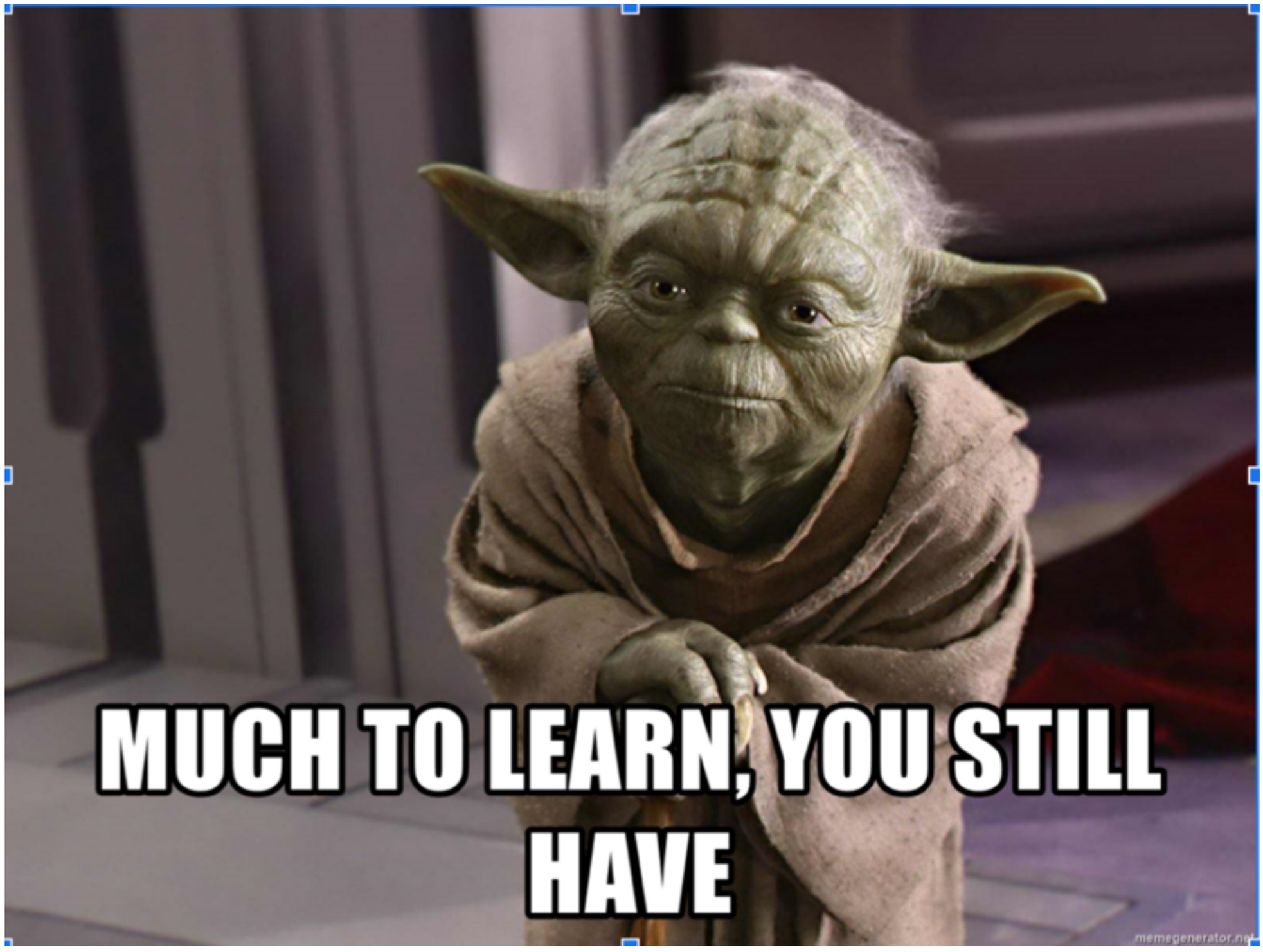
A Comparative Corpus Analysis of PP Ordering in English and Chinese

Zoey Liu
Computational Linguistics Lab
Department of Linguistics, UC Davis


MUCH TO LEARN, YOU STILL
HAVE
memegenerator.net

- Dependency length minimization (DLM)

- Dependency length minimization (DLM)
  - grammars minimize dependency length

- Dependency length minimization (DLM)
  - grammars minimize dependency length
  - constituent orderings in syntactic alternations

- Predictions of DLM on constituent orderings

- Predictions of DLM on constituent orderings

(a) Zoey **presented** [**on** something linguistic] [**to** her professors and colleagues]

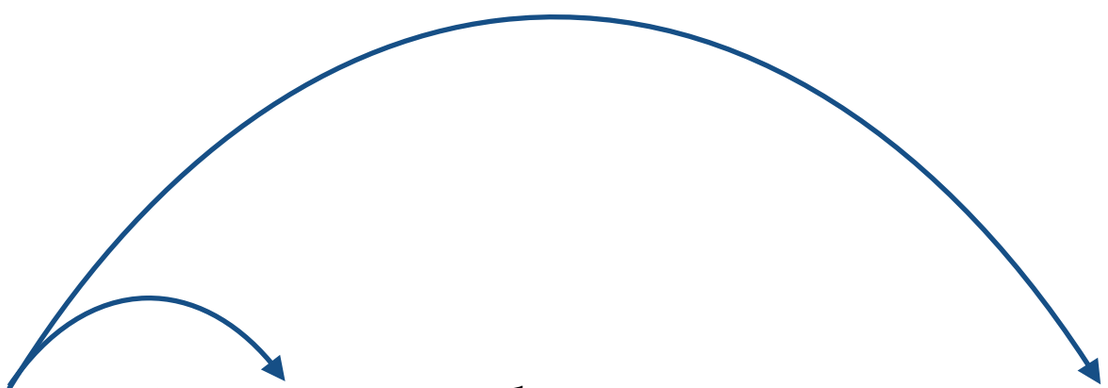
(b) Zoey **presented** [**to** her professors and colleagues] [**on** something linguistic]

- Predictions of DLM on constituent orderings

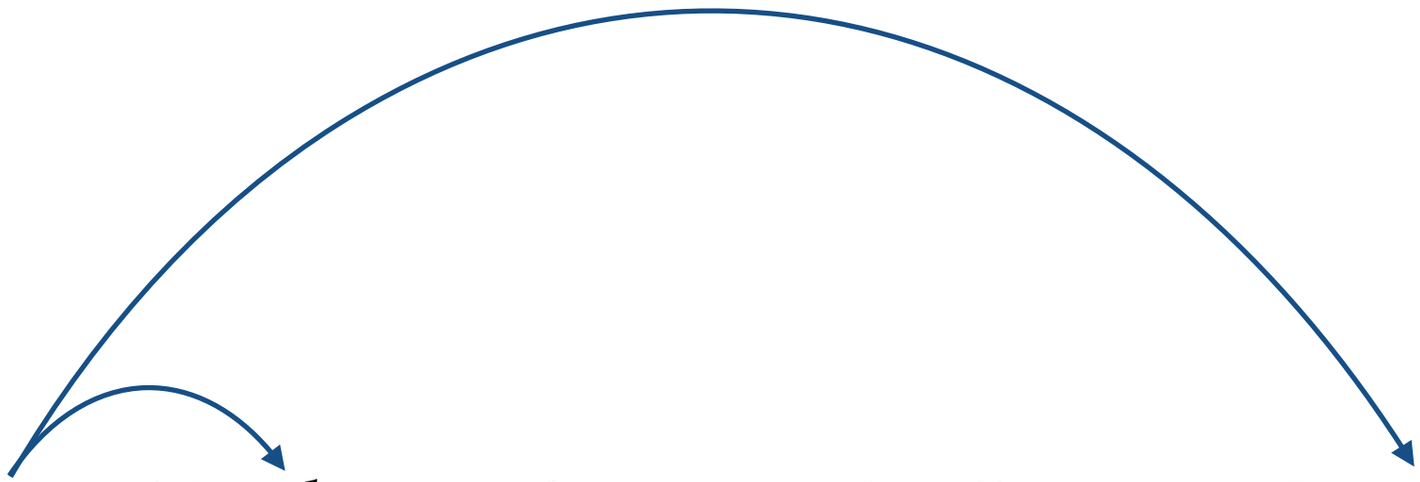
(a) Zoey **presented** [**on** something linguistic] [**to** her professors and colleagues]

(b) Zoey **presented** [**to** her professors and colleagues] [**on** something linguistic]

- Predictions of DLM on constituent orderings



(a) Zoey **presented** [**on** something linguistic] [**to** her professors and colleagues]



(b) Zoey **presented** [**to** her professors and colleagues] [**on** something linguistic]

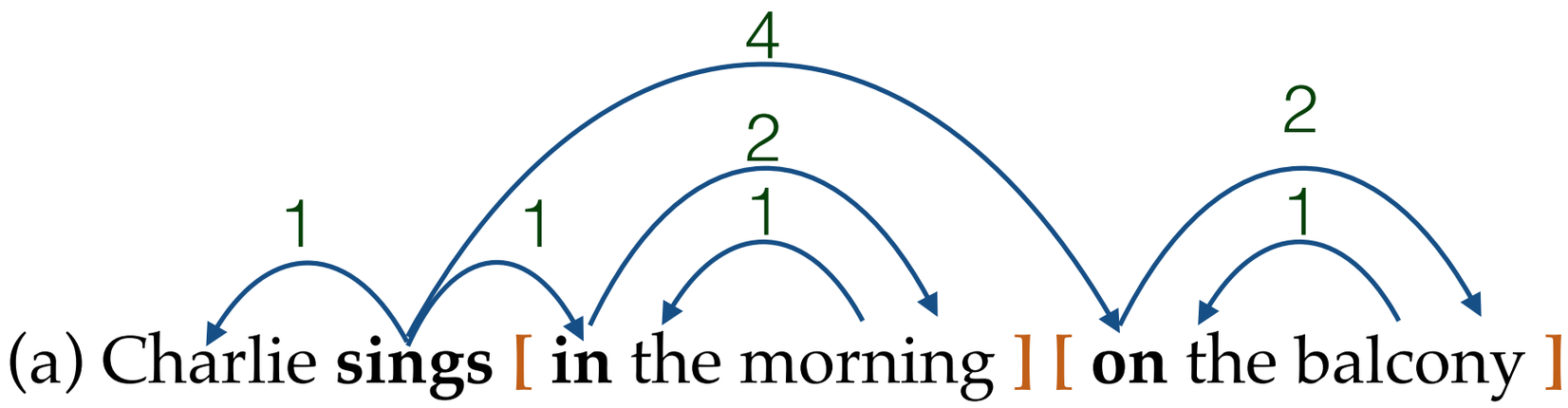
But …

But …

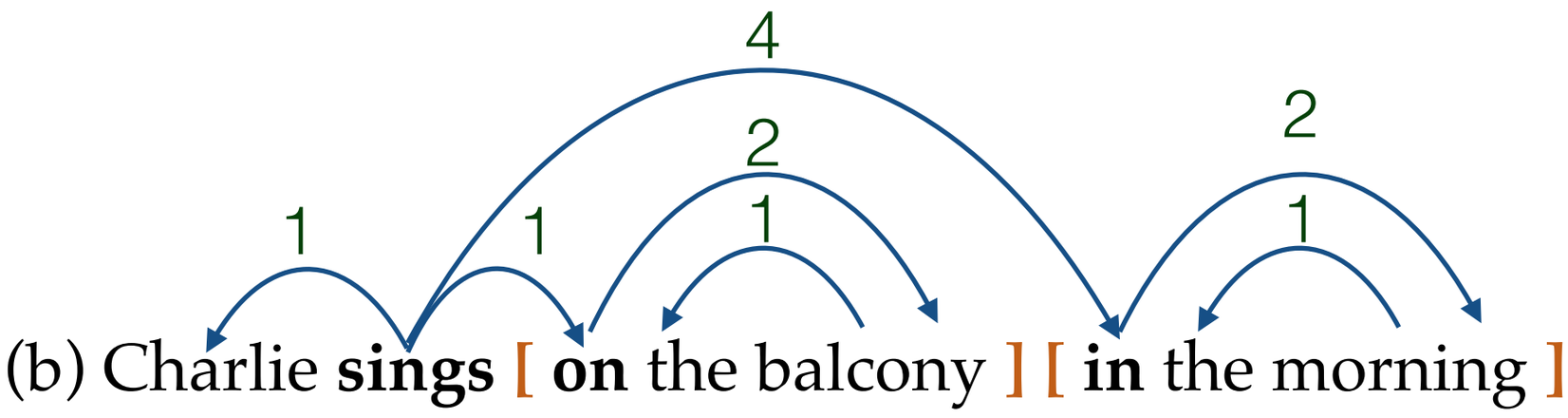
- not able to explain when equal dependency length

(a) Charlie **sings** [**in** the morning] [**on** the balcony]

(b) Charlie **sings** [**on** the balcony] [**in** the morning]



Total dependency length = 12



Total dependency length = 12

But …

- not able to explain when equal dependency length
- longer dependencies are easier to process sometimes

But …

- not able to explain when equal dependency length
- longer dependencies are easier to process sometimes
- efficacy of DLM varies across languages

But …

- not able to explain when equal dependency length
- longer dependencies are easier to process sometimes
- efficacy of DLM varies across languages

Competing and cooperating motivations

Testbed: PP order in English and Chinese

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?
- Argument status
  - Does more argument-like PP appear closer?

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?
- Argument status
  - Does more argument-like PP appear closer?
- Manner Place Time (MPT)

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?
- Argument status
  - Does more argument-like PP appear closer?
- Manner Place Time (MPT)
  - *Vera danced **elegantly***

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?
- Argument status
  - Does more argument-like PP appear closer?
- Manner Place Time (MPT)
  - *Vera danced **elegantly** **on the dance floor***

Testbed: PP order in English and Chinese

- Dependency length
  - Does short PP appear closer?
- Argument status
  - Does more argument-like PP appear closer?
- Manner Place Time (MPT)
  - *Vera danced **elegantly** **on the dance floor** **at night***

Data

- English: Penn Treebank (PTB)
- Chinese: Penn Chinese Treebank (CTB)

PP ordering in Chinese

PP ordering in Chinese

(a) 他们 将 [在 生产 电子产品 方面] [和 中国] 合作

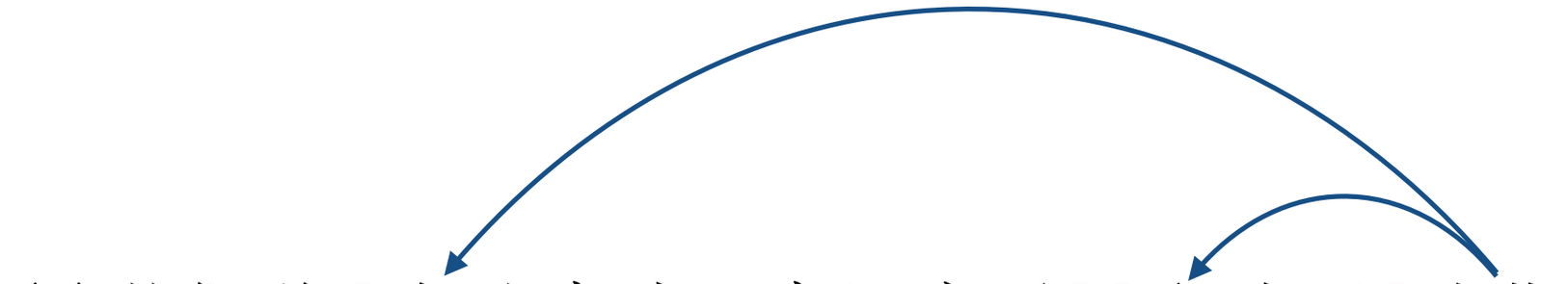
They will [**in** the aspects of electronic device production] [**with** China] **collaborate**

(b) 他们 将 [和 中国] [在 生产 电子产品 方面] 合作

They will [**with** China] [**in** the aspects of electronic device production] **collaborate**

They will collaborate with China in the aspects of electronic device production

PP ordering in Chinese



(a) 他们 将 [在 生产 电子产品 方面] [和 中国] 合作

They will [**in** the aspects of electronic device production] [**with** China] **collaborate**

(b) 他们 将 [和 中国] [在 生产 电子产品 方面] 合作

They will [**with** China] [**in** the aspects of electronic device production] **collaborate**

They will collaborate with China in the aspects of electronic device production

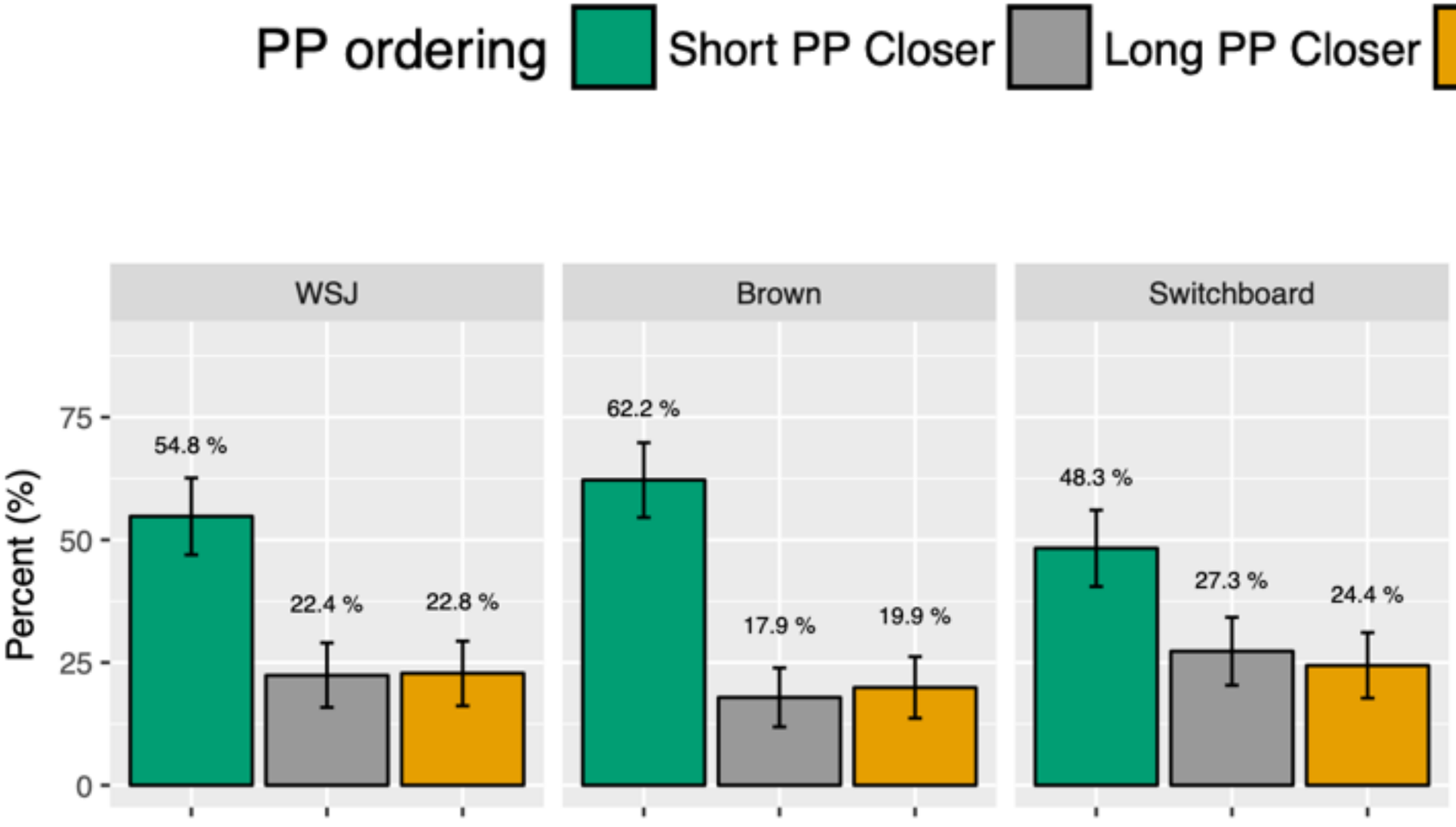
Measures for dependency length

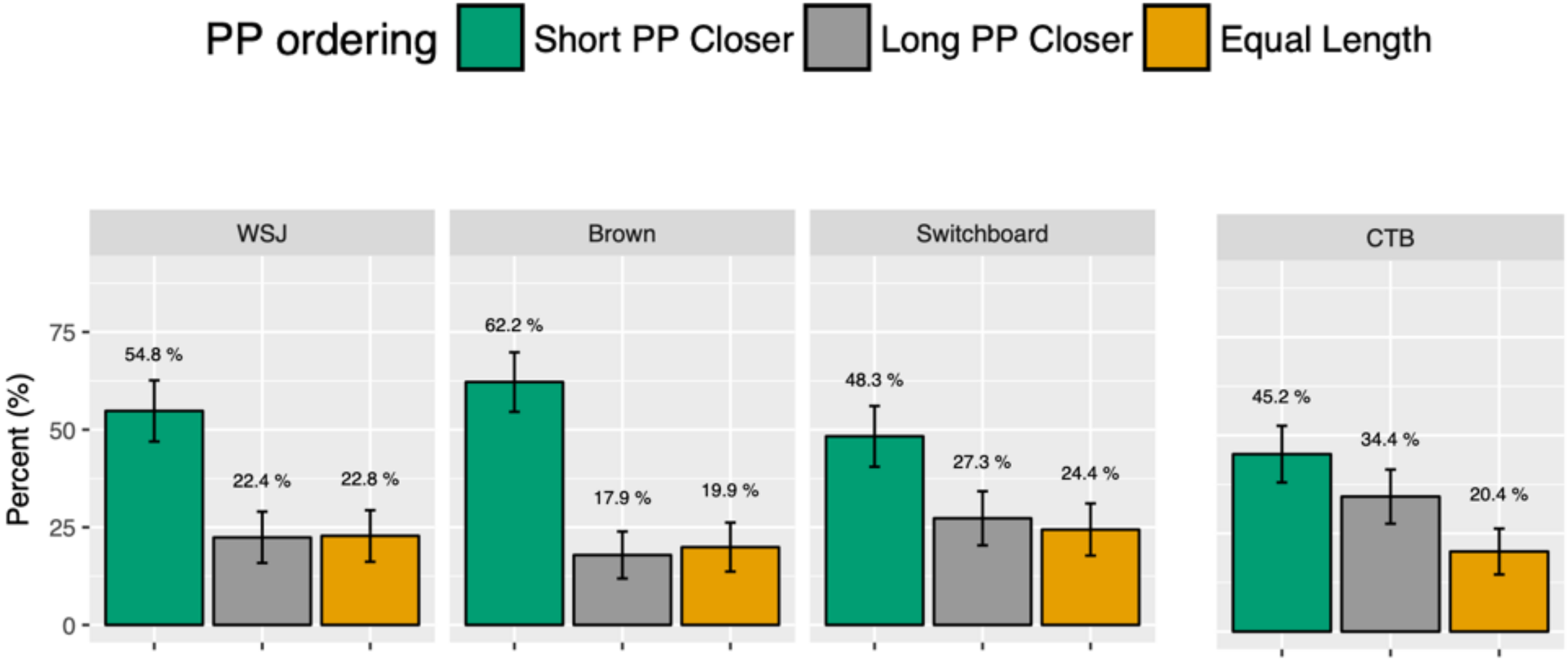
Measures for dependency length

- the number of tokens within each PP

Measures for dependency length

- the number of tokens within each PP
- proportion of short PP closer, long PP closer, equal length


PP ordering
Short PP Closer
Long PP Closer
Equal Length
WSJ
Brown
Switchboard
Percent (%)
75
50
25
0
54.8 %
22.4 %
22.8 %
62.2 %
17.9 %
19.9 %
48.3 %
27.3 %
24.4 %


PP ordering
Short PP Closer
Long PP Closer
Equal Length
WSJ
Brown
Switchboard
CTB
Percent (%)
75
50
25
0
54.8 %
22.4 %
22.8 %
62.2 %
17.9 %
19.9 %
48.3 %
27.3 %
24.4 %
45.2 %
34.4 %
20.4 %

Measures for argument status

Measures for argument status

- Merlo and Ferrer (2006)

Measures for argument status

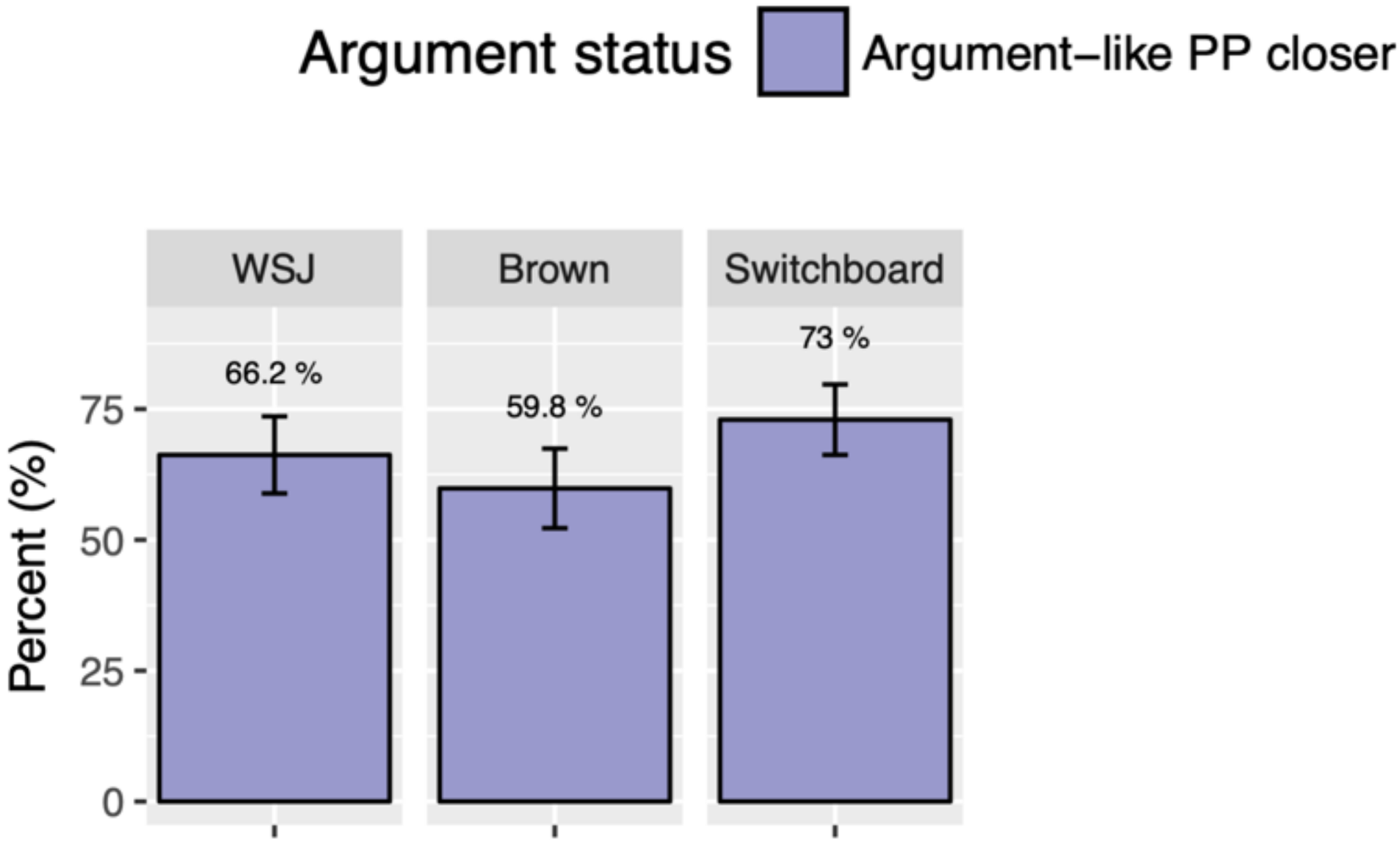
- Merlo and Ferrer (2006)
- how *argument-like* or *adjunct-like*

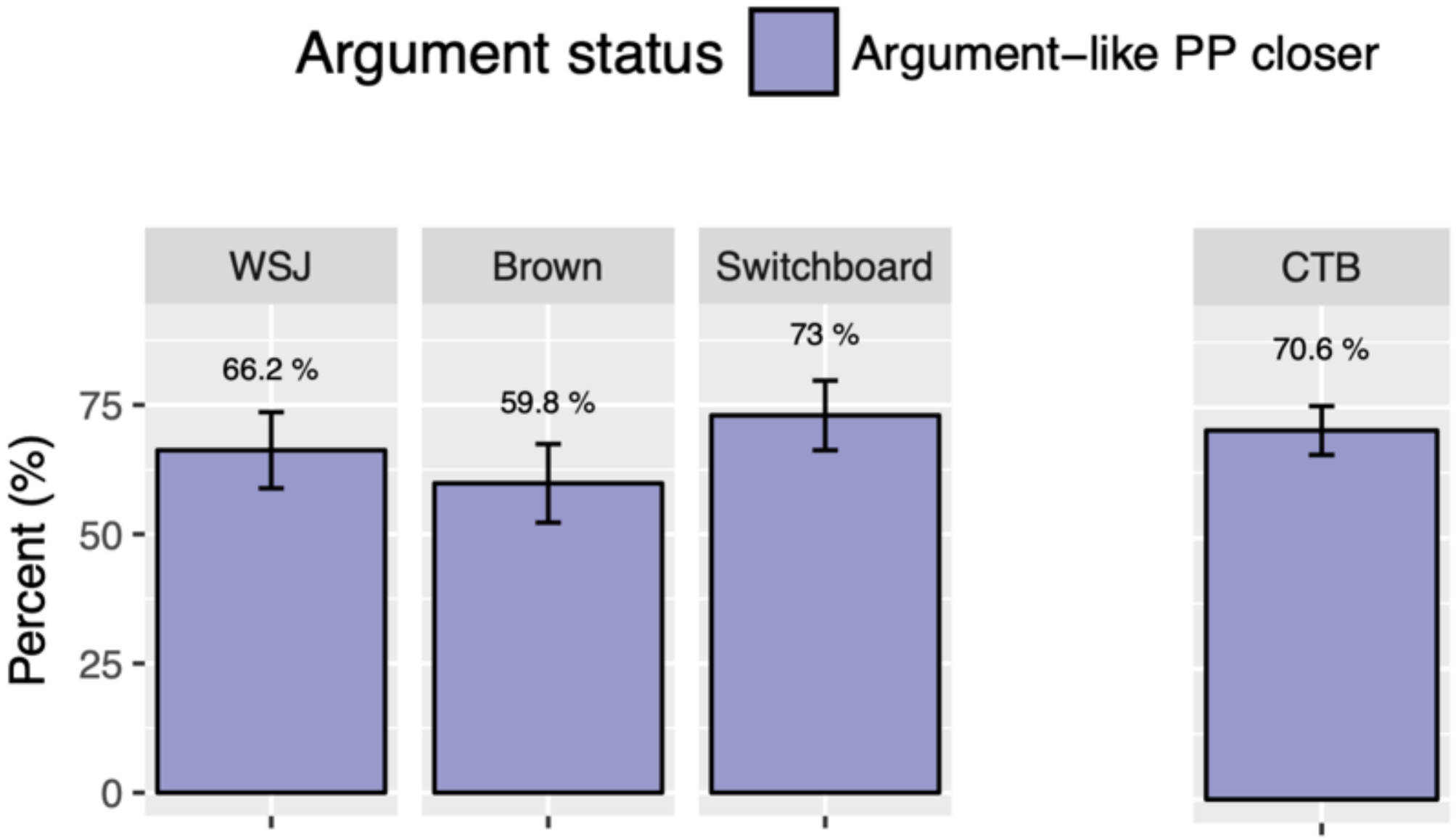
Measures for argument status

- Merlo and Ferrer (2006)

- how *argument-like* or *adjunct-like*

- proportion of argument-like PP closer


Argument status
Argument–like PP closer
WSJ
Brown
Switchboard
66.2 %
59.8 %
73 %
Percent (%)
75
50
25
0


Argument status
Argument−like PP closer
WSJ
Brown
Switchboard
CTB
66.2 %
59.8 %
73 %
70.6 %
Percent (%)
75
50
25
0

Measures for MPT

Measures for MPT

- manner (PP-MNR), place (PP-LOC), time (PP-TMP)

Measures for MPT

- manner (PP-MNR), place (PP-LOC), time (PP-TMP)
- English: MPT; Chinese: TPM

Measures for MPT

- manner (PP-MNR), place (PP-LOC), time (PP-TMP)
- English: MPT; Chinese: TPM

WSJ	89.3%
Brown	100%
Switchboard	100%

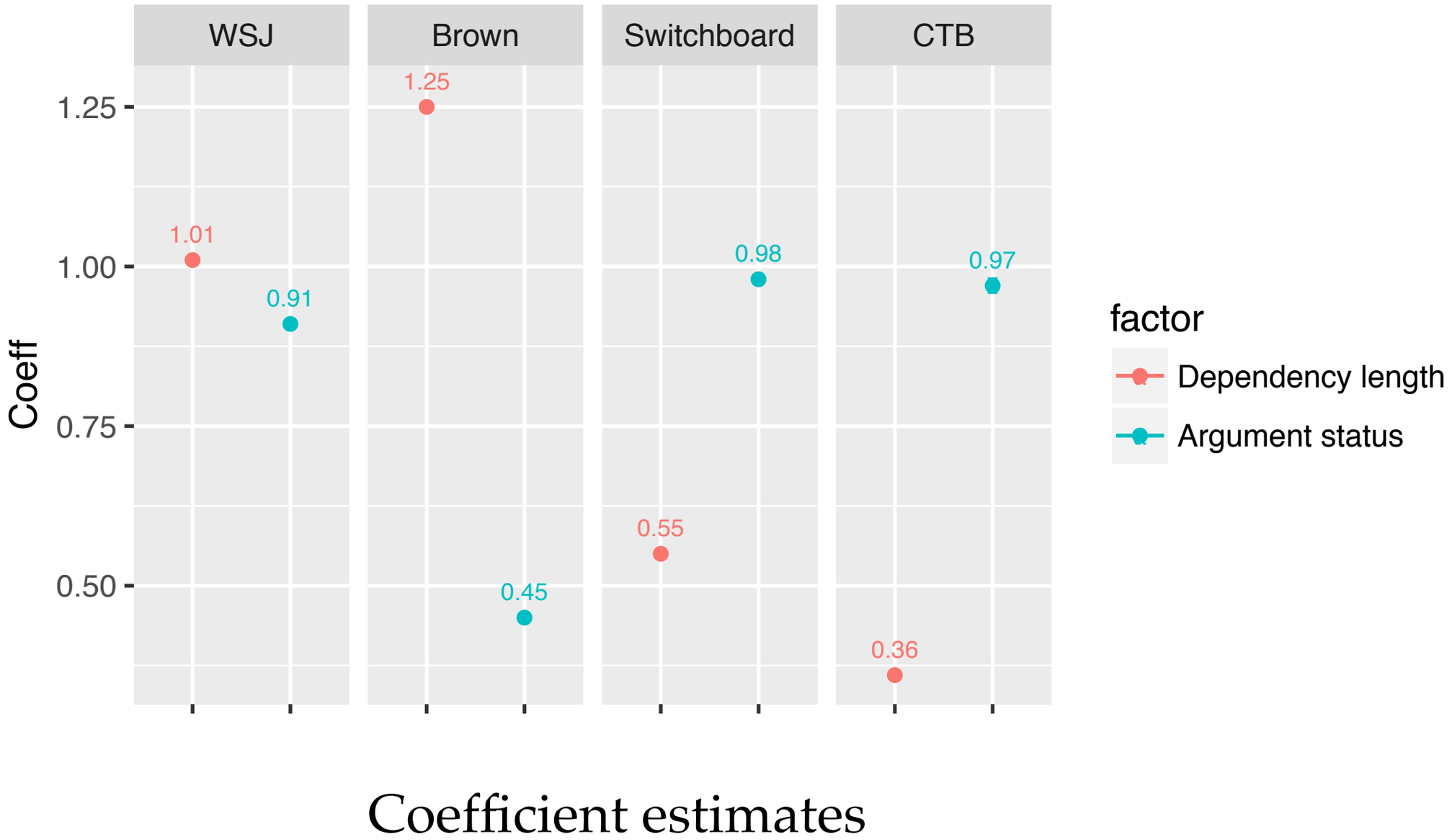
Cooperation and competition between dependency length and argument status

Cooperation and competition between dependency length and argument status

- Logistic regression

Factors	1	-1	0
dependency length	short PP closer	long PP closer	equal length
argument status	argument-like PP closer	adjunct-like PP closer	same argument status

Cooperation and competition between dependency length and argument status



Questions revisited

Questions revisited

- Does short PP appear closer across the two languages?

Questions revisited

- Does short PP appear closer across the two languages?
  - Yes

Questions revisited

- Does short PP appear closer across the two languages?
  - Yes
  - strongest preference when matching headedness

Questions revisited

- Does short PP appear closer across the two languages?
  - Yes
  - strongest preference when matching headedness
  - stronger in written than spoken corpora

Questions revisited

- Does the argument-like PP appear closer?

Questions revisited

- Does the argument-like PP appear closer?
  - Yes

Questions revisited

- Does the argument-like PP appear closer?
  - Yes
  - stronger in spoken English and Mandarin Chinese

Pondering over

- Why DLM / Argument closer than adjuncts?
  - Minimize domains (Hawkins 2004); processing typology
- Why are preverbal and postverbal DLM preferences different?
  - listeners vs. speakers perspective
  - predictability-based accounts
- Variations of grammatical constructions / functions


THANK YOU
I SAY