# Understanding phrasal weight effect on Mandarin word order variation

from a Minimalist parsing perspective

Lei Liu

Stony Brook University

December 14, 2019 @ University College Cork

## The talk in bullet points

- Whether a BA construction [Sbj BA NP VP] is preferred in sentence processing is better understood when taking both NP and VP into consideration.
- Minimalist Grammar parsers predict a short-before-long preference for Chinese BA constructions.
- Corpus data confirms the parser's prediction.

## The talk in bullet points

- Whether a BA construction [Sbj BA NP VP] is preferred in sentence processing is better understood when taking both NP and VP into consideration.
- Minimalist Grammar parsers predict a short-before-long preference for Chinese BA constructions.
- Corpus data confirms the parser's prediction.

# Outline

# Introduction - English heavy NP shift

(1) Word order flexibility and preference: English Heavy NP Shift (HNPS)

a. i. Chris put [$_{NP}$ a box] [$_{PP}$ in the car].

ii. *Chris put [$_{PP}$ in the car] [$_{NP}$ a box].

# Introduction - English heavy NP shift

(1) Word order flexibility and preference:
English Heavy NP Shift (HNPS)

a. i. Chris put [$_{NP}$ a box] [$_{PP}$ in the car].

ii. *Chris put [$_{PP}$ in the car] [$_{NP}$ a box].

(2) Chris put [$_{PP}$ in the car] [$_{NP}$ a box filled with magical pens that help with dissertation writing].

# Introduction - English heavy NP shift

(1) Word order flexibility and preference:
English Heavy NP Shift (HNPS)

a. i. Chris put [$_{NP}$ a box] [$_{PP}$ in the car].

ii. *Chris put [$_{PP}$ in the car] [$_{NP}$ a box].

(2) Chris put [$_{PP}$ in the car] [$_{NP}$ a box filled with magical pens that help with dissertation writing].

(3) ??Chris put [$_{NP}$ a box filled with magical pens that help with dissertation writing] [$_{PP}$ in the car].

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)
  - Information structure (Liu 2007)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)
  - Information structure (Liu 2007)
- Minimalist Grammar parsing proposals for HNPS

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)
  - Information structure (Liu 2007)
- Minimalist Grammar parsing proposals for HNPS
  - procrastinate to save memory (Liu 2018)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)
  - Information structure (Liu 2007)
- Minimalist Grammar parsing proposals for HNPS
  - procrastinate to save memory (Liu 2018)

# Introduction - English heavy NP shift

- Syntactic proposals for HNPS
  - Rightward movement (Ross 1986)
  - PP movement (Kayne 1994)
  - Remnant movement (Rochemont and Culicover 1997)
  - Light predicate raising (Larson 1998; 2014)
- Human processing proposals for HNPS
  - Minimized domain (Hawkins 2004)
  - Information structure (Liu 2007)
- Minimalist Grammar parsing proposals for HNPS
  - procrastinate to save memory (Liu 2018)

**Short-before-long preference**

# Introduction - Mandarin Chinese BA vs non-BA

(4) Word order flexibility in Mandarin

a. Zhangsan ba [NP yifu daocha] fang [PP zai zhuozi shangmian]
Z. BA one.CL knife.fork put ZAI table surface

b. Zhangsan [PP zai zhuozi shangmian] fang [NP yifu daocha]
Z. ZAI table surface put one.CL knife.fork

c. Zhangsan fang [PP zai zhuozi shangmian] [NP yifu daocha]
Z. put ZAI table surface one.CL knife.fork

d. Zhangsan fang (le) [NP yifu daocha] [PP zai zhuozi
Z. put (LE) one.CL knife.fork ZAI table
shangmian]
surface
'Z put a set of knife and fork on the table.'

# Introduction - Mandarin Chinese BA vs non-BA

- Human processing preference:

# Introduction - Mandarin Chinese BA vs non-BA

- Human processing preference:
  - zoeng1 sentences have shorter reading time and higher accuracy when NP is heavy (Chung (2006) cited in Francis et al. 2011)

# Introduction - Mandarin Chinese BA vs non-BA

- Human processing preference:
  - zoeng1 sentences have shorter reading time and higher accuracy when NP is heavy (Chung (2006) cited in Francis et al. 2011)
  - ba sentences are preferred when NP is “very” short or “very” long (Yao 2018)

# Introduction - Mandarin Chinese BA vs non-BA

- Human processing preference:
  - zoeng1 sentences have shorter reading time and higher accuracy when NP is heavy (Chung (2006) cited in Francis et al. 2011)
  - ba sentences are preferred when NP is “very” short or “very” long (Yao 2018)

# Introduction - Mandarin Chinese BA vs non-BA

- Human processing preference:
  - zoeng1 sentences have shorter reading time and higher accuracy when NP is heavy (Chung (2006) cited in Francis et al. 2011)
  - ba sentences are preferred when NP is “very” short or “very” long (Yao 2018)

**What does an Minimalist Parser predict?**

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
    - lexical items

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations
    - `merge, move`

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations
    - `merge, move`

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations
    - `merge, move`

(5) Max packed boxes.

- Max :: $D^-$ $nom^-$
  `pronunciation` `category` `movement`

# Minimalist Grammar

- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations
    - `merge, move`

(5) Max packed boxes.

- Max :: $D^-$ $nom^-$
  `pronunciation` `category` `movement`
- packed :: $D^+$ $D^+$ $V^-$
  `pronunciation` `selection` `selection` `category`

# Minimalist Grammar

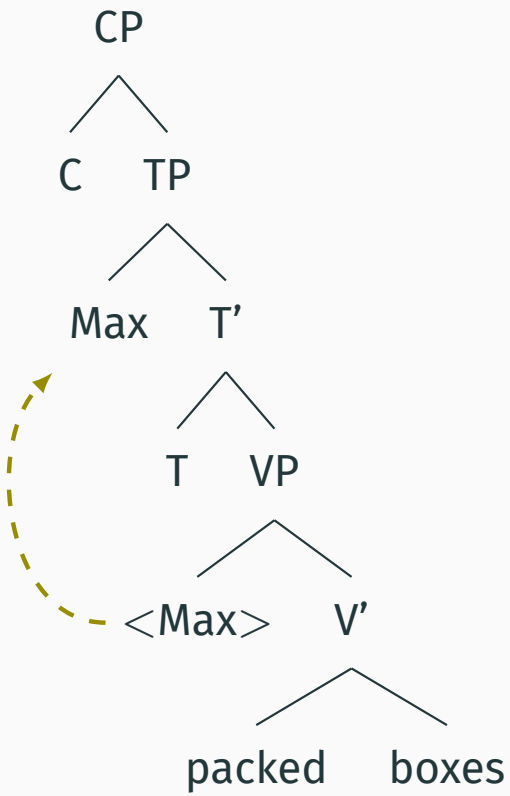
- Minimalist Grammars (Stabler 1996; 2011, Graf 2013): grammar formalisms based on the Minimalist Program (Chomsky 2014).
  - lexical items
    - feature bundles *pronunciation*, *category*, *movement*
  - operations
    - `merge, move`

(5) Max packed boxes.

- Max :: $D^-$ $nom^-$
  `pronunciation` `category` `movement`
- packed :: $D^+$ $D^+$ $V^-$
  `pronunciation` `selection` `selection` `category`
- boxes :: $D^-$
  `pronunciation` `category`

# Minimalist Grammar

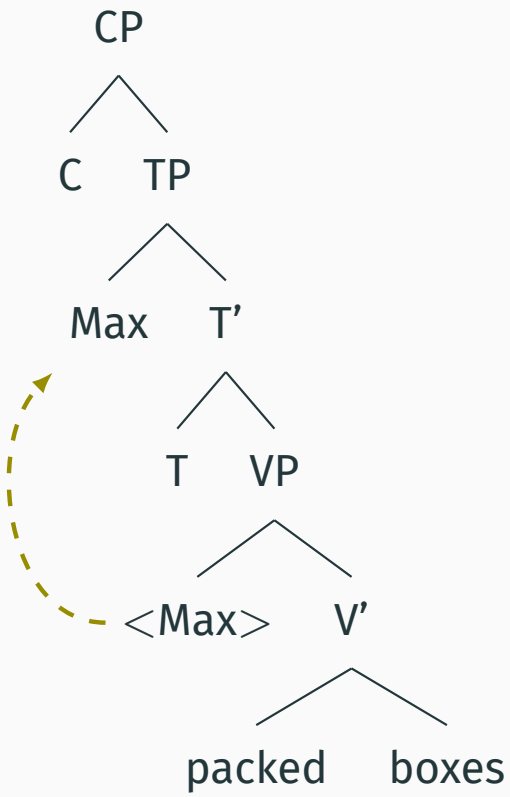
(3) Max packed boxes.

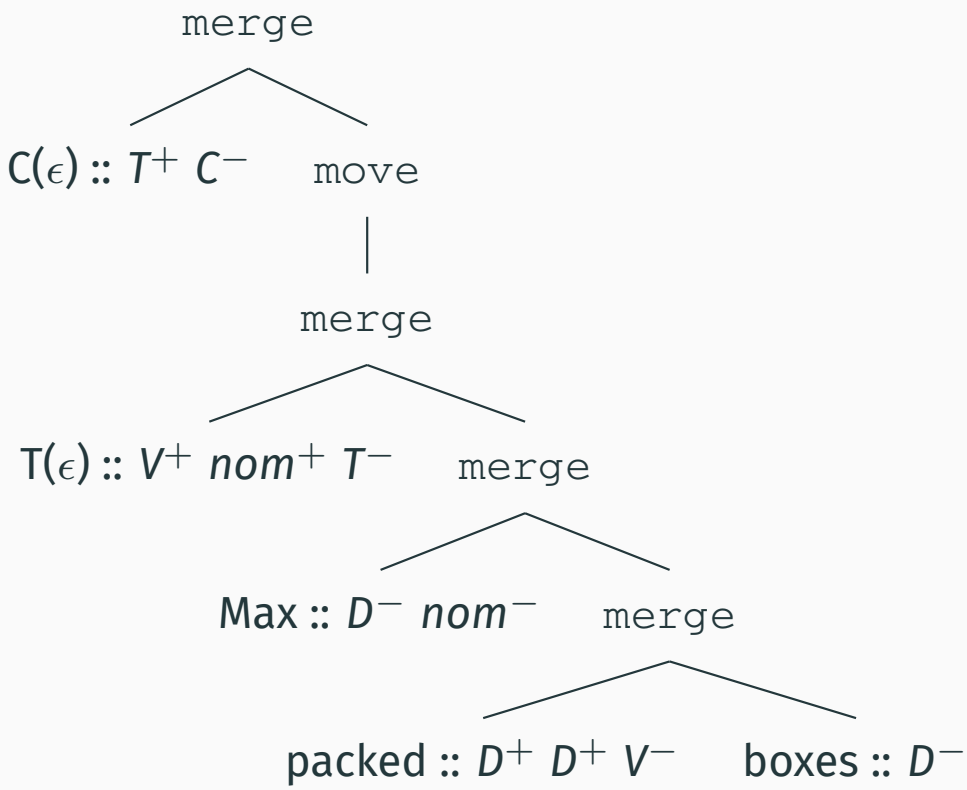



*derived tree*

# Minimalist Grammar

(3) Max packed boxes.




*derived tree*

*derivation tree*

# MG parser

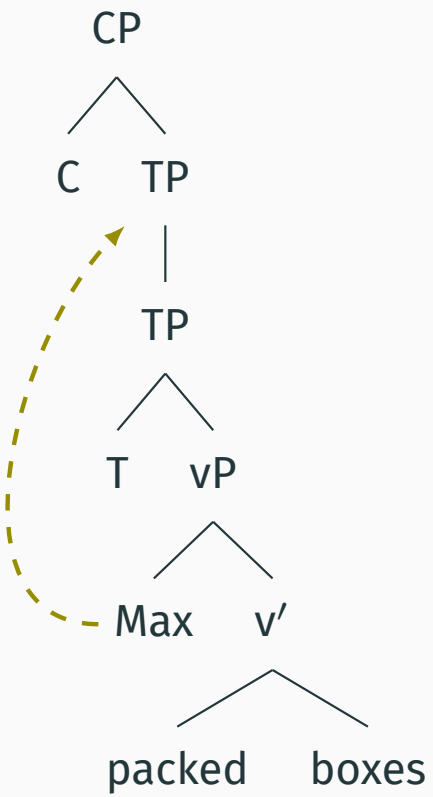
An MG parser (Stabler 2013, Graf et al. 2015a) is a recursive-descent parsers that:

- takes as input a string with pronounced and unpronounced nodes and,
- outputs derivation trees

# MG - parser

An annotated tree is a record of a parser's behavior

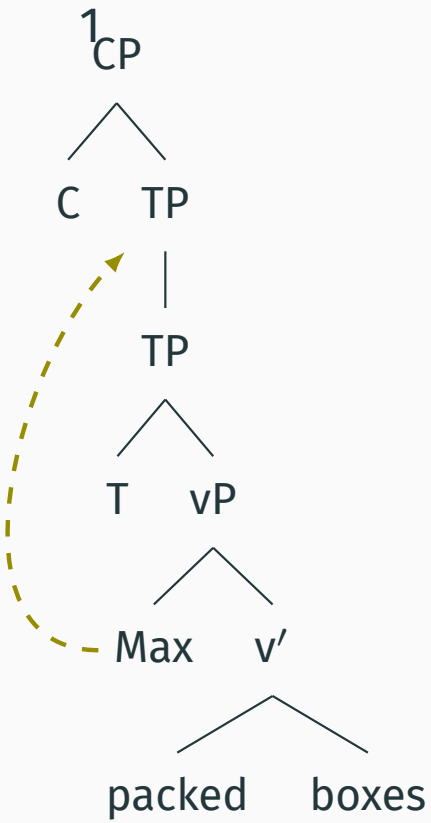
(3) ▷ C Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

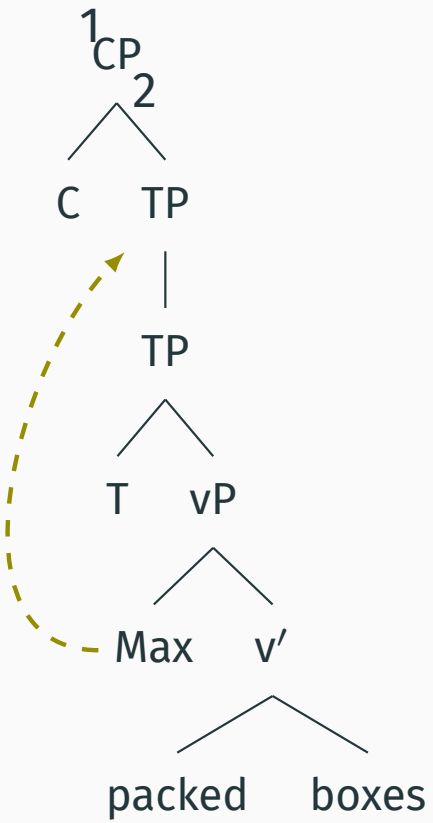
(3) ▷ C Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

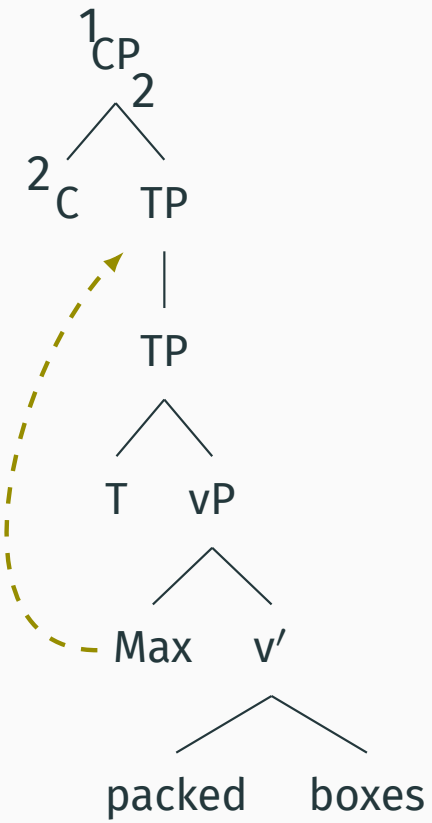
(3) ▷ C Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

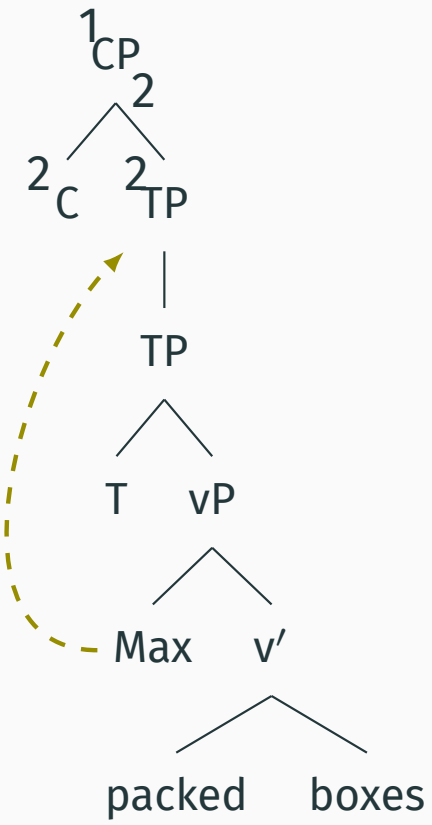
(3) ▷ C Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

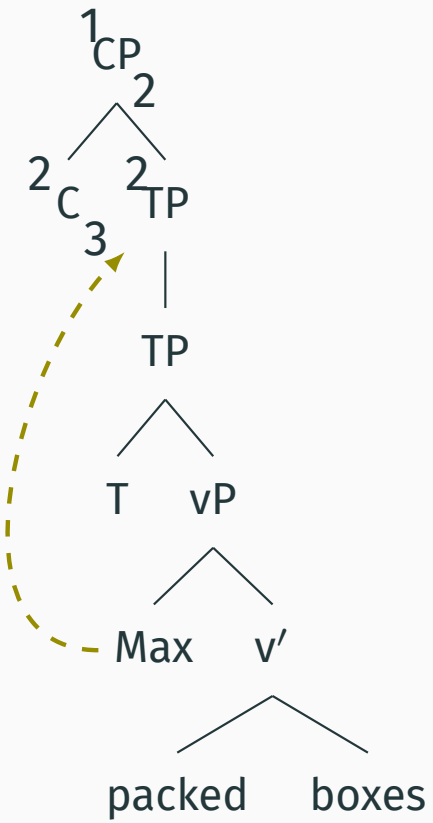
(3) ▷ C Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

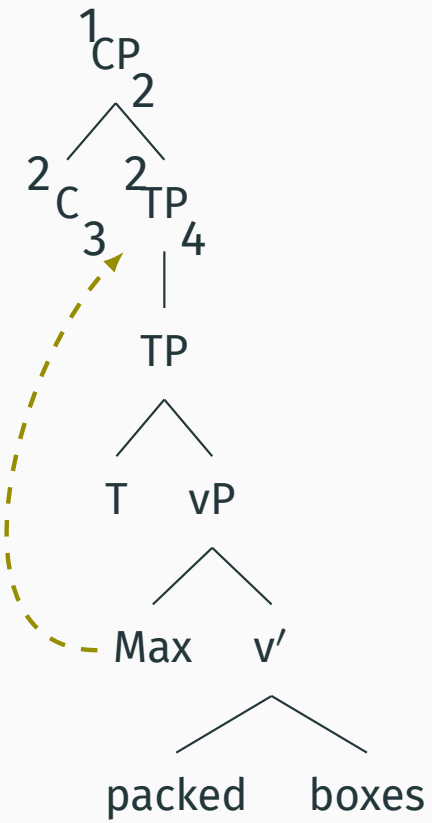
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

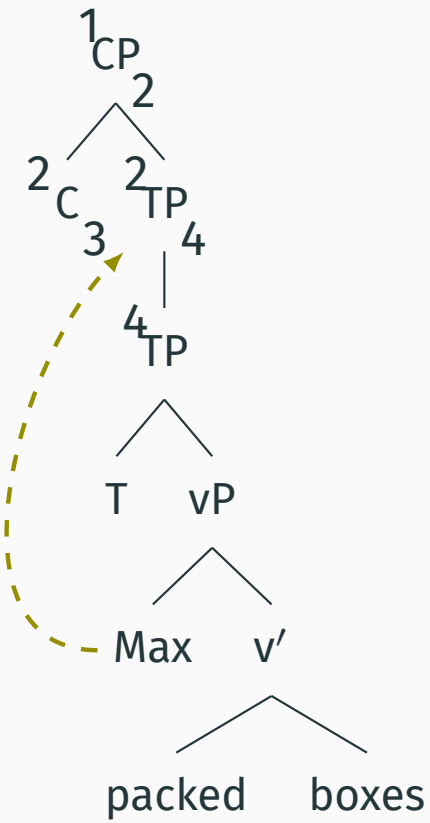
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

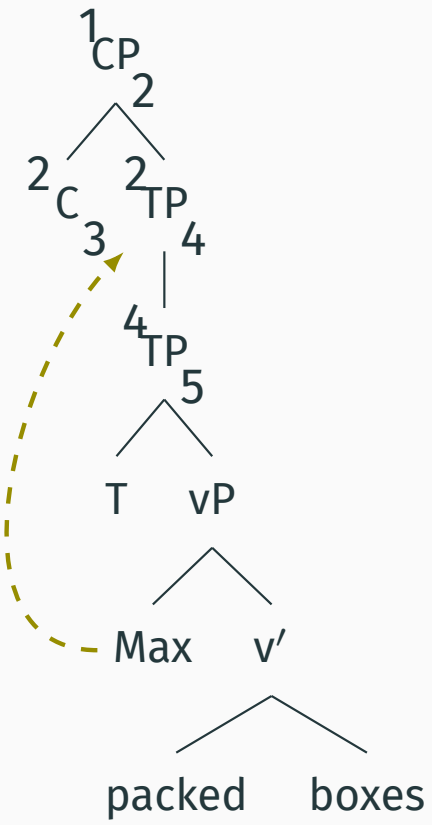
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

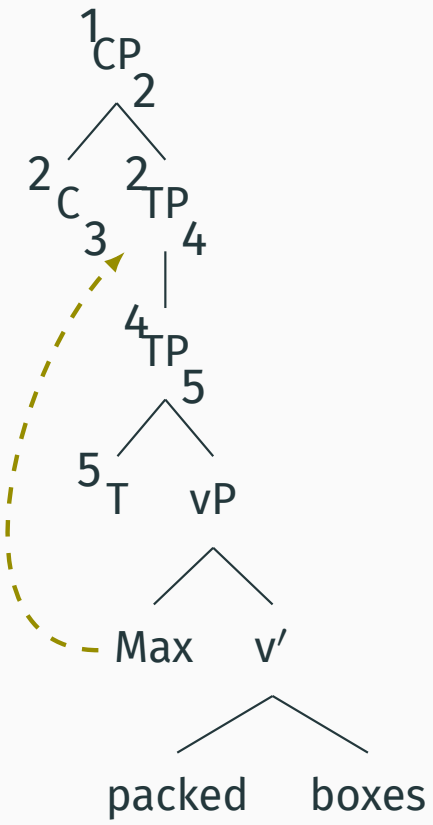
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

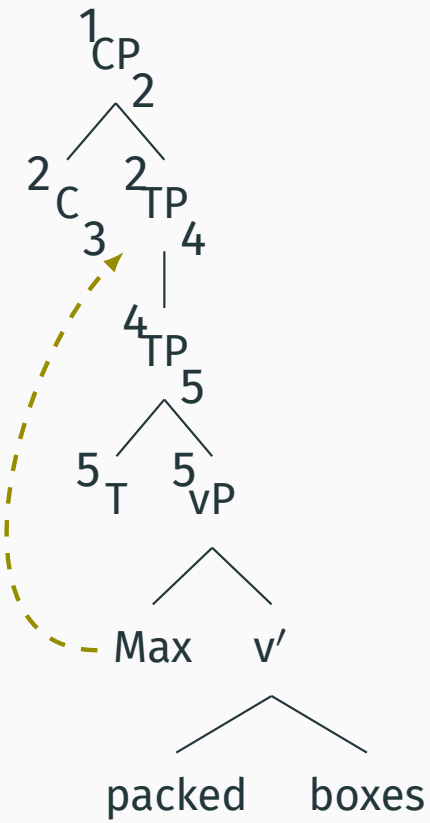
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

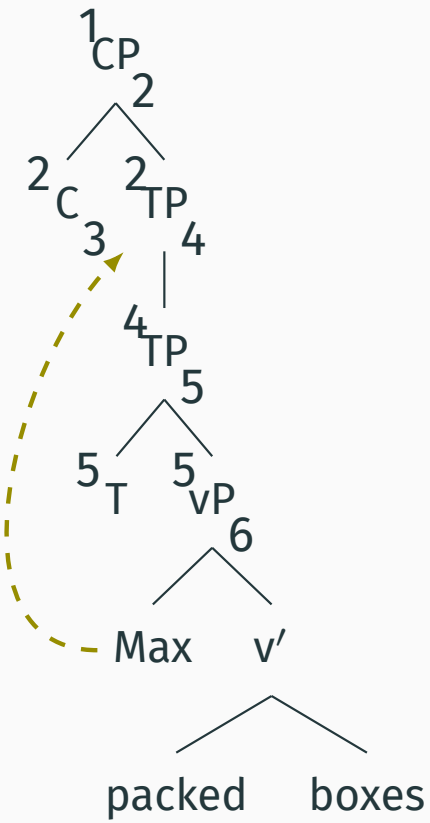
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

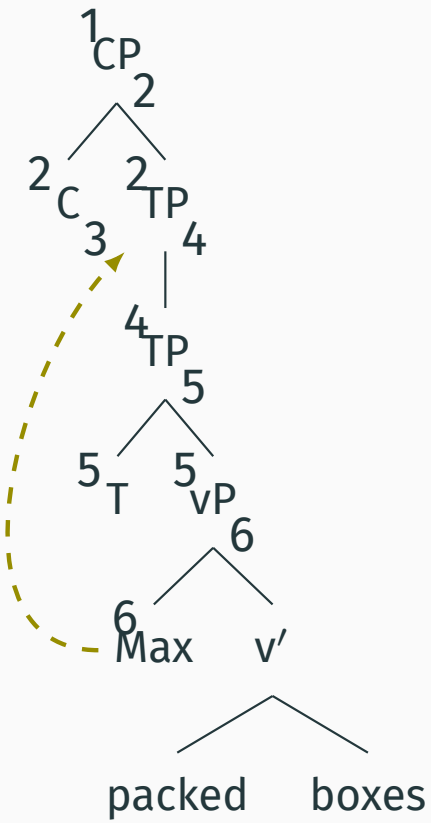
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

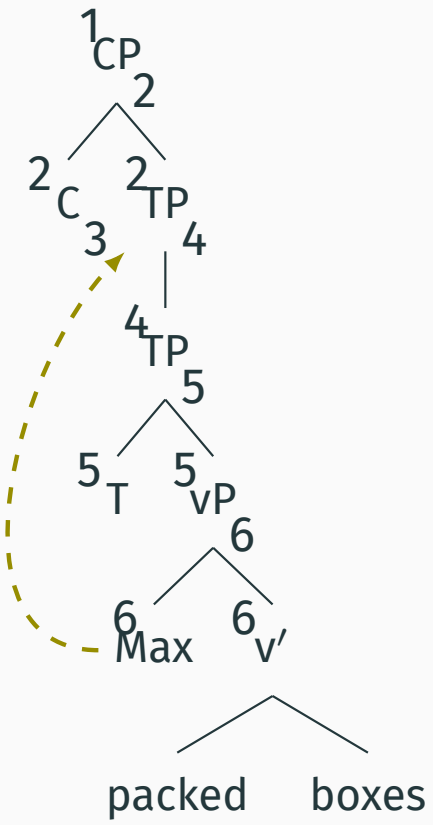
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

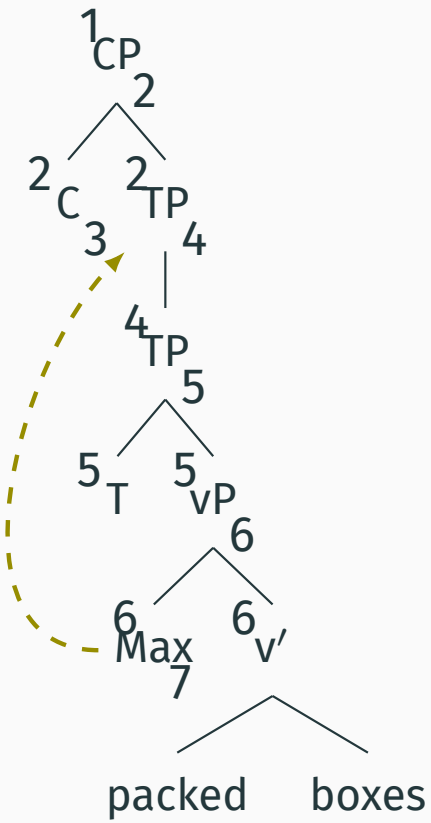
(3) C ▷ Max T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

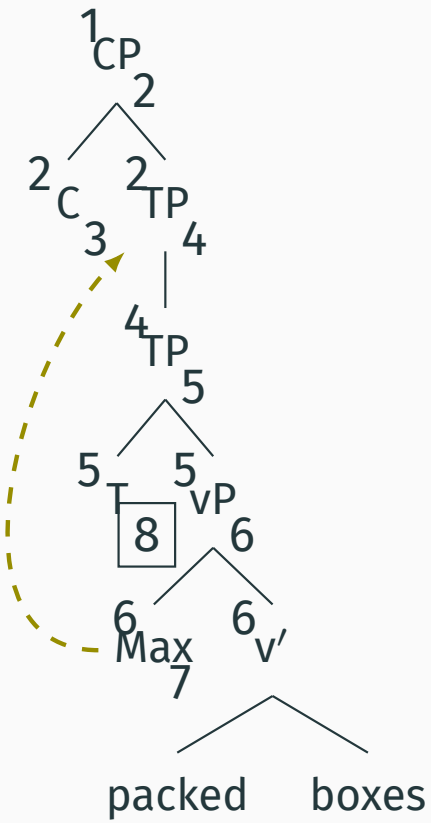
(3) C Max ▷ T packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

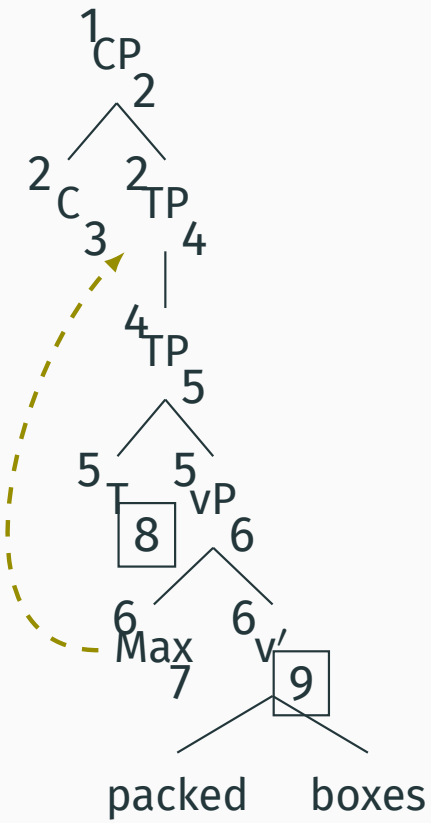
(3) C Max T ▷ packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

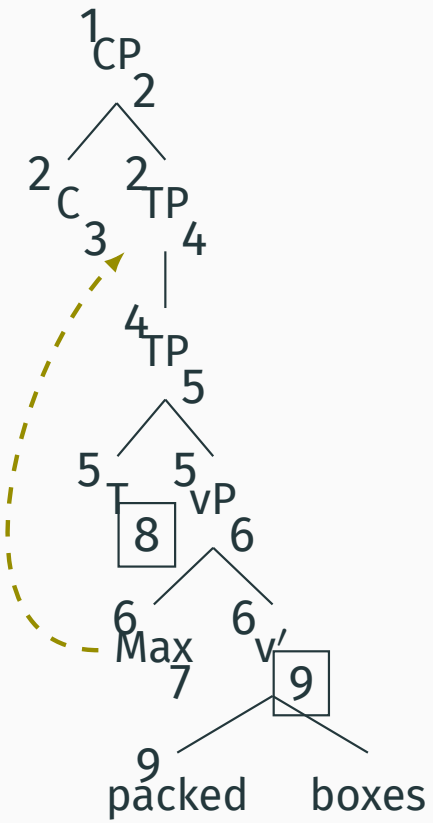
(3) C Max T ▷ packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

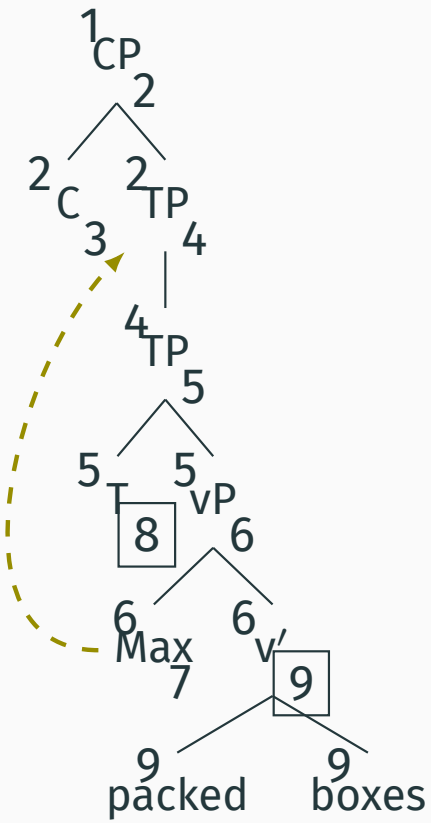
(3) C Max T ▷ packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

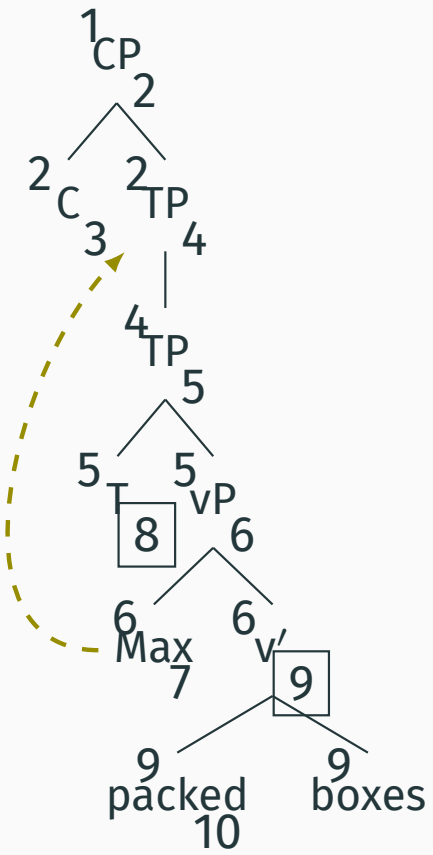
(3) C Max T ▷ packed boxes.



# MG - parser

An annotated tree is a record of a parser's behavior

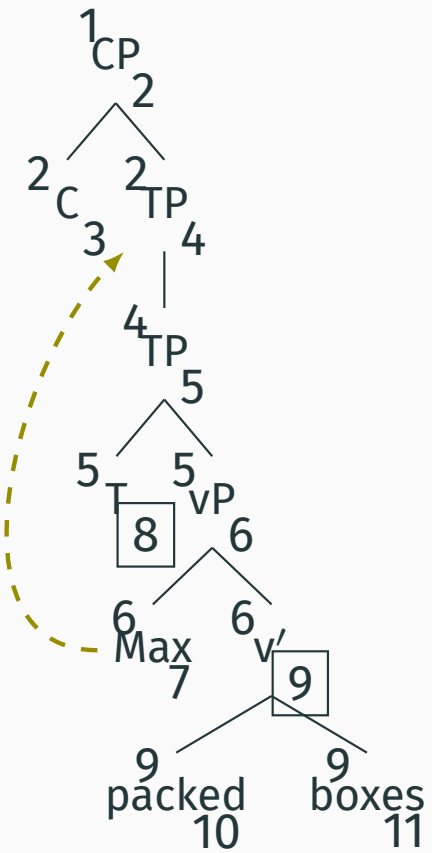
(3) C Max T packed ▷ boxes.



# MG - parser

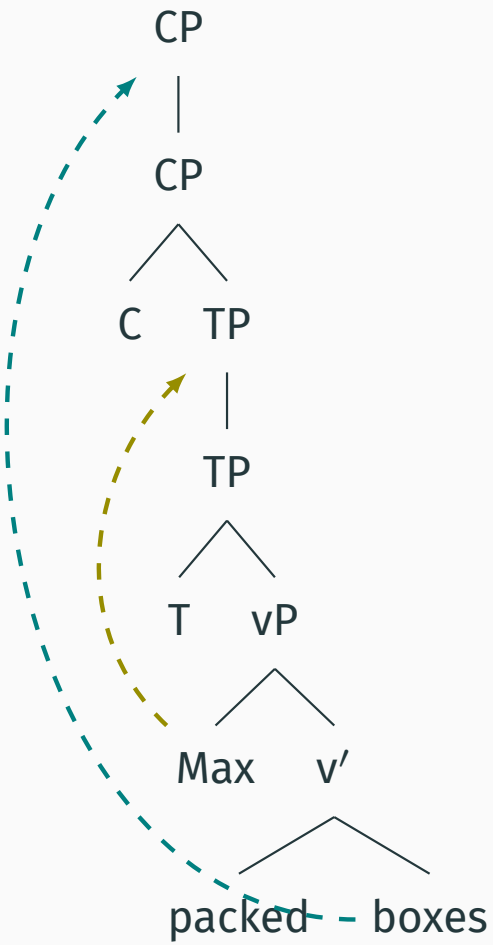
An annotated tree is a record of a parser's behavior

(3) C Max T packed boxes. ⋉



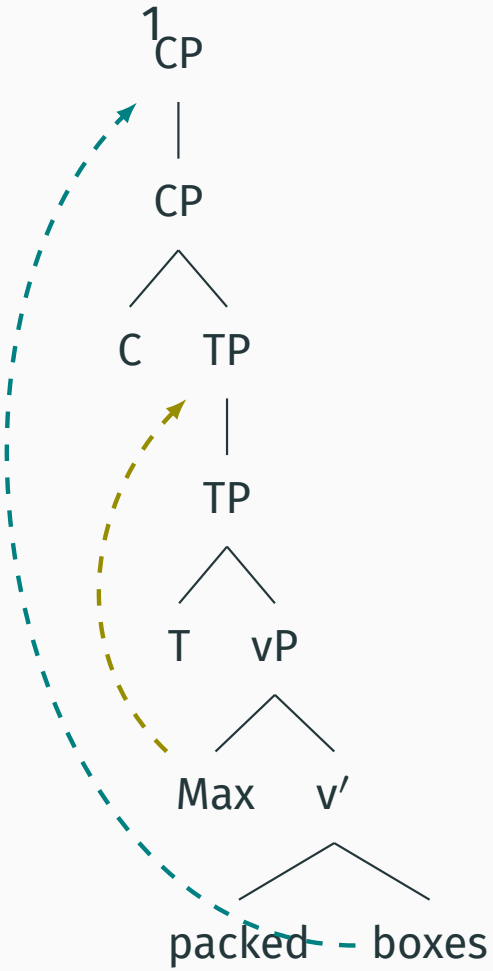
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



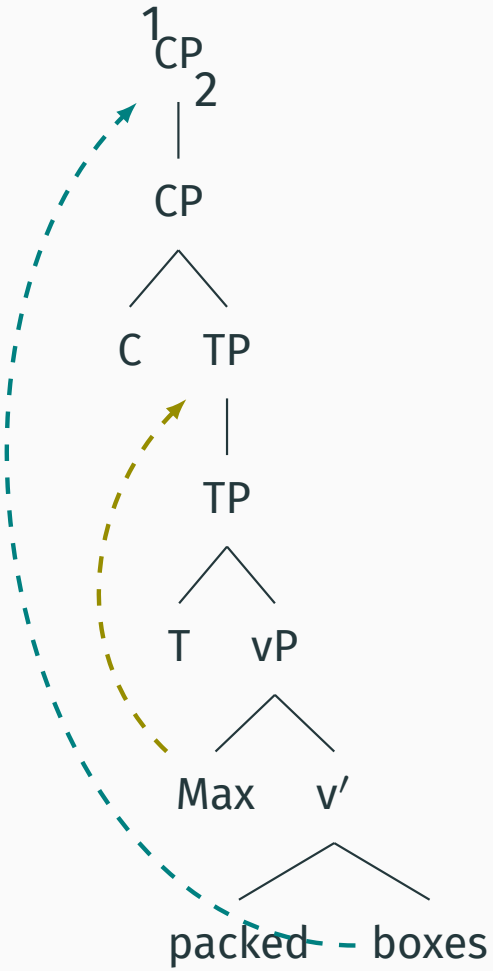
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



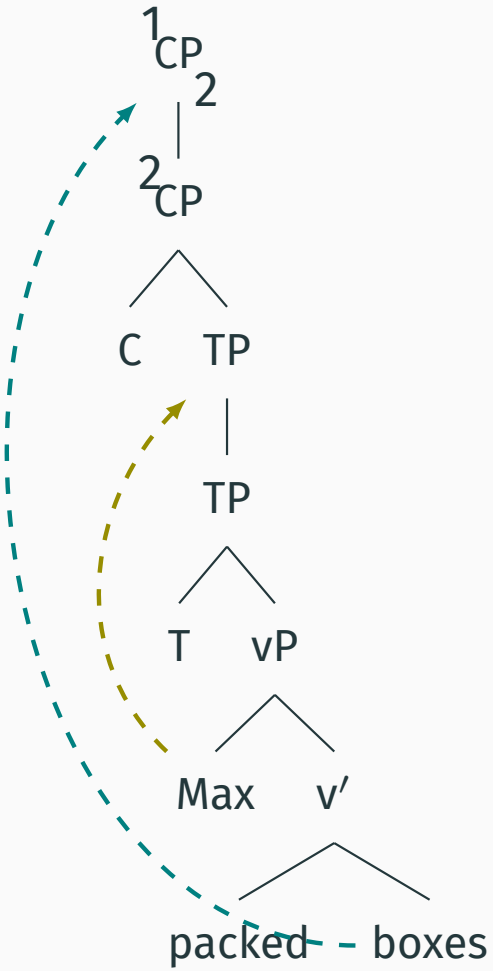
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



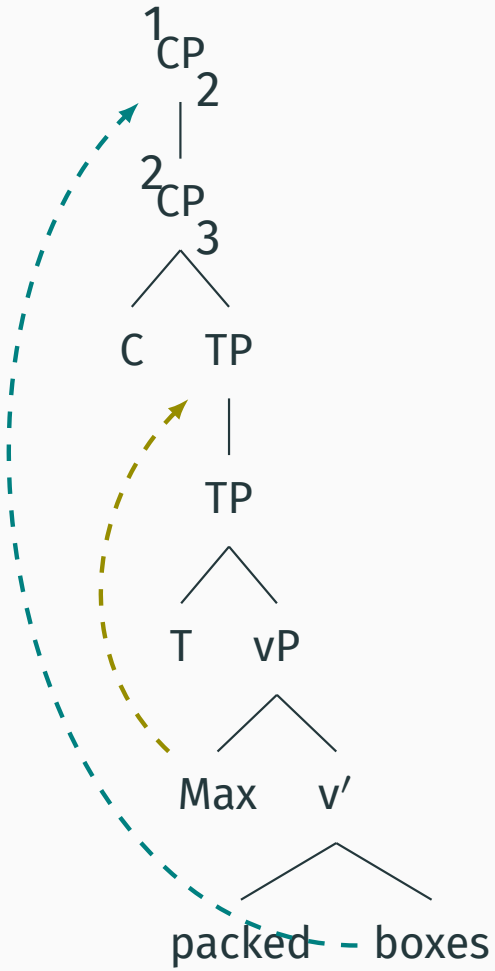
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



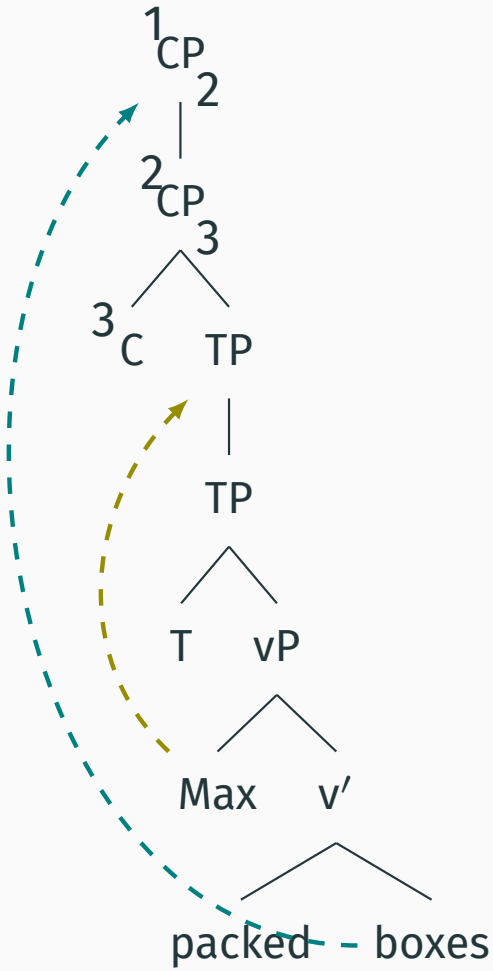
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



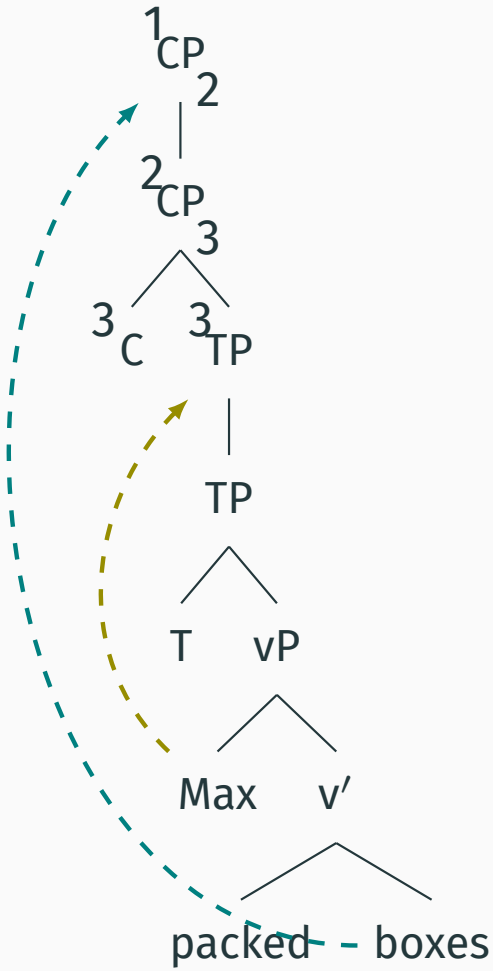
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



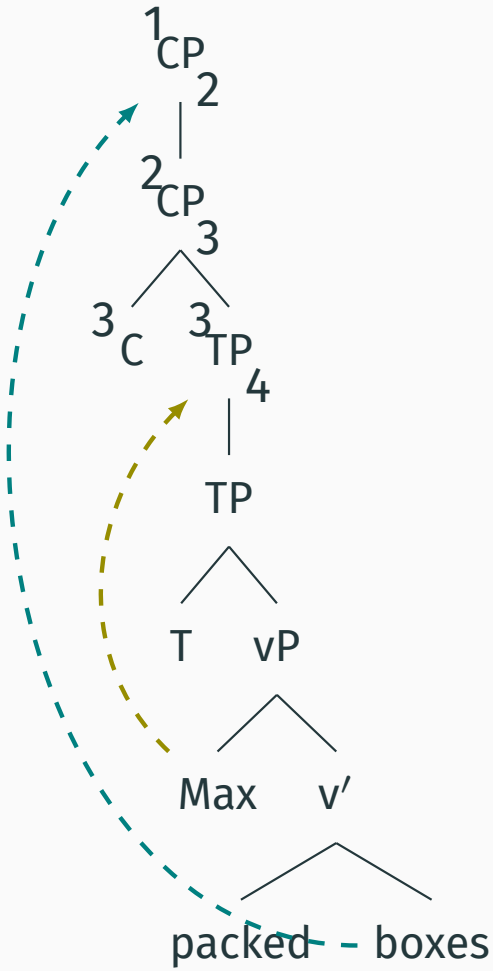
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



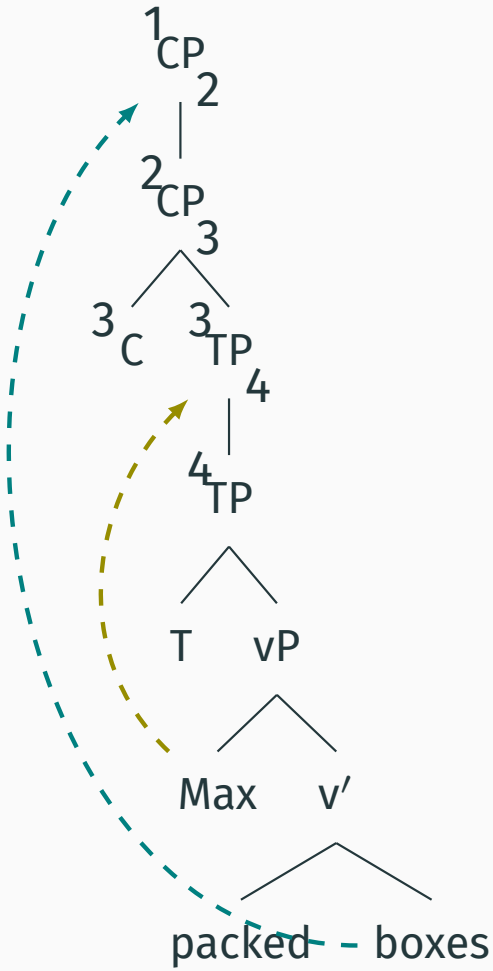
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



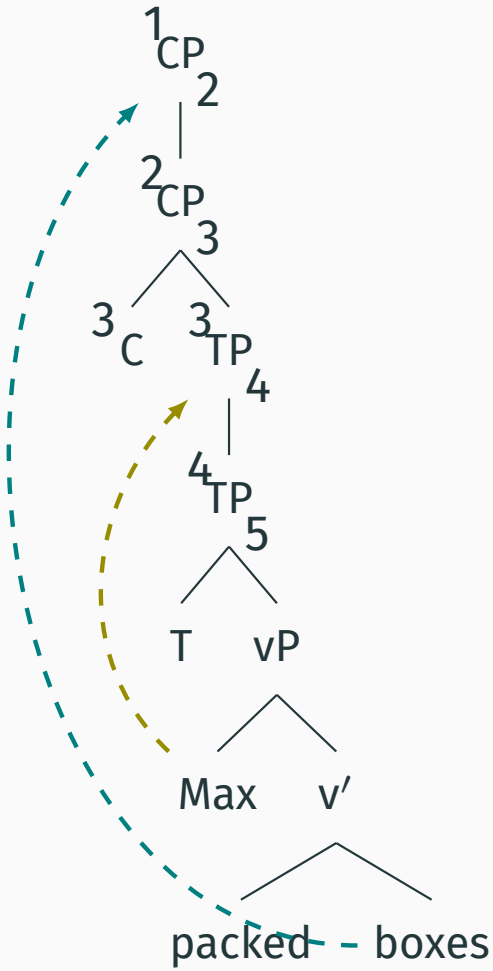
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



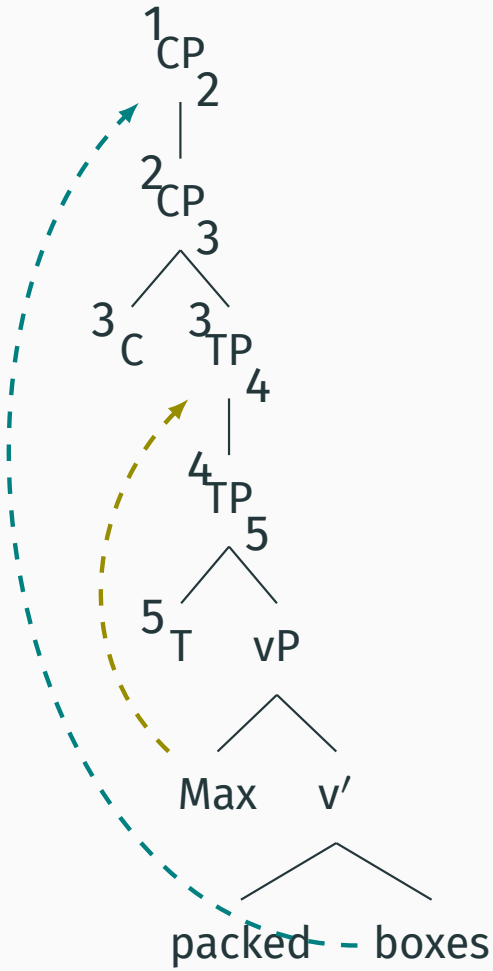
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



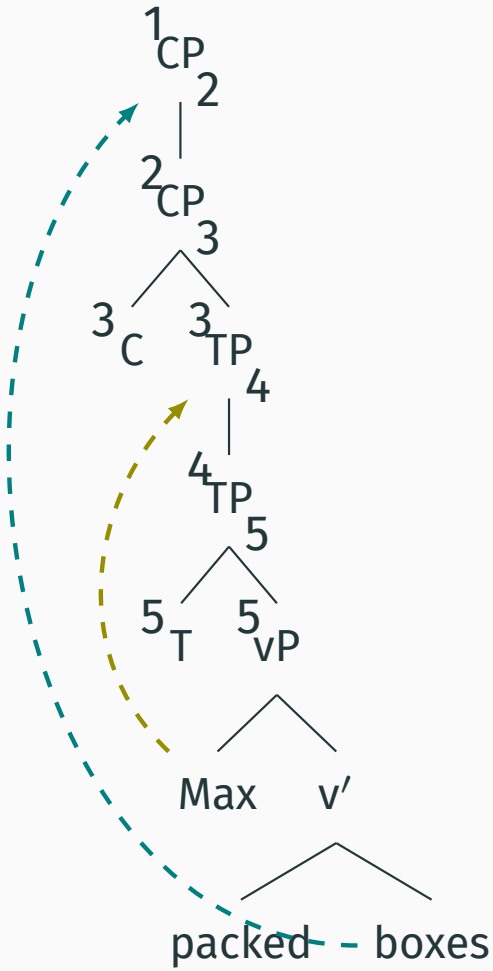
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



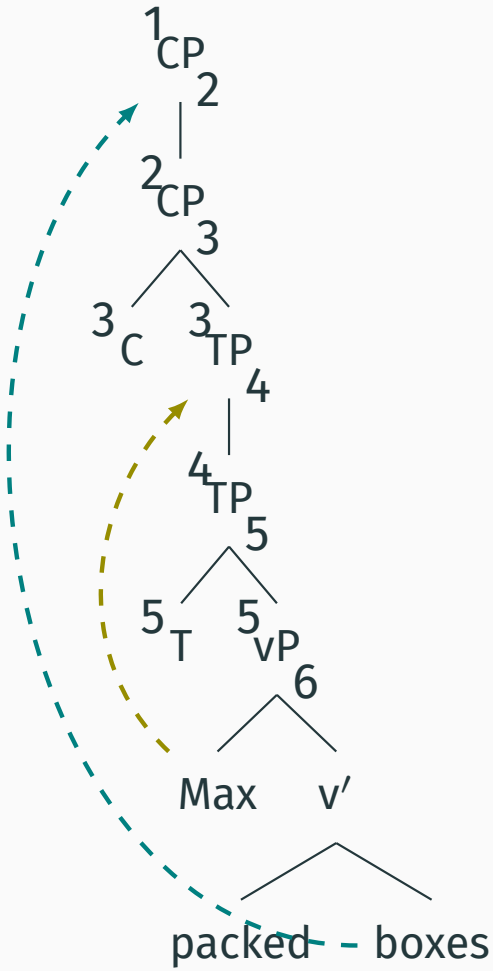
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



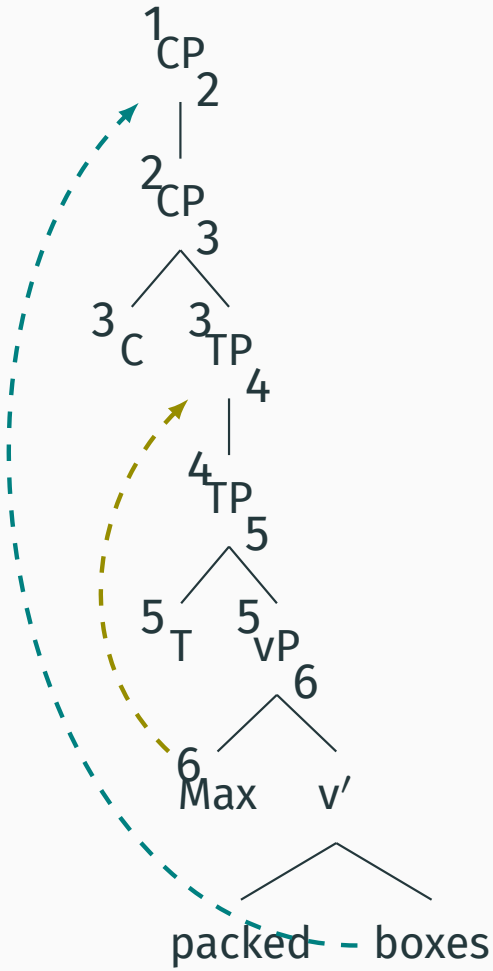
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



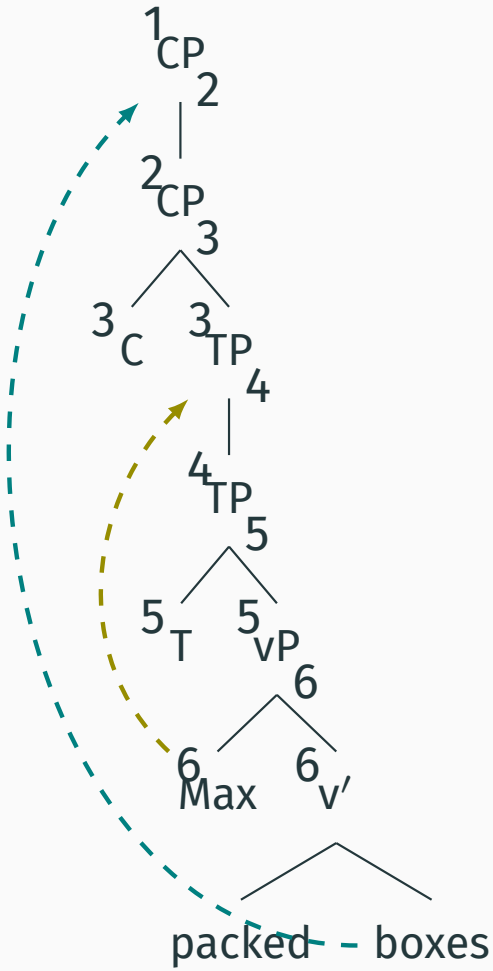
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



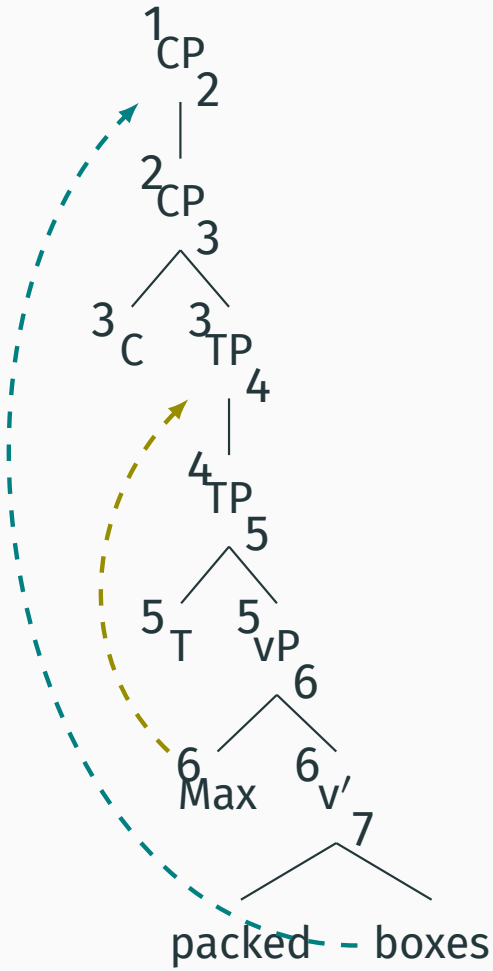
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



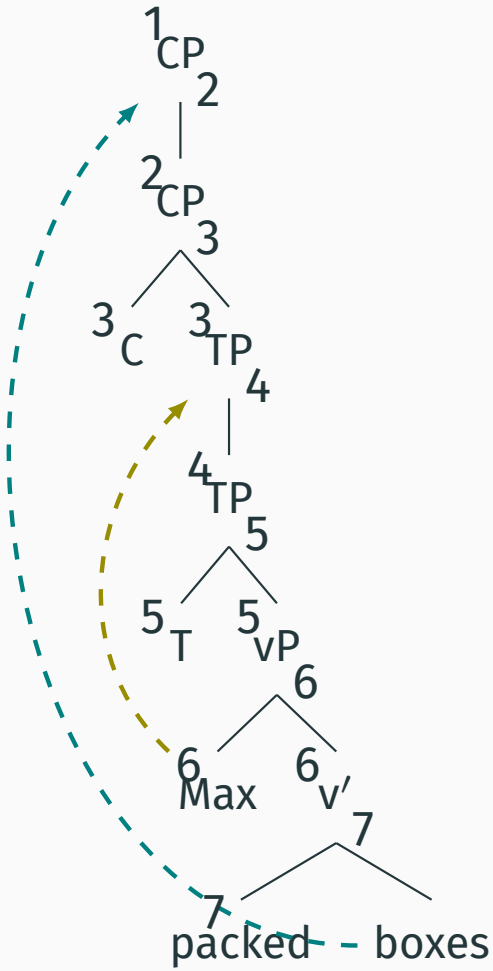
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



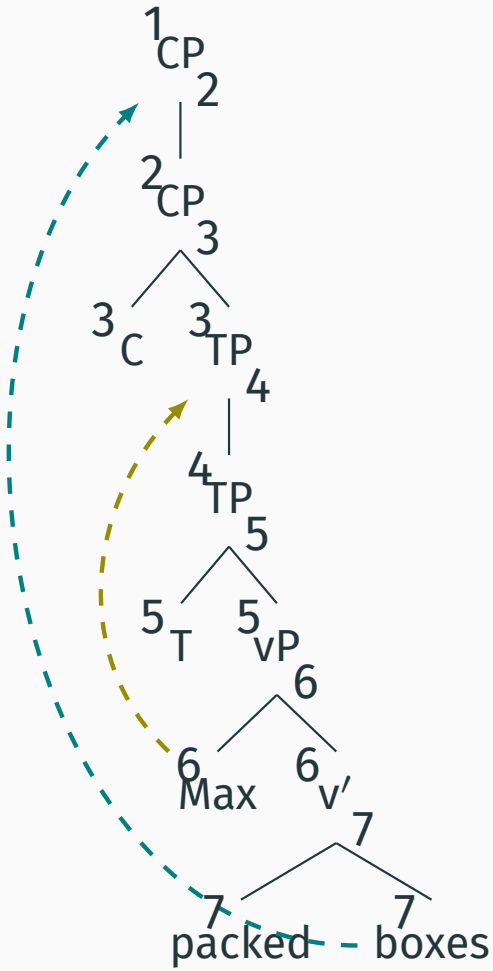
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



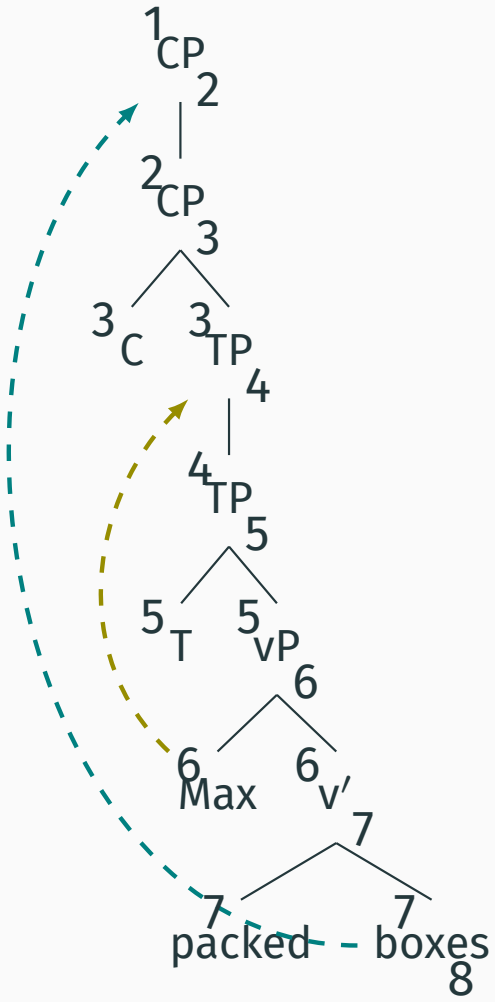
# MG parser

(4) ▷ Boxes C, Max T packed *t*.



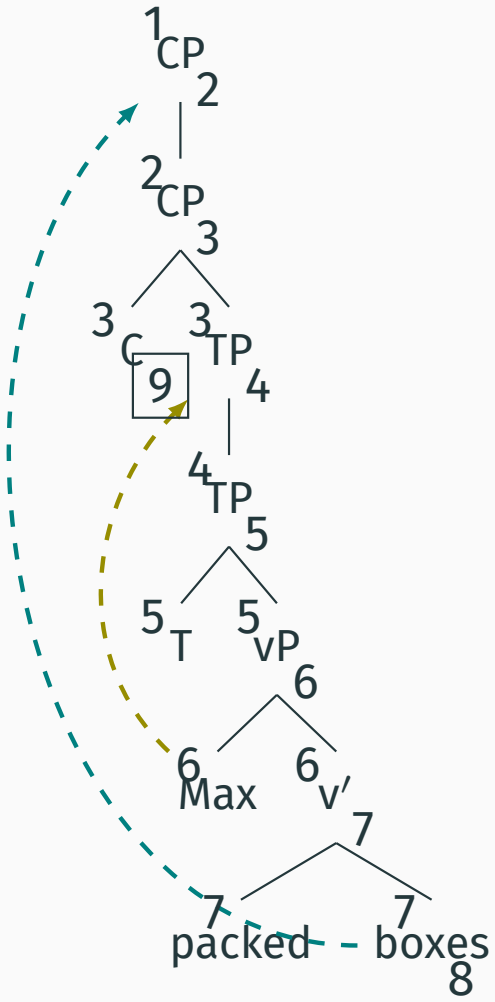
# MG parser

(4) Boxes ▷ C, Max T packed *t*.



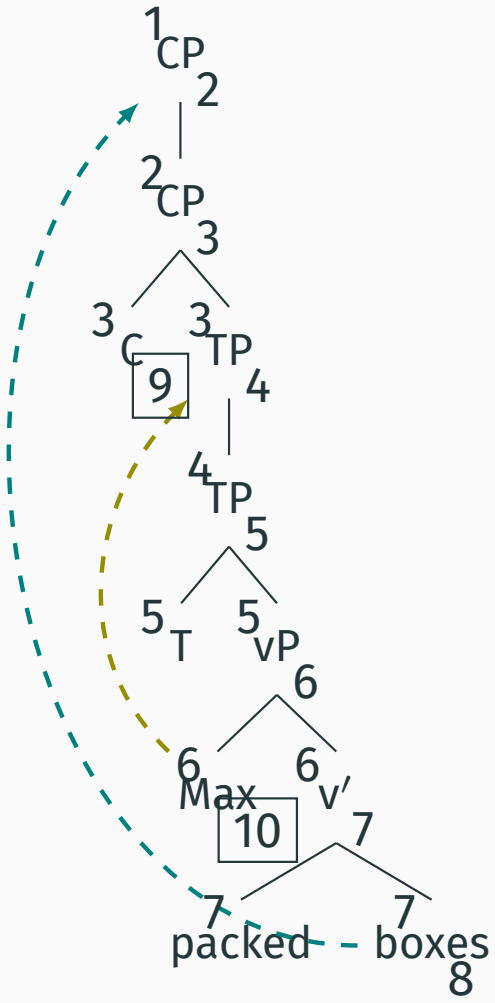
# MG parser

(4) Boxes C, ▷ Max T packed *t*.



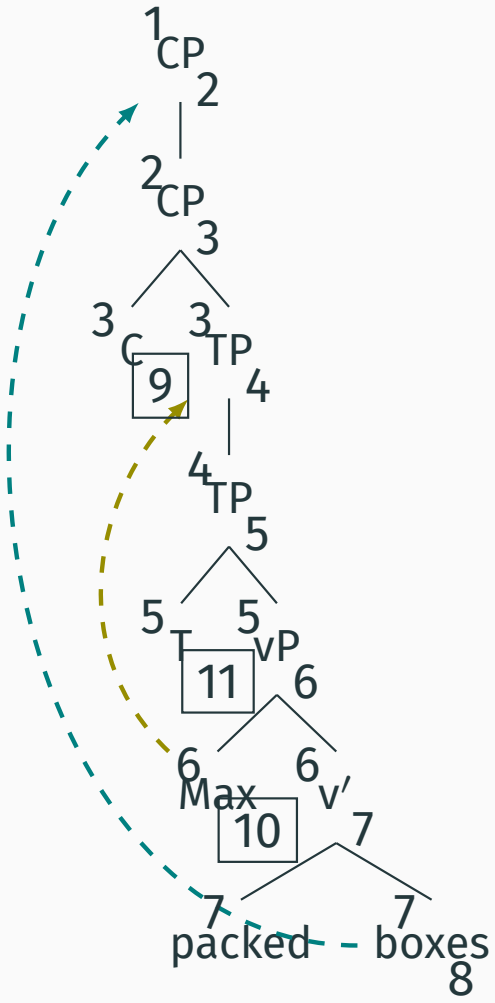
# MG parser

(4) Boxes C, Max ▷ T packed *t*.



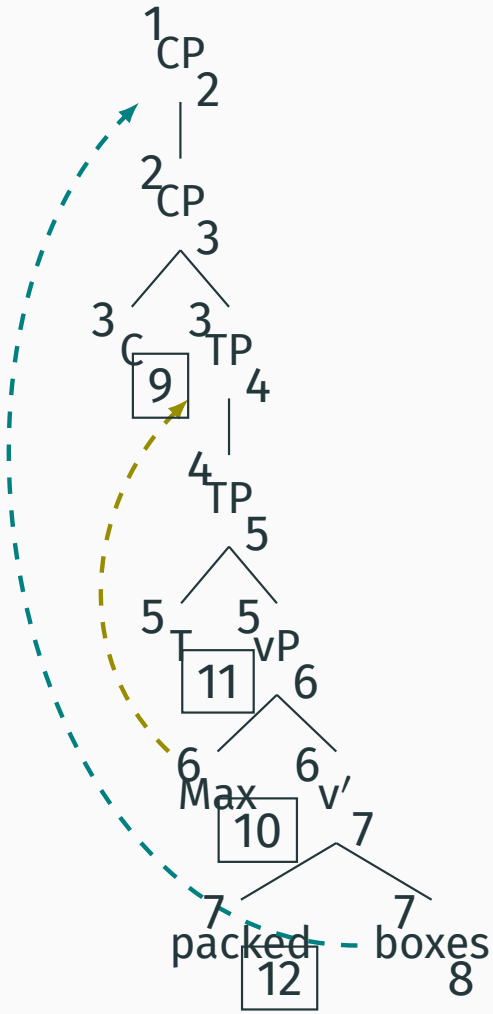
# MG parser

(4) Boxes C, Max T ▷ packed *t*.

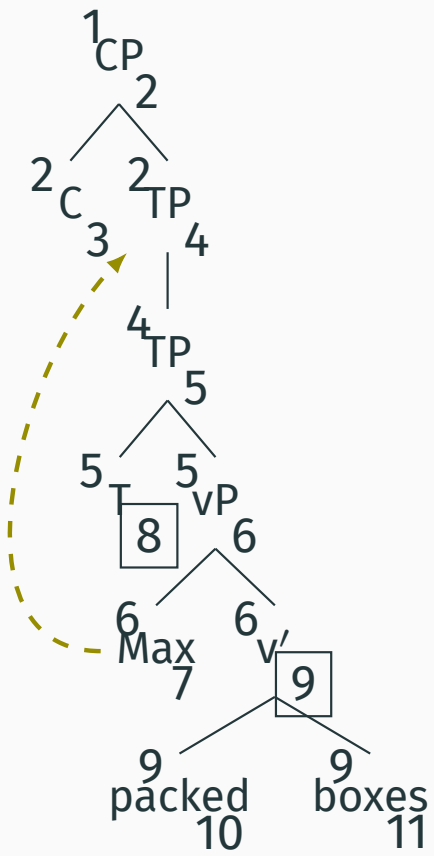


# MG parser

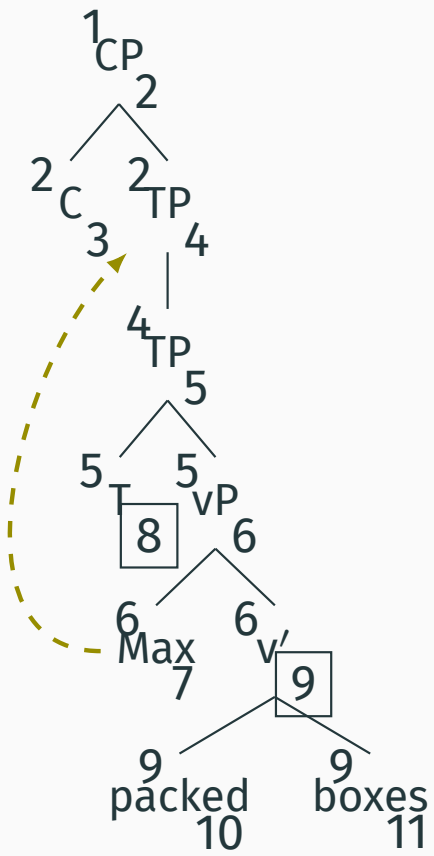
(4) Boxes C, Max T packed $t_{\ltimes}$.



# MG - complexity metrics

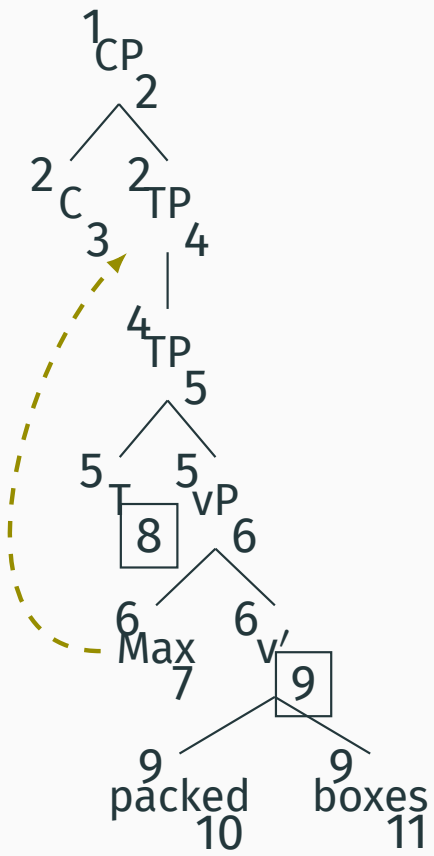


# MG - complexity metrics



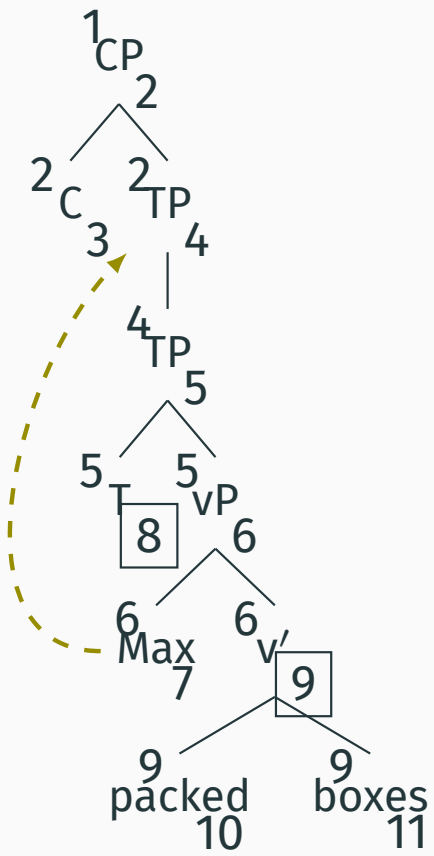

- Complexity metrics (Graf et al. 2015b)

# MG - complexity metrics



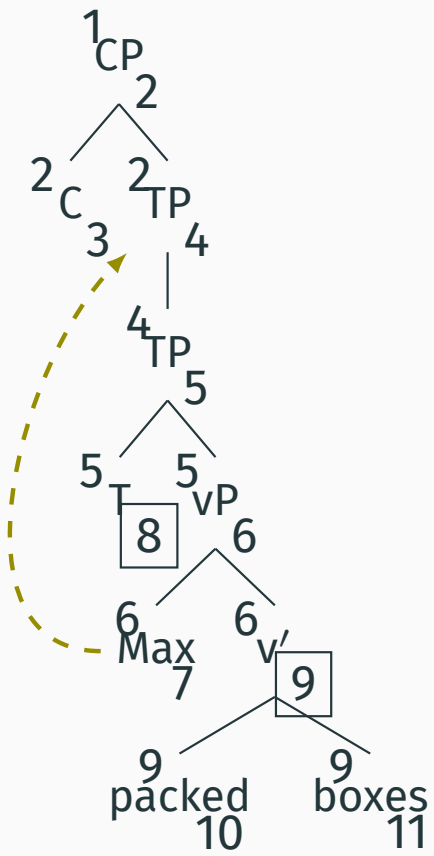

- Complexity metrics (Graf et al. 2015b)
  - `Tenure`: how long a parse item is held in memory

# MG - complexity metrics



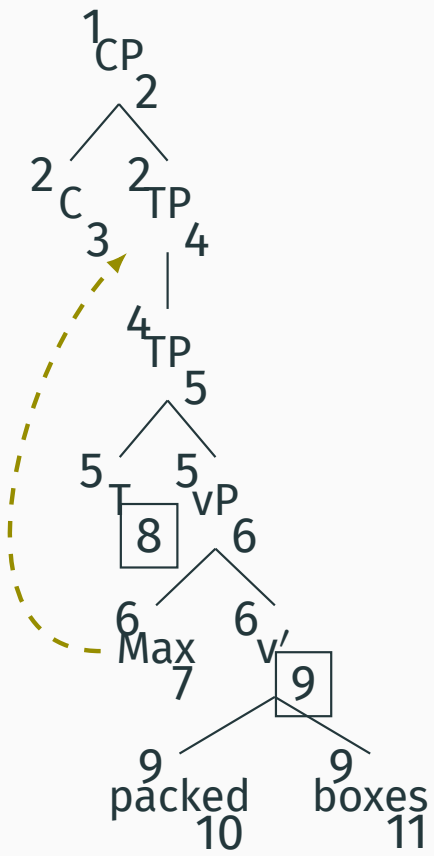

- Complexity metrics (Graf et al. 2015b)
  - `Tenure`: how long a parse item is held in memory
  - `Payload`: how many items are held in memory throughout a parse

# MG - complexity metrics



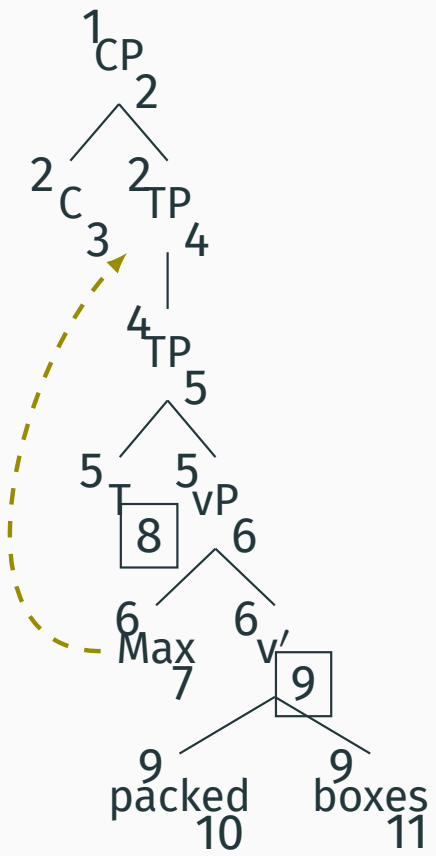

- Complexity metrics (Graf et al. 2015b)
  - `Tenure`: how long a parse item is held in memory
  - `Payload`: how many items are held in memory throughout a parse
  - `Size`: the length of movement dependencies

# MG - complexity metrics




- Complexity metrics (Graf et al. 2015b)
  - `Tenure`: how long a parse item is held in memory
  - `Payload`: how many items are held in memory throughout a parse
  - `Size`: the length of movement dependencies
  - a total of 20 metrics were used

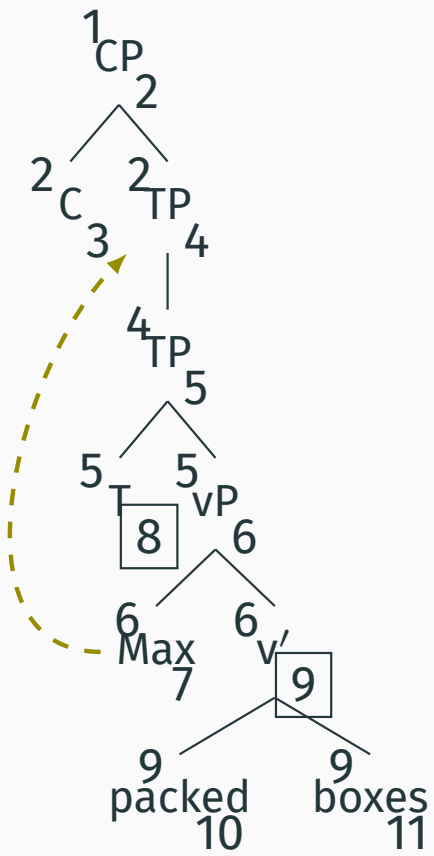
# MG - complexity metrics



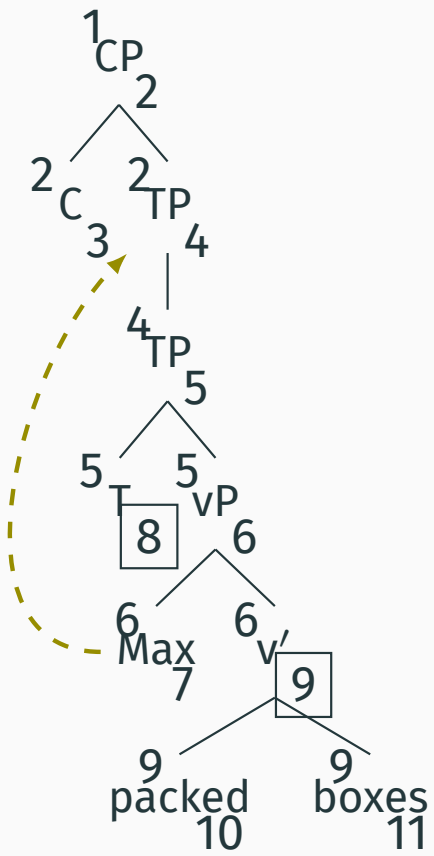

- Complexity metrics (Graf et al. 2015b)
  - `Tenure`: how long a parse item is held in memory
  - `Payload`: how many items are held in memory throughout a parse
  - `Size`: the length of movement dependencies
  - a total of 20 metrics were used

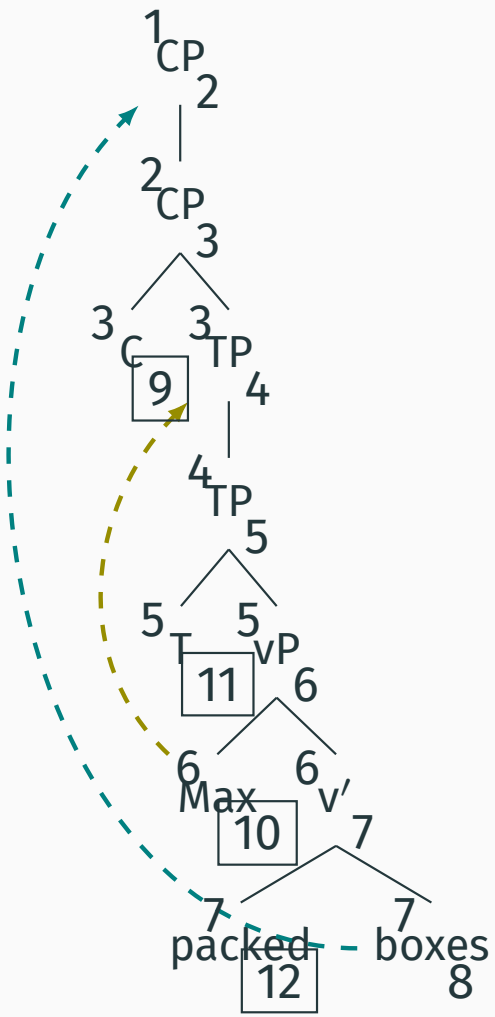
| example metrics | in this tree |
|---|---|
| **MaxT** | $3 = 9 - 6 (= 8 - 5)$ |
| **SumT** | $6 = 3 + 3$ |
| **BoxT** | 2 on $T, v'$ |
| **AvgT** | $3 = \frac{6}{2}$ ($\frac{\text{SumT}}{\text{BoxT}}$) |

# Comparisons



# Comparisons



# Comparisons

- Minimalist parsing models
  - operate on Minimalist Grammars (Stabler 2011, Graf 2012)
  - replicate sentence processing difficulties for human
  - have been shown to capture

| | |
|---|---|
| Center embeddings vs. Left/right embedding | Graf et al. (2017) |
| Inverse scope readings vs. Surface scope readings | Pasternak and Graf (in prep) |
| Objective relative clauses vs. Subject rel. clauses | Zhang (2017) |
| English heavy NP shift | Liu (2018) |
| Gradient Acceptability in Syntactic Islands | De Santo (to appear) |

# Comparisons

- What does a Minimalist parser predict?
  - RQ: When compared with a non-BA construction, is a BA construction predicted to be preferred with varying NP and PP condition?

# Comparisons

- What does a Minimalist parser predict?
  - RQ: When compared with a non-BA construction, is a BA construction predicted to be preferred with varying NP and PP condition?

Comparisons are conducted...

- with `mgproc`
- by specifying
  - target sentences
  - syntactic structure
  - human processing bias

# Comparisons - target sentences

- Pair-wise comparison were made between two word orders
  - BA: [Sbj BA NP (V PP)]
  - non-BA: [Sbj (PP V) NP]
- The target sentences were controlled for DP (short, long, rel), PP (short, long), and BA (yes, no) ($3 * 2 * 2 = 12$)

# Comparisons - target sentences

- Pair-wise comparison were made between two word orders
  - BA: [Sbj BA NP (V PP)]
  - non-BA: [Sbj (PP V) NP]
- The target sentences were controlled for DP (short, long, rel), PP (short, long), and BA (yes, no) ($3 * 2 * 2 = 12$)

(5) short DP short PP no BA

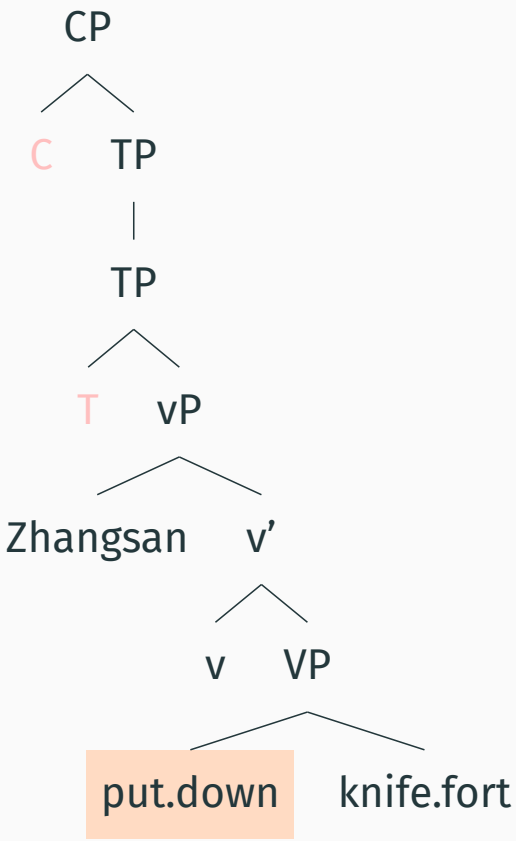
Zhangsan fang [$_{PP}$ xia] [$_{DP}$ daocha]
Z. put down knife.fork

(6) rel DP long PP yes BA

Zhangsan ba [$_{DP}$ yong guijinshu zhizuo de daocha] fang [$_{PP}$ zai yong jishu zhizuo de zhuozi shangmian]
Z ba use precious.metal make DE knife.fork put ZAI use metal make DE table surface

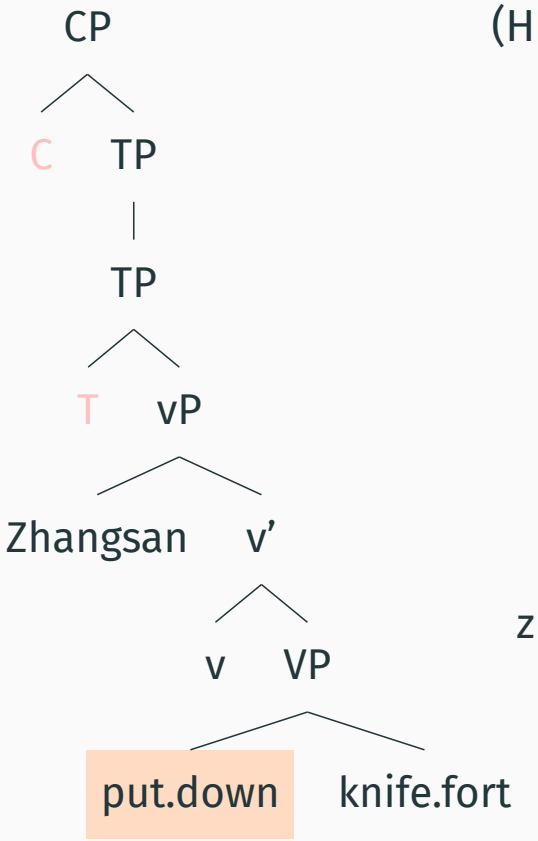
# Comparisons - syntactic structures

(7) short PP



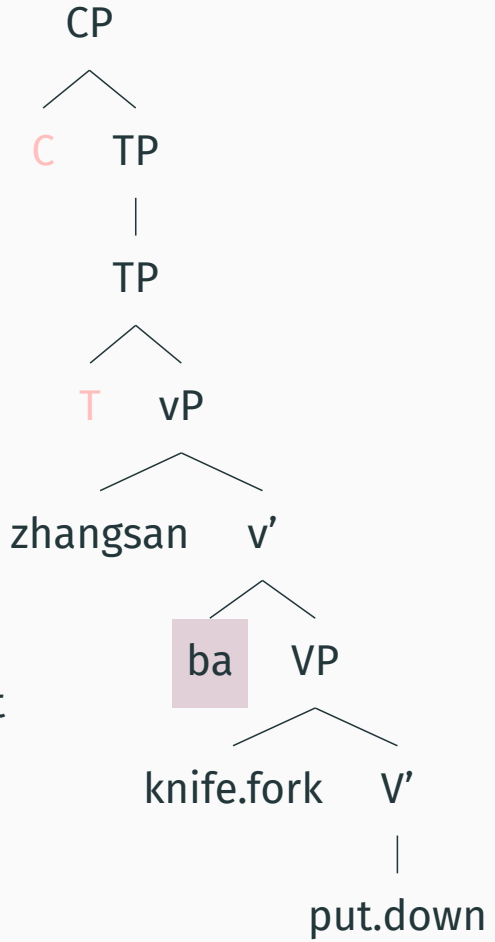
# Comparisons - syntactic structures

(7) short PP



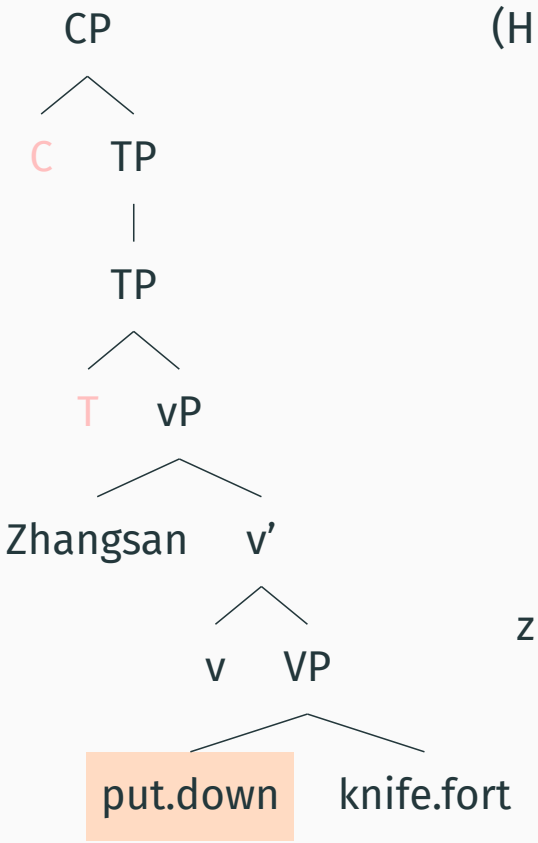

(8) BA

(Huang et al. 2009)

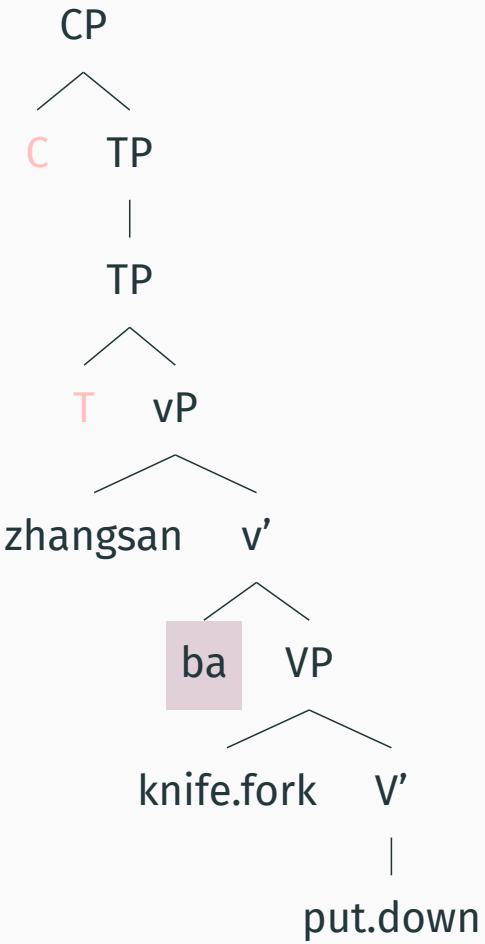


# Comparisons - syntactic structures

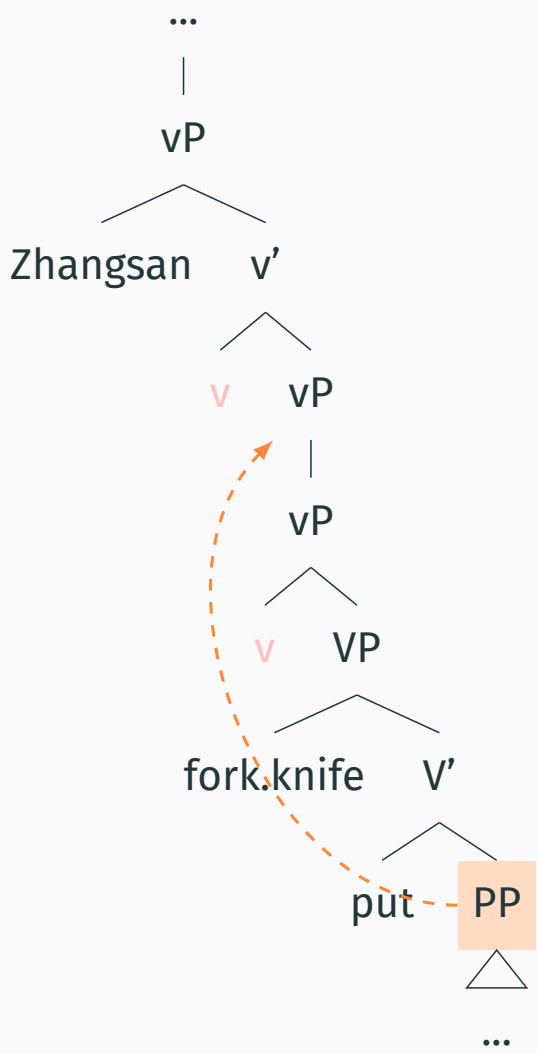
(7) short PP




(8) BA (Huang et al. 2009)




(9) PP shift (Larson 2014, Liu 2017)



# Comparisons - processing bias

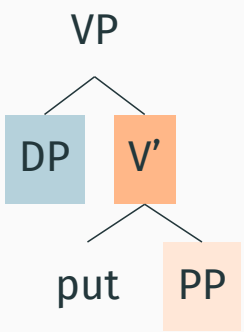
- *DP* $> 5$, the longer the more ba ; *DP* $< 5$, the shorter the more ba (Yao 2018)
- comparison
  - length of DP: $2$ and $9$
  - ba sentences should be preferred in
    - long PP sentences, and
    - short PP sentences

# Results - summary

- Attempt 1. BA sentence are preferred when PP is long.
  - When PP is short, none of the $20$ base metrics predicts ba preference
  - When PP is long $16$ out of the $20$ base metrics predicts ba preference

# Results

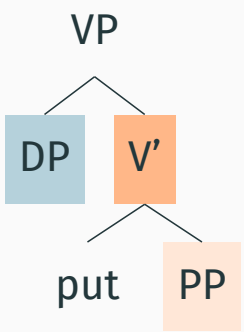
(10)




- Memory burden occurs when there is a “overweight” sister node

# Results

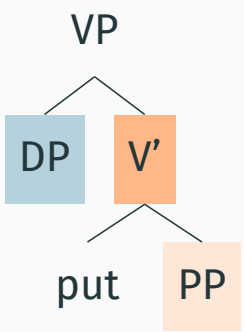
(10)




- Memory burden occurs when there is a “overweight” sister node
- The favored structure builds heavy phrases at the last.

# Results

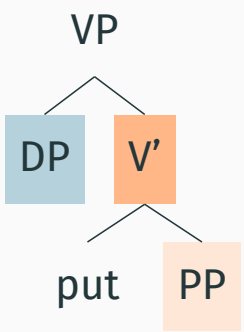
(10)




- Memory burden occurs when there is a “overweight” sister node
- The favored structure builds heavy phrases at the last.
- [BA DP V PP]

# Results

(10)




- Memory burden occurs when there is a “overweight” sister node
- The favored structure builds heavy phrases at the last.
- [BA DP V PP]
- Summary attempt 2. BA structure is favored when PP is heavier than DP (short-before-long)

# BA in Corpus

- Hypothesis:

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree
- Methodology

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree
- Methodology
  - Corpus: Chinese Treebank 8.0 (Xue et al. 2013)

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree
- Methodology
  - Corpus: Chinese Treebank 8.0 (Xue et al. 2013)
    - $\sim$ 1.5 million annotated and parsed words from Chinese newswire

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree
- Methodology
  - Corpus: Chinese Treebank 8.0 (Xue et al. 2013)
    - $\sim$ 1.5 million annotated and parsed words from Chinese newswire
    - 1835 BA out of 71373 (2.57%) (257 out of 10000 in sinica treebank (Huang et al. 2000))

# BA in Corpus

- Hypothesis:
  - BA structure is favored when PP is heavier than DP
    - $\rightarrow$ BA structures on average have VPs that are heavier than DP
    - weight: number of characters, number of phrases, height of tree
- Methodology
  - Corpus: Chinese Treebank 8.0 (Xue et al. 2013)
    - $\sim$ 1.5 million annotated and parsed words from Chinese newswire
    - 1835 BA out of 71373 (2.57%) (257 out of 10000 in sinica treebank (Huang et al. 2000))
    - 1744 BA left for analysis

# BA in corpus

(11) Examples of currently excluded cases (N = 91 (5%))

a. ba public debate fang zai qianmian
BA public debate put ZAI front

b. ba zhege bandaodianshitai ne keyi shuo zuole yige jieshao
BA this Al.Jazeera NE may say do.LE one.CL introduction

c. ba xiaohai gei yao huilai
BA child GEI(Filler) ask.for back

# BA in corpus

- Methodology (continued)
  - Measurement
    - number of characters in VP (`num.vp.char`)
    - number of characters in DP (`num.np.char`)
    - number of phrases in VP (`num.vp.phrase`)
    - number of phrases in DP (`num.np.phrase`)
    - height of VP (`height.vp`)
    - height of DP (`height.np`)

# BA in corpus

- Preliminary results
  - Mean `num.vp.char` = 6.82
    Mean `num.np.char` = 4.73 ($p < 0.001$)
  - Mean `num.vp.phrase` = 4.8
    Mean `num.np.phrase` = 3.1 ($p < 0.001$)
  - Mean `height.vp` = 5.3
    Mean `height.np` = 4.3 ($p < 0.001$)
- $\rightarrow$ BA structures found in the corpus generally follow the short-before-long scheme.

# Ref. i


Chomsky, N. (2014). *The minimalist program*. MIT press.

Francis, E. J., Matthews, S., Wong, R. W. Y., and Kwan, S. W. M. (2011). Effects of weight and syntactic priming on the production of cantonese verb-doubling. *Journal of psycholinguistic research*, 40(1):1–28.

Graf, T. (2012). Locality and the complexity of minimalist derivation tree languages. In *Formal Grammar*, pages 208–227. Springer.

Graf, T. (2013). *Local and transderivational constraints in syntax and semantics*. PhD thesis, UCLA.

Graf, T., Fodor, B., Monette, J., Rachiele, G., Warren, A., and Zhang, C. (2015a). A refined notion of memory usage for minimalist parsing. In *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, pages 1–14, Chicago, USA. Association for Computational Linguistics.

Graf, T., Fodor, B., Monette, J., Rachiele, G., Warren, A., and Zhang, C. (2015b). A refined notion of memory usage for minimalist parsing. In *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, pages 1–14.

Graf, T., Monette, J., and Zhang, C. (2017). Relative clauses as a benchmark for Minimalist parsing. *Journal of Language Modelling*, 5:57–106.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press on Demand.

Huang, C.-R., Chen, F.-Y., Chen, K.-J., Gao, Z.-m., and Chen, K.-Y. (2000). Sinica treebank: Design criteria, annotation guidelines, and on-line interface. In *Second Chinese Language Processing Workshop*, pages 29–37, Hong Kong, China. Association for Computational Linguistics.

Huang, C.-T. J., Li, Y.-h. A., and Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press Cambridge.

Kayne, R. S. (1994). *The antisymmetry of syntax*. Number 25. mit Press.

Larson, R. K. (1998). *Light predicate raising*. MIT. Center for Cognitive Science.

Larson, R. K. (2014). *On shell structure*. Routledge.

# Ref. ii


Liu, F.-h. (2007). Word order variation and ba sentences in chinese. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 31(3):649–682.

Liu, L. (2017). Applicative shift in mandarin sov word order. *talk at North American Conference on Chinese Linguistics (NACCL) 29*.

Liu, L. (2018). Minimalist parsing of heavy np shift. In *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation*.

Rochemont, M. and Culicover, P. W. (1997). Deriving dependent right adjuncts in english. *Rightward movement*, pages 279–300.

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

Stabler, E. (1996). Derivational minimalism. In *International Conference on Logical Aspects of Computational Linguistics*, pages 68–95. Springer.

Stabler, E. P. (2011). Computational perspectives on minimalism. *Oxford handbook of linguistic minimalism*, pages 617–643.

Stabler, E. P. (2013). Two models of minimalist, incremental syntactic analysis. *Topics in cognitive science*, 5(3):611–633.

Xue, N., Zhang, X., Jiang, Z., Palmer, M., Xia, F., Chiou, F.-D., and Chang, M. (2013). Chinese treebank 8.0. *Linguistic Data Consortium*.

Yao, Y. (2018). Np weight effects in word order variation in mandarin chinese. *Lingua Sinica*, 4(1):5.

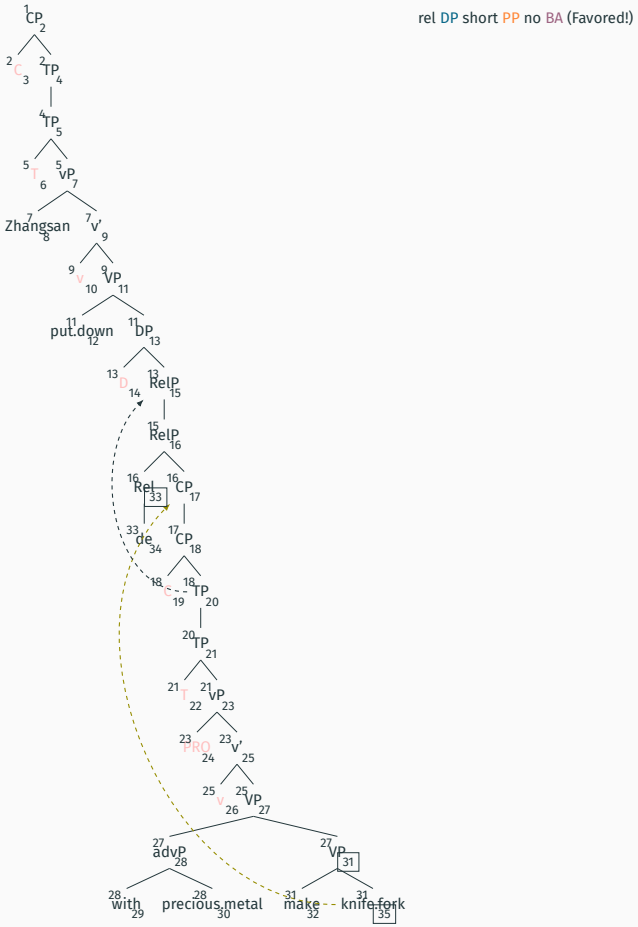
Zhang, C. (2017). *Stacked Relatives: Their Structure, Processing and Computation*. PhD thesis, State University of New York at Stony Brook.

# Comparisons

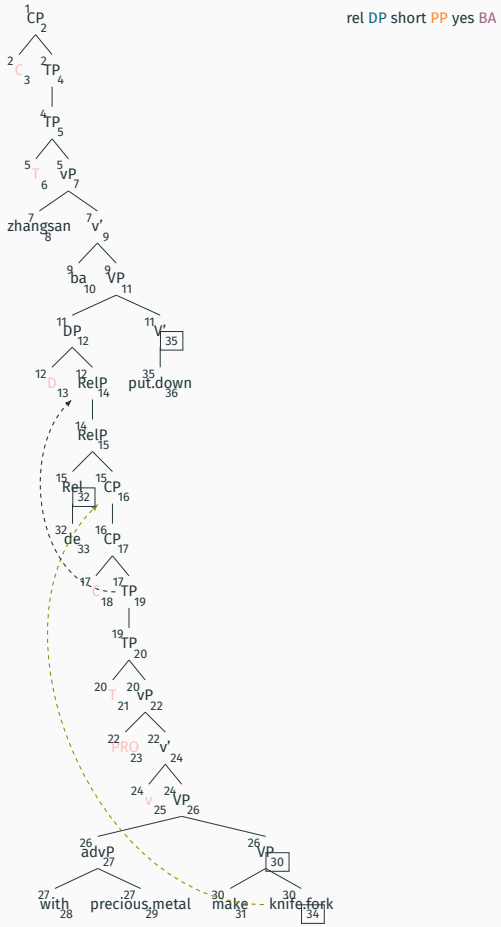
Comparisons are conducted...

- by specifying
  - target sentences
  - syntactic structure
  - human processing bias
- $\rightarrow$ testing processing predictions of different target sentences
  - BA vs. non-BA
- $\rightarrow$ evaluating different structures of the same target sentences
  - HNPS
  - rightward movement for the win!

# Results - derivation trees




rel DP short PP no BA (Favored!)




rel DP short PP yes BA

# Results - derivation trees