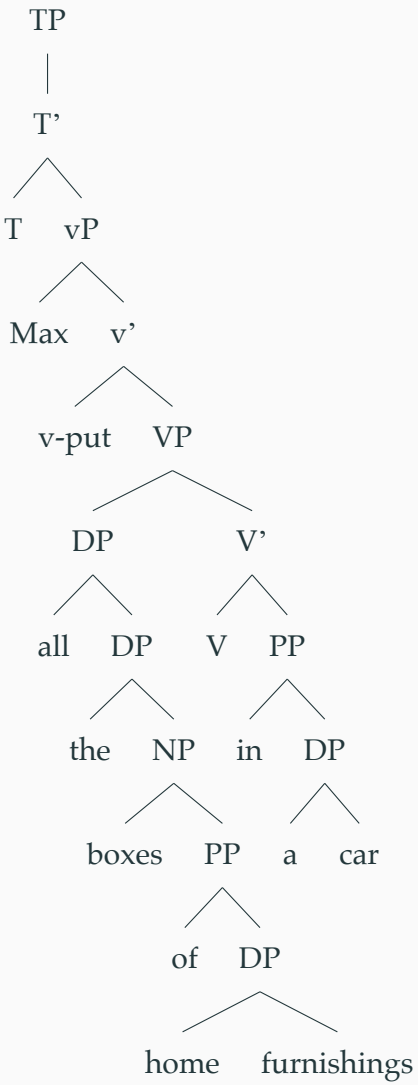
English HNPS: Syntax and modeling

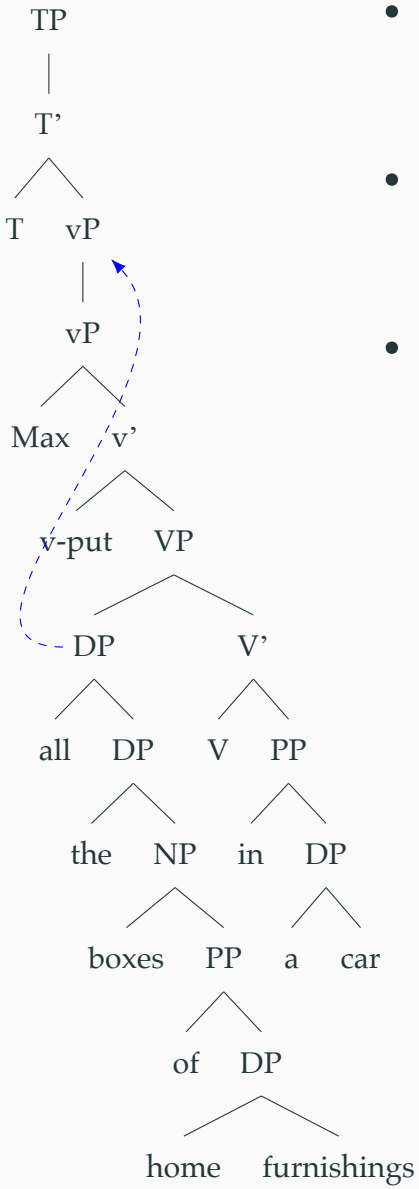




- Shifted order: rightward movement (Ross 1986)
- Ignored: V-to-v movement, ArgO-movement
- Results:

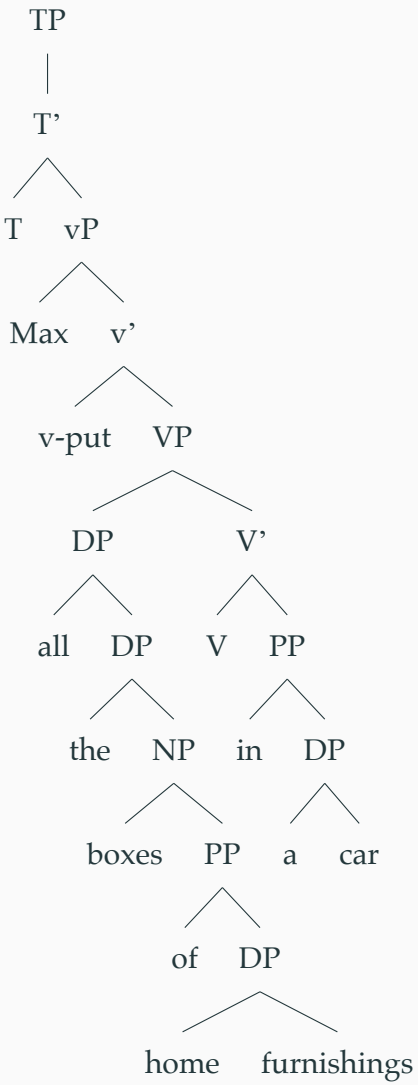
English HNPS: Syntax and modeling

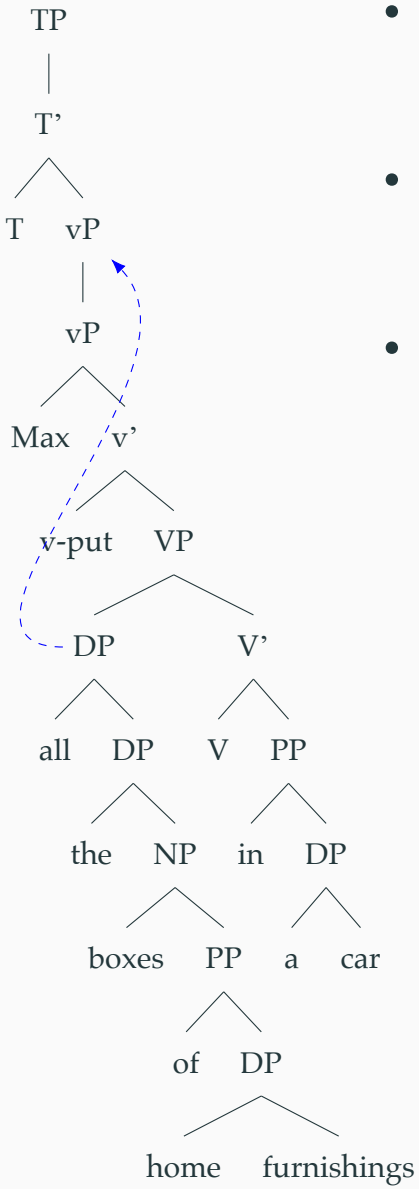




- Shifted order: rightward movement (Ross 1986)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 8

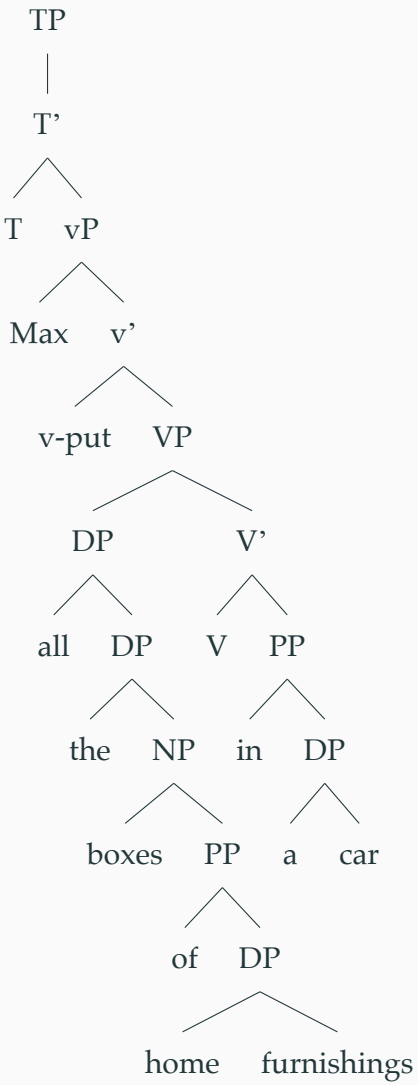
English HNPS: Syntax and modeling

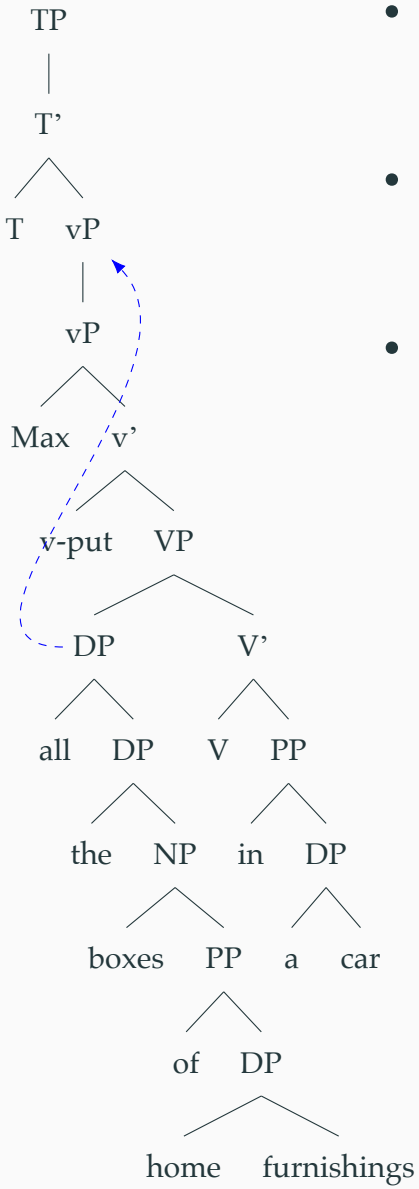




- Shifted order: rightward movement (Ross 1986)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 8
  - Shifted order: MaxTitem = 12

English HNPS: Syntax and modeling

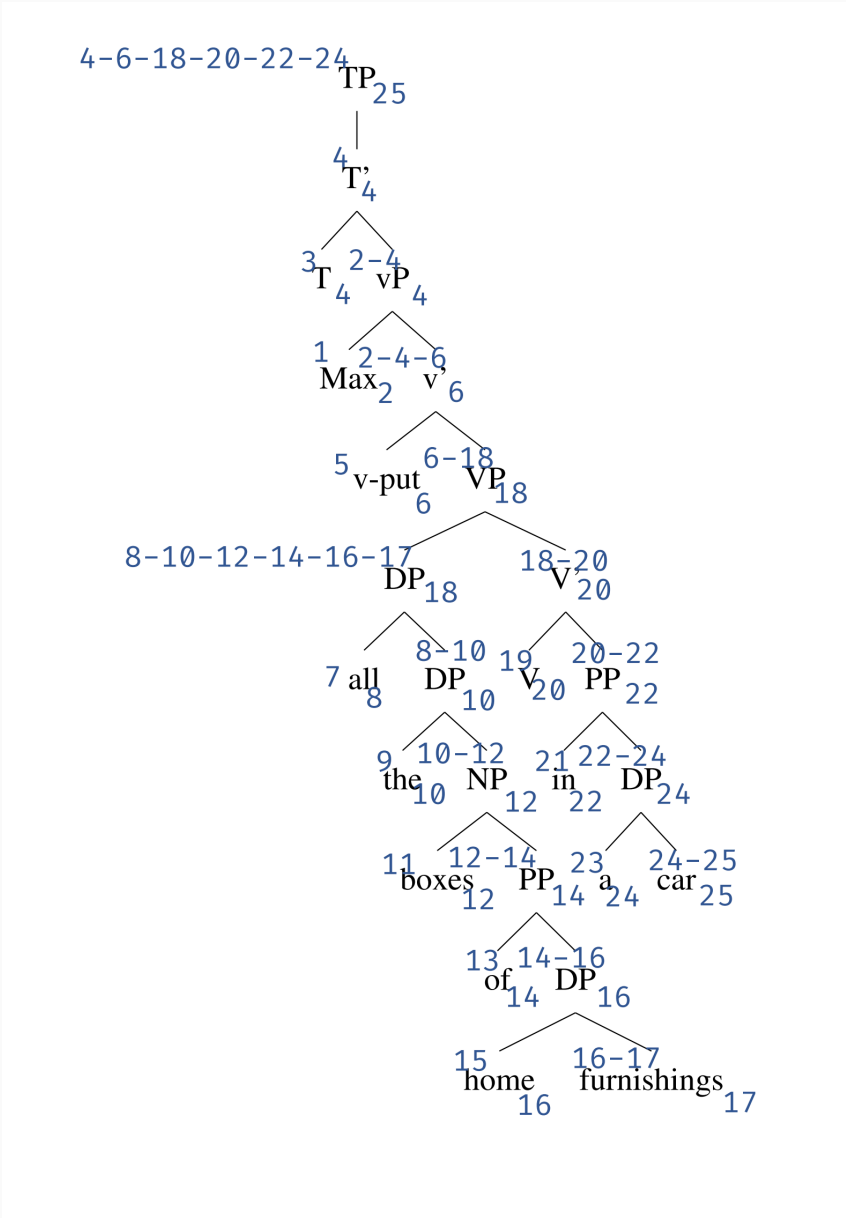




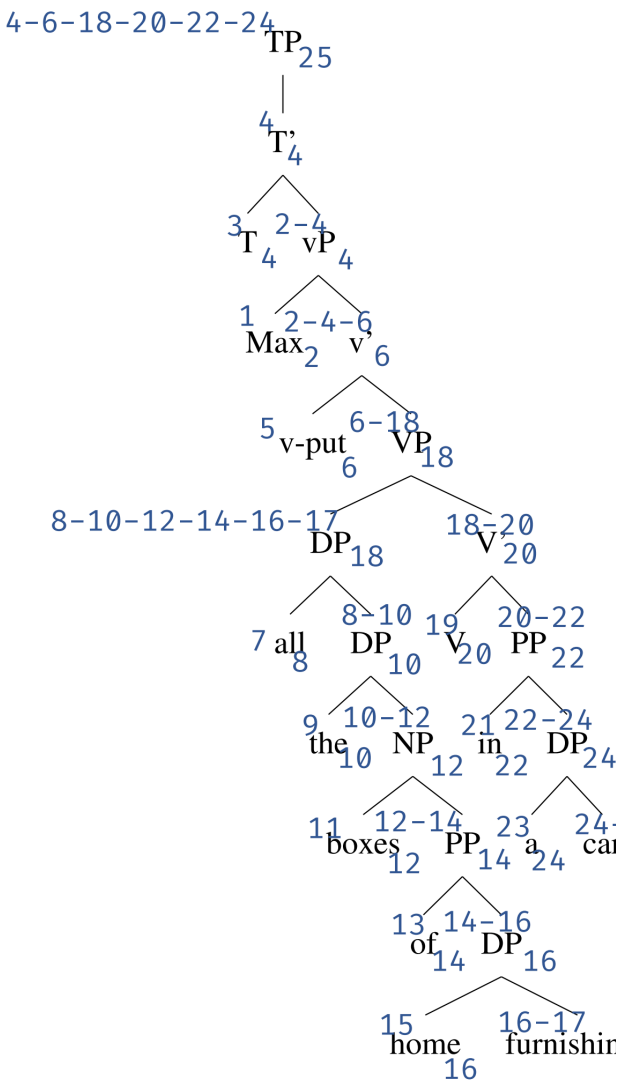
- Shifted order: rightward movement (Ross 1986)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 8
  - Shifted order: MaxTitem = 12

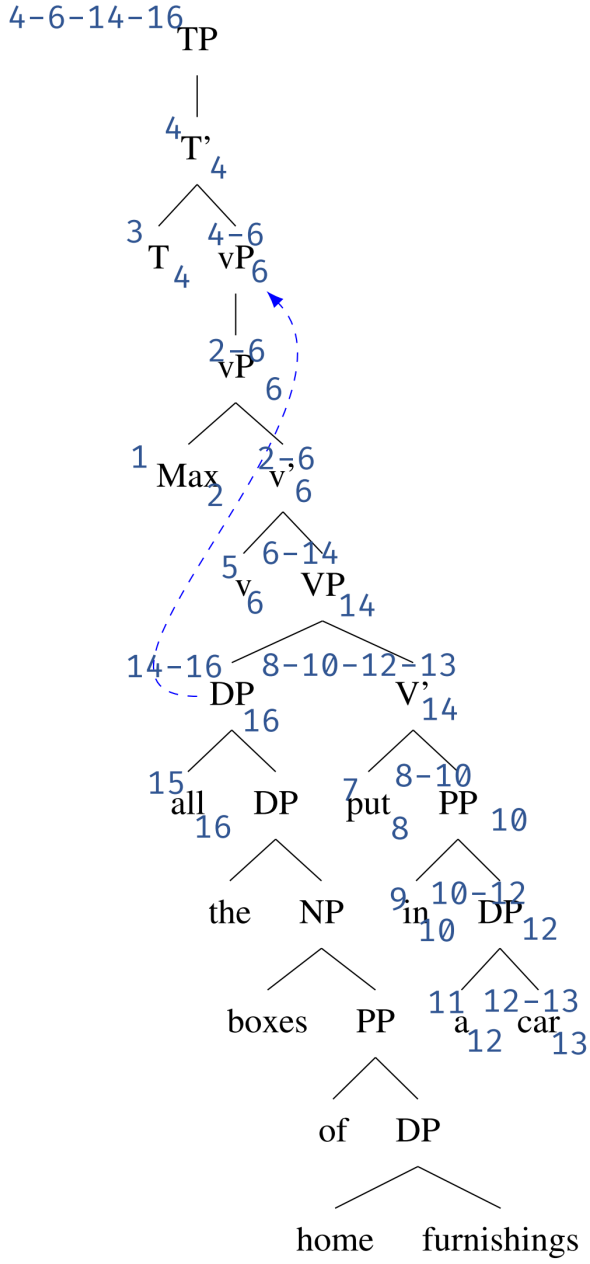
  => Shifted order is easier to process

English HNPS: Modeling results



English HNPS: Modeling results





English HNPS and similar cases

- Similar short-before-long preferences predicted in
  - English heavy NP shift (HNPS)
  - English post-verbal PP adjuncts
  - Mandarin *ba*-construction and its alternative

English HNPS and similar cases

- Similar short-before-long preferences predicted in
  - English heavy NP shift (HNPS)
  - English post-verbal PP adjuncts
  - Mandarin *ba*-construction and its alternative
- Next steps
  - Shift iff heavy

Japanese transitive: Data

(3) SOV

[keezi-ga] [Se-ga takakute gassiri sita hanni-o] oikaketa
detective-nom height-nom tall-and big-boned suspect-acc chased

'The detective chased the suspect who is tall and big-boned.'

Japanese transitive: Data

(3) SOV

[keezi-ga] [Se-ga takakute gassiri sita hanni-o] oikaketa
detective-nom height-nom tall-and big-boned suspect-acc chased

'The detective chased the suspect who is tall and big-boned.'

(4) OSV

[Se-ga takakute gassiri sita hanni-o] [keezi-ga] oikaketa
height-nom tall-and big-boned suspect-acc detective-nom chased

Japanese transitive: Data

(3) SOV

[keezi-ga] [Se-ga takakute gassiri sita hanni-o] oikaketa
detective-nom height-nom tall-and big-boned suspect-acc chased

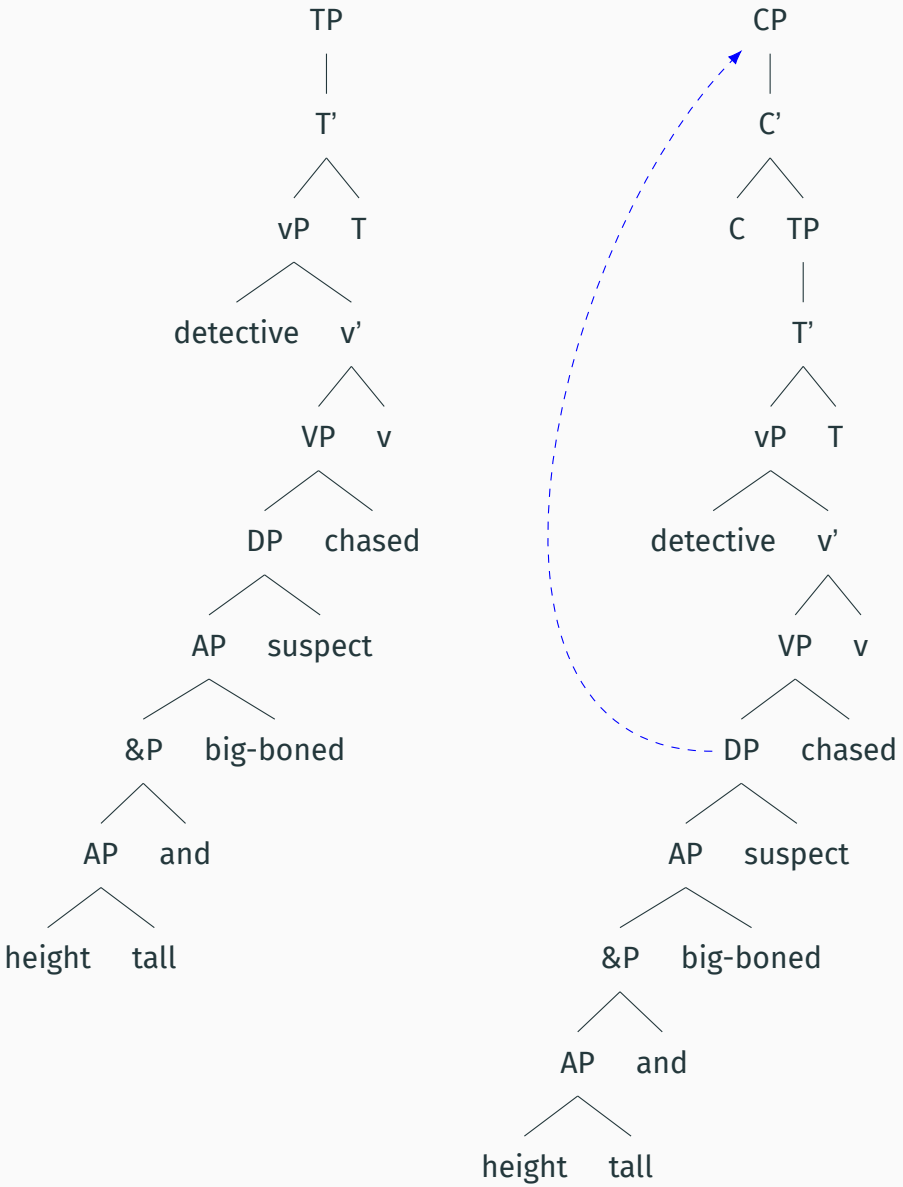
'The detective chased the suspect who is tall and big-boned.'

(4) OSV

[Se-ga takakute gassiri sita hanni-o] [keezi-ga] oikaketa
height-nom tall-and big-boned suspect-acc detective-nom chased

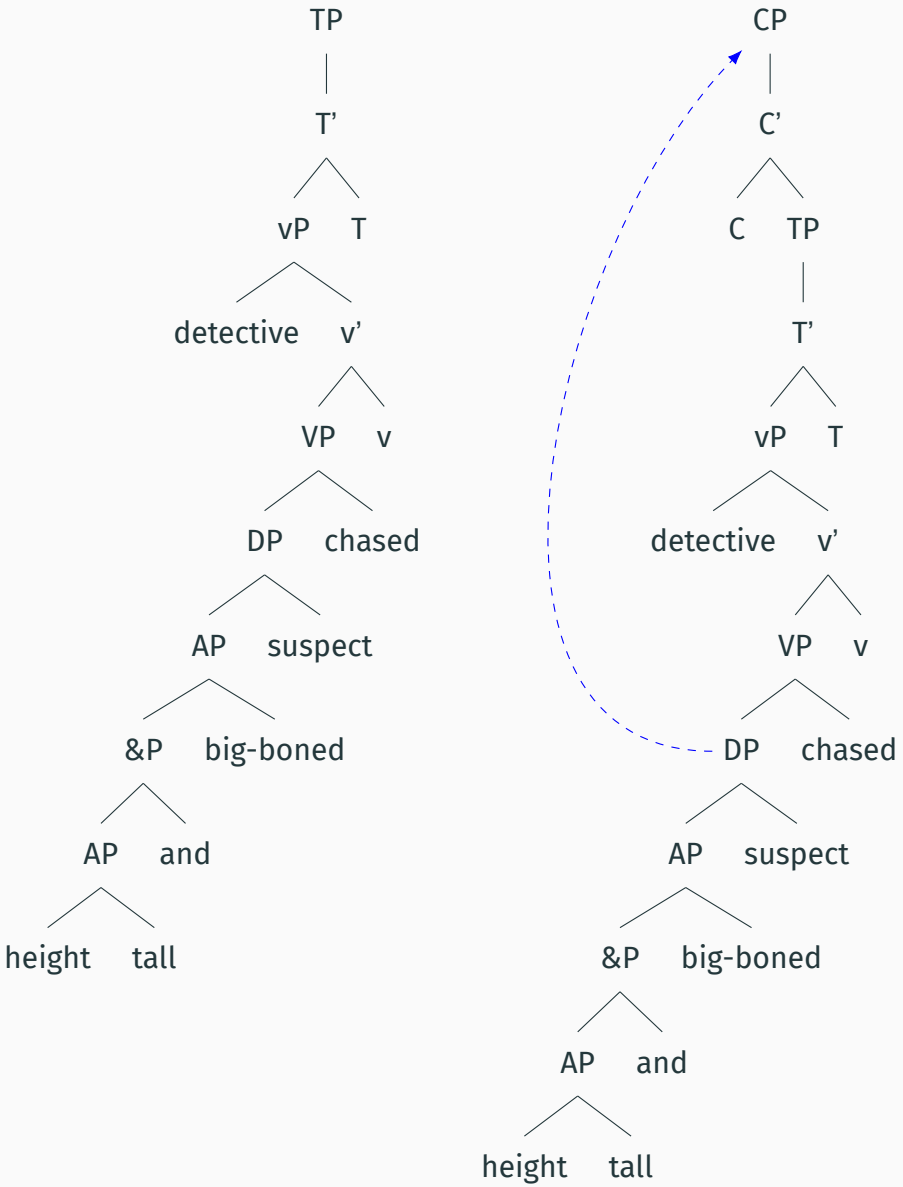
"Japanese speakers tend to shift long arguments ahead of short ones in an on-line task." (Yamashita and Chang 2001)

Japanese transitive: Syntax and modeling



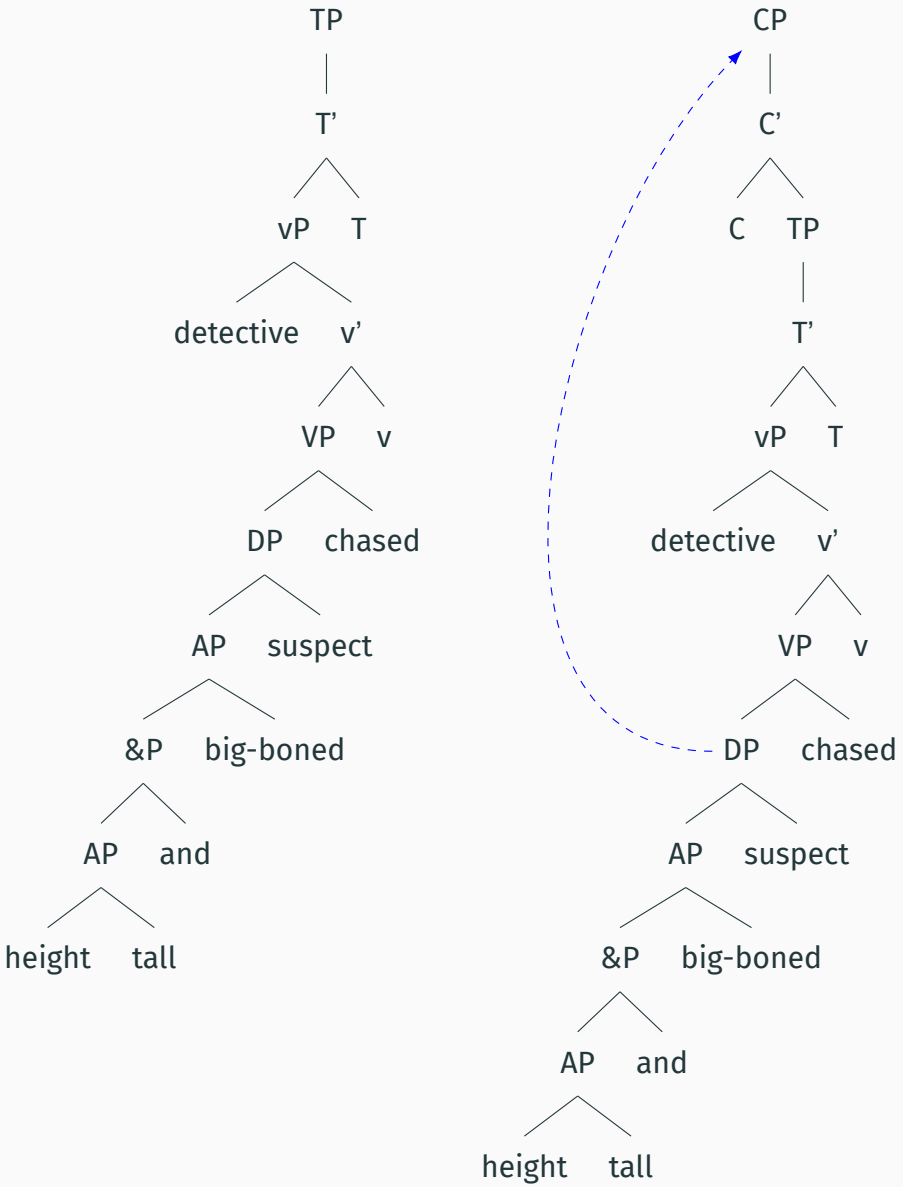
- Shifted order: scrambling to CP (Saito 1992)

Japanese transitive: Syntax and modeling



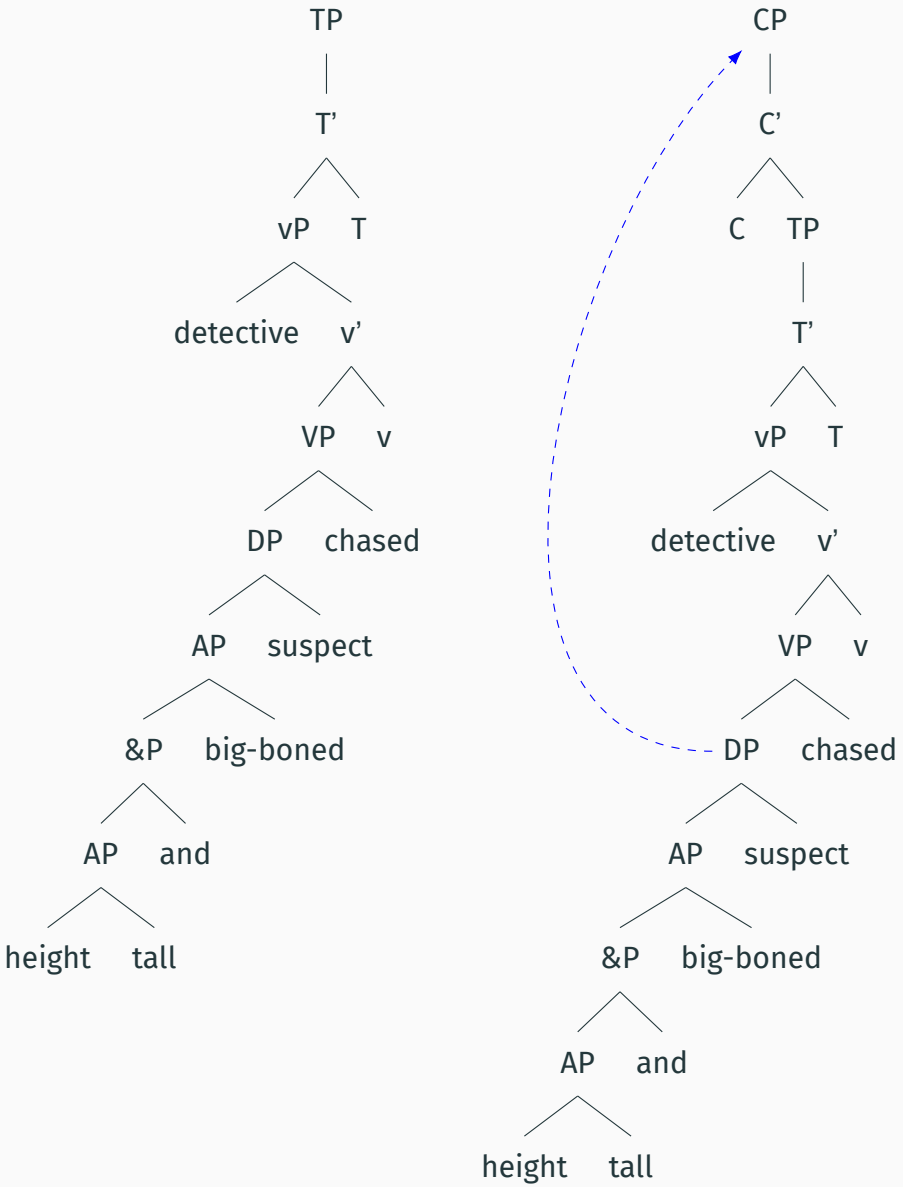
- Shifted order: scrambling to CP (Saito 1992)
- Ignored: V-to-v movement, ArgO-movement

Japanese transitive: Syntax and modeling



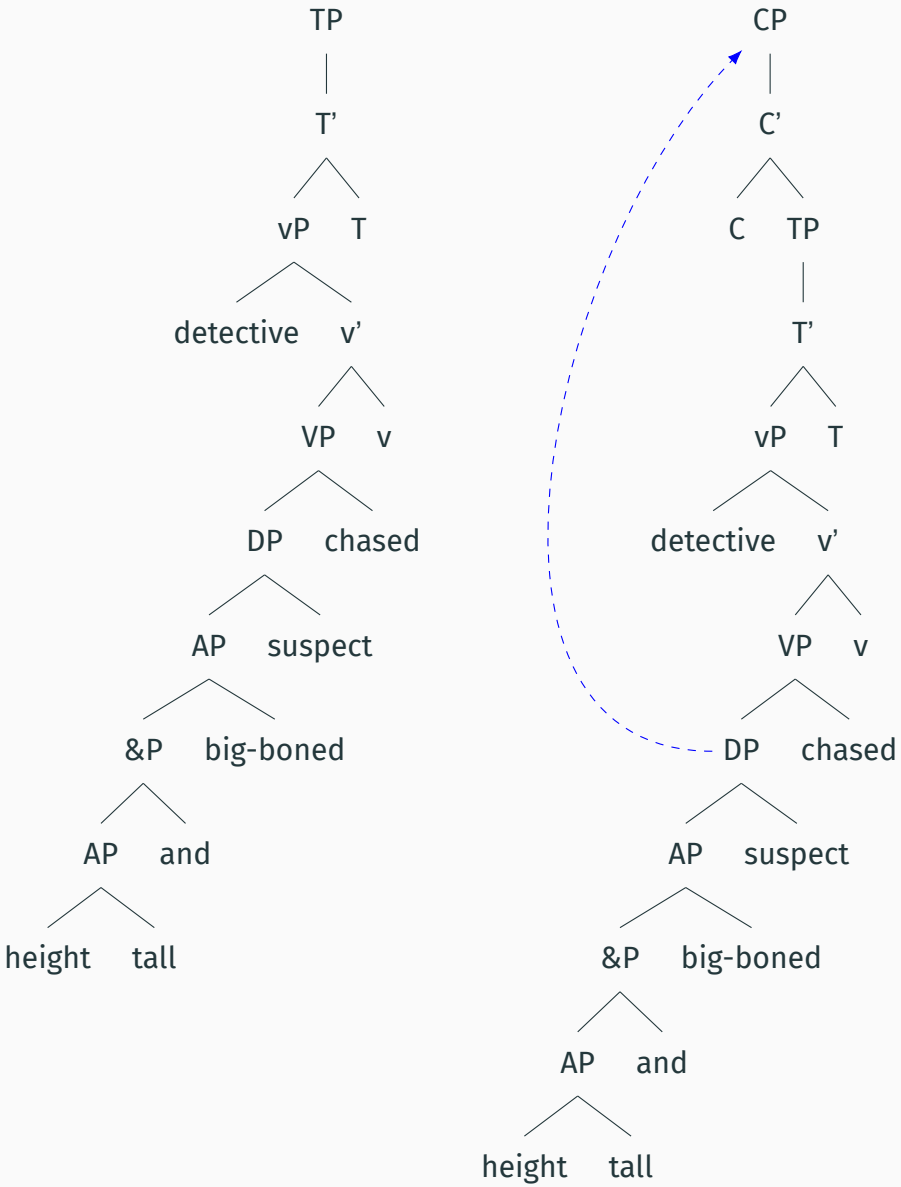
- Shifted order: scrambling to CP (Saito 1992)
- Ignored: V-to-v movement, ArgO-movement
- Results:

Japanese transitive: Syntax and modeling



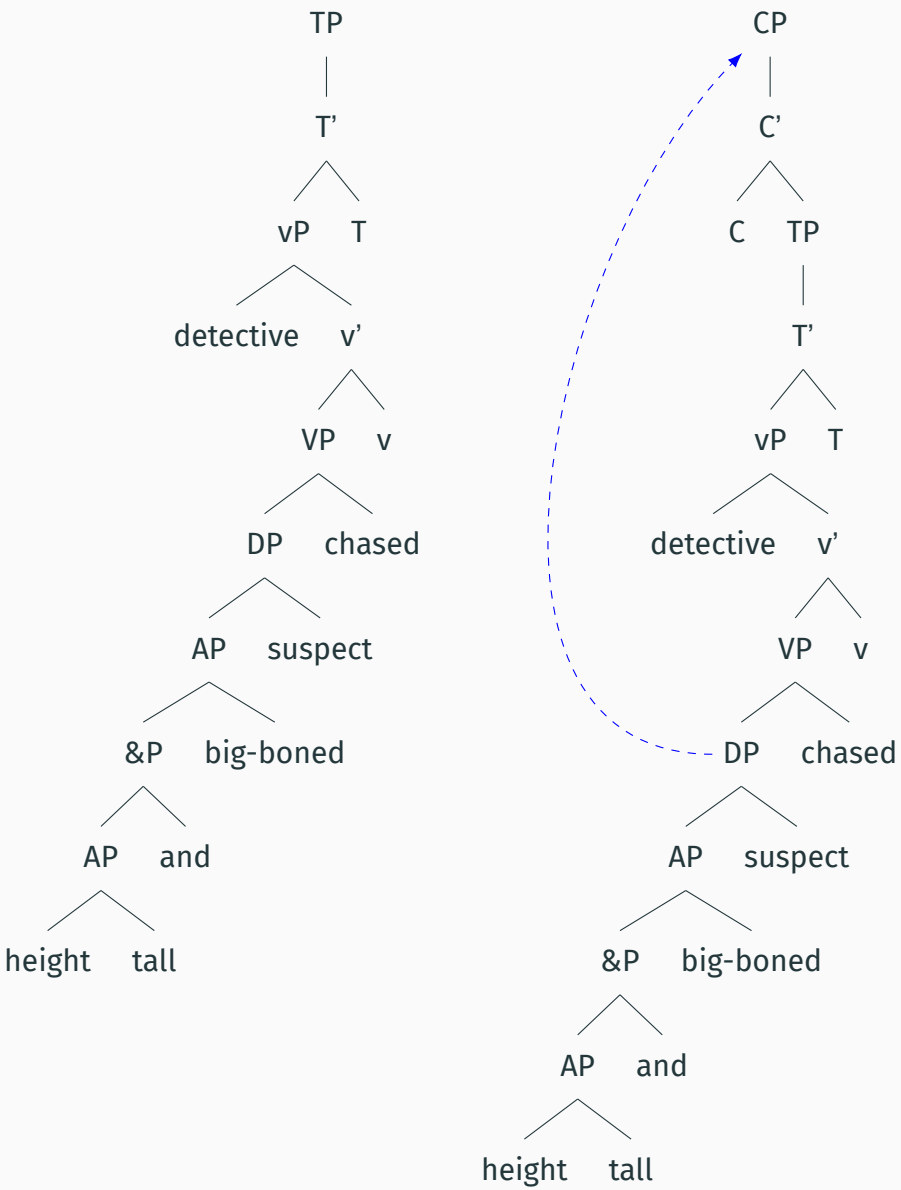
- Shifted order: scrambling to CP (Saito 1992)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 12

Japanese transitive: Syntax and modeling



- Shifted order: scrambling to CP (Saito 1992)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 12
  - Shifted order: MaxTitem = 7

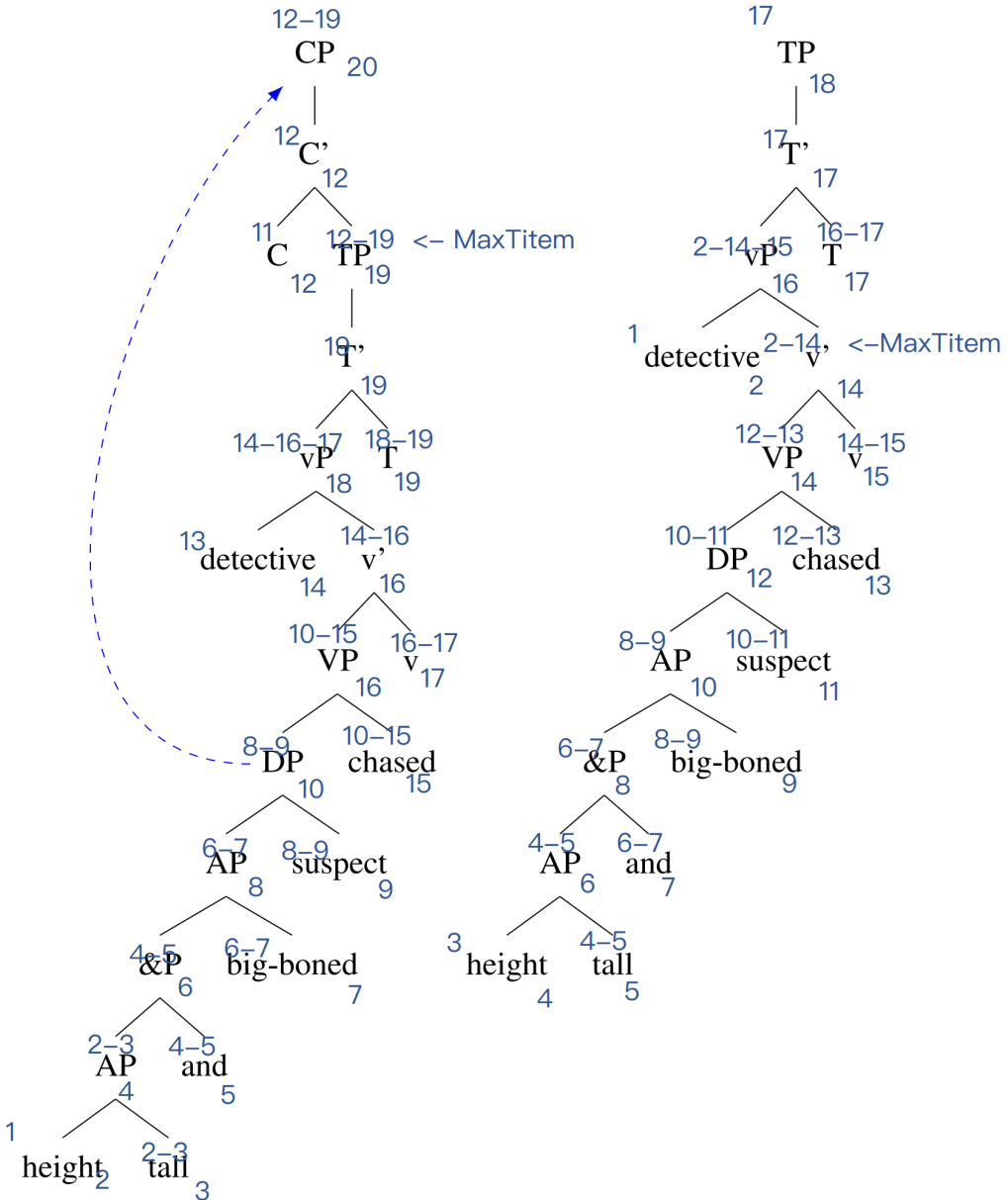
Japanese transitive: Syntax and modeling



- Shifted order: scrambling to CP (Saito 1992)
- Ignored: V-to-v movement, ArgO-movement
- Results:
  - canonical order: MaxTitem = 12
  - Shifted order: MaxTitem = 7

=> Shifted order is easier to process

Japanese transitive: Modeling results



Japanese transitive and similar cases

- Similar long-before-short preferences to model in

Japanese transitive and similar cases

- Similar long-before-short preferences to model in
  - Korean PP dative (Choi 2007)

Japanese transitive and similar cases

- Similar long-before-short preferences to model in
  - Korean PP dative (Choi 2007)
- Also

Japanese transitive and similar cases

- Similar long-before-short preferences to model in
  - Korean PP dative (Choi 2007)
- Also
  - Heavy $\Rightarrow$ shift

Japanese transitive and similar cases

- Similar long-before-short preferences to model in
  - Korean PP dative (Choi 2007)
- Also
  - Heavy $\Rightarrow$ shift
  - Other syntactic proposals?

Next steps and open questions

Next steps and open questions

- Alternative complexity metrics

Next steps and open questions

- Alternative complexity metrics
- Alternative syntactic analyses

Next steps and open questions

- Alternative complexity metrics
- Alternative syntactic analyses
- Shift iff heavy vs. Heavy => shift

Next steps and open questions

- Alternative complexity metrics
- Alternative syntactic analyses
- Shift iff heavy vs. Heavy => shift
    - Optimization for parsing

Next steps and open questions

- Alternative complexity metrics
- Alternative syntactic analyses
- Shift iff heavy vs. Heavy => shift
    - Optimization for parsing

Next steps and open questions

- Alternative complexity metrics
- Alternative syntactic analyses
- Shift iff heavy vs. Heavy => shift
    - Optimization for parsing

(5) Niuean transitive wh-questions (Tollan et al. 2019)

a. Ko e pusi fē ne tutuli tūmau e lapiti?
PRED cat which past chase always ABS rabbit
'Which cat always chased the rabbit?'

b. Ko e pusi fē ne tutuli tūmau he kulī?
PRED cat which past chase always ERG dog
'Which cat did the dog always chased?'

References

Choi, H.-W. (2007). Length and order: A corpus study of korean dative-accusative construction. 담화와인지, 14(3):207–227.

Liu, L. (2022). *Phrasal Weight Effect on Word Order*. PhD thesis, State University of New York at Stony Brook.

Liu, Z. (2020). Mixed evidence for crosslinguistic dependency length minimization. *STUF-Language Typology and Universals*, 73(4):605–633.

Ross, J. R. (1986). *Infinite syntax*. Ablex Publishing Corporation.

Saito, M. (1992). Long distance scrambling in japanese. *Journal of East Asian Linguistics*, 1(1):69–118.

Tollan, R., Massam, D., and Heller, D. (2019). Effects of case and transitivity on processing dependencies: Evidence from niuean. *Cognitive science*, 43(6):e12736.

Wasow, T. (1997). Remarks on grammatical weight. *Language variation and change*, 9(1):81–105.

Yamashita, H. and Chang, F. (2001). "long before short" preference in the production of a head-final language. *Cognition*, 81(2):B45–B55.