Syntactic Transfer Using a Bilingual Lexicon

Greg Durrett, Adam Pauls, and Dan Klein
UC Berkeley

Parsing a New Language

Mosambik hofft auf Handel mit anderen Mitgliedern

Parsing a New Language

Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

Parsing a New Language

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

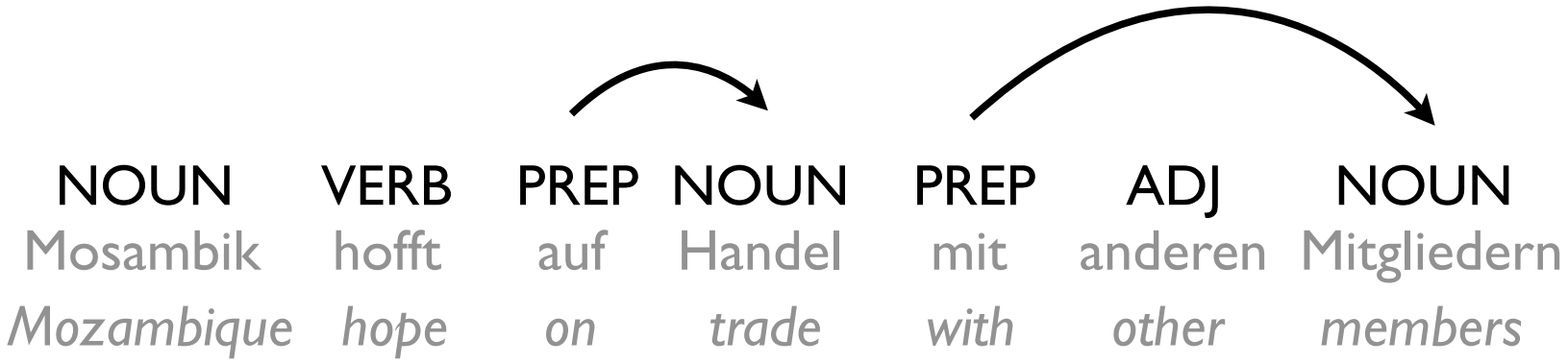
Parsing a New Language

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

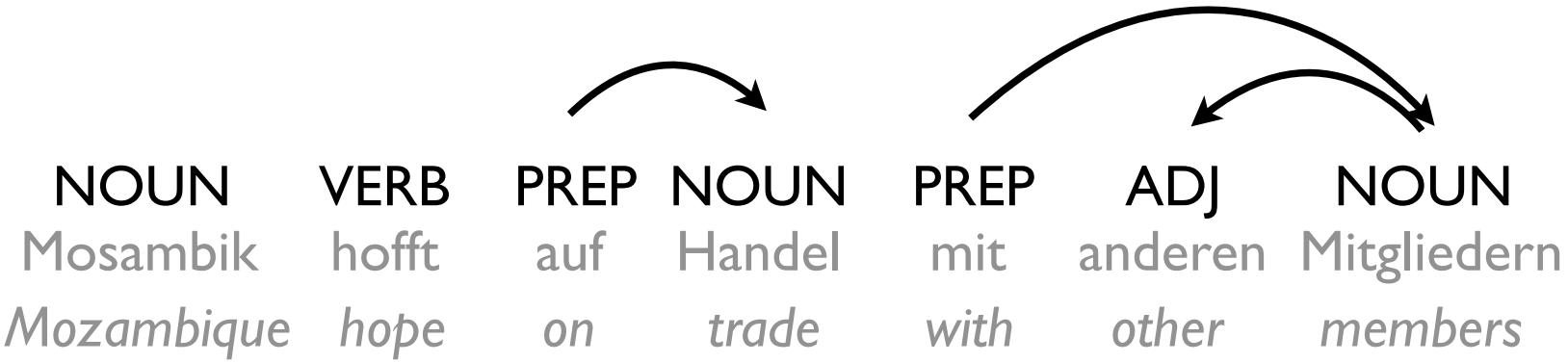
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

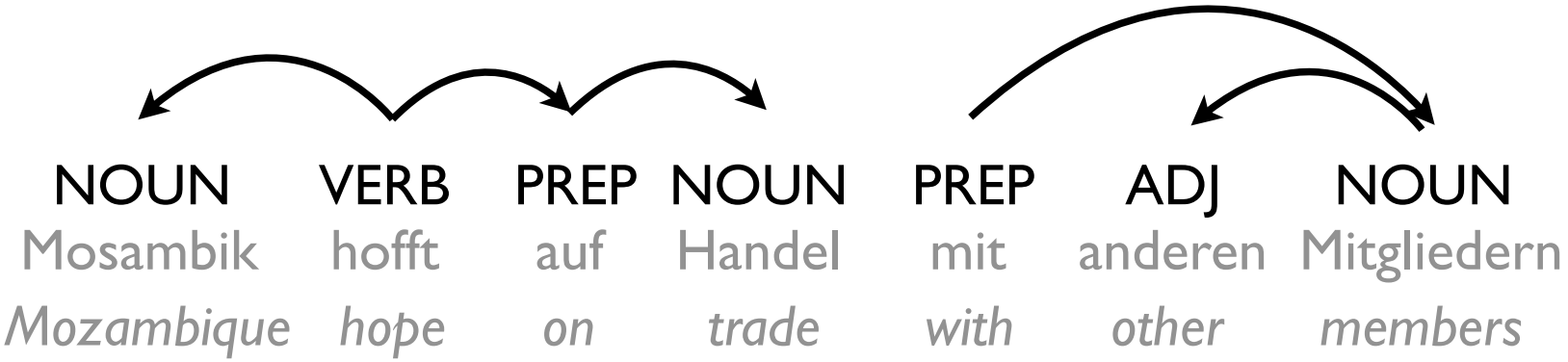
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

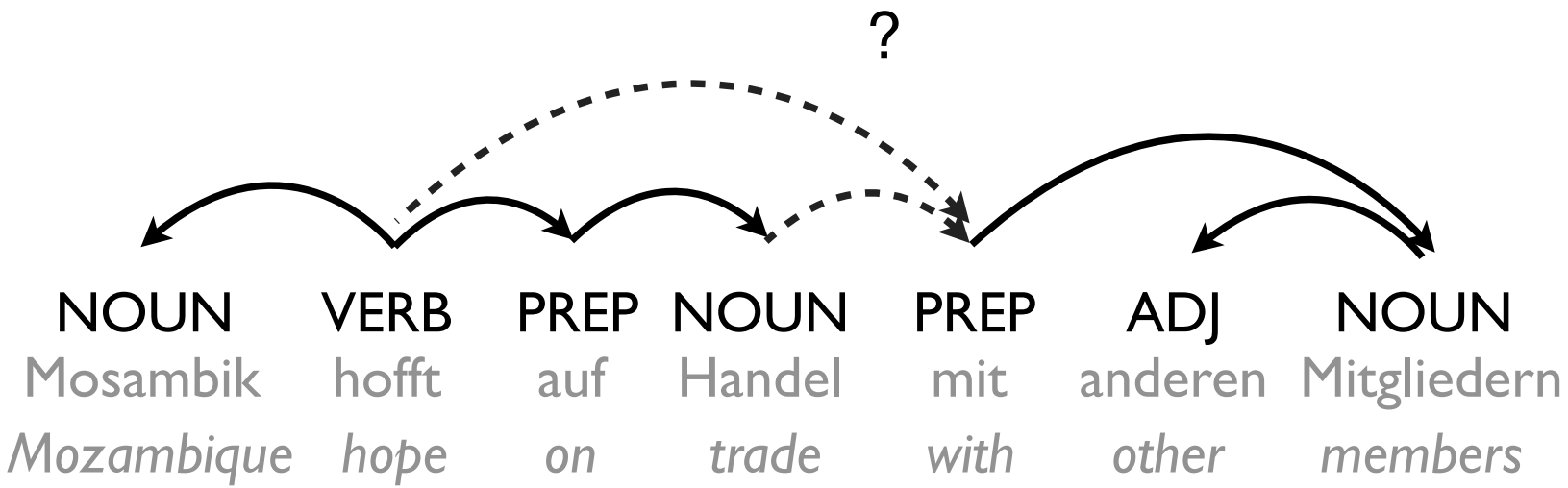
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

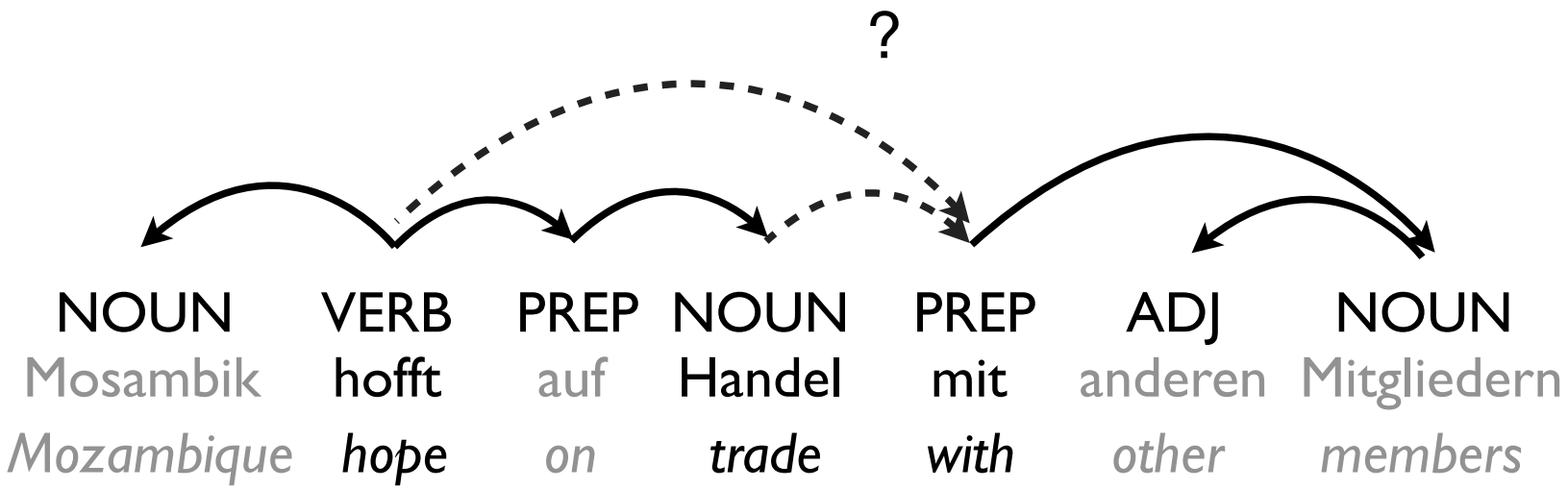
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

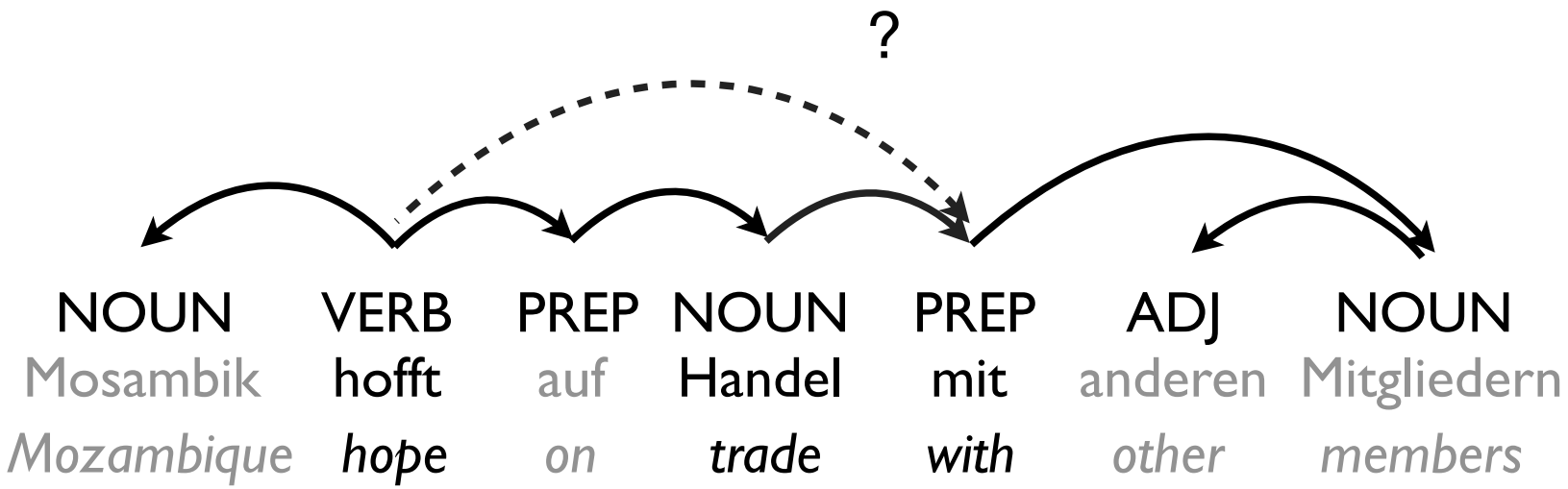
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

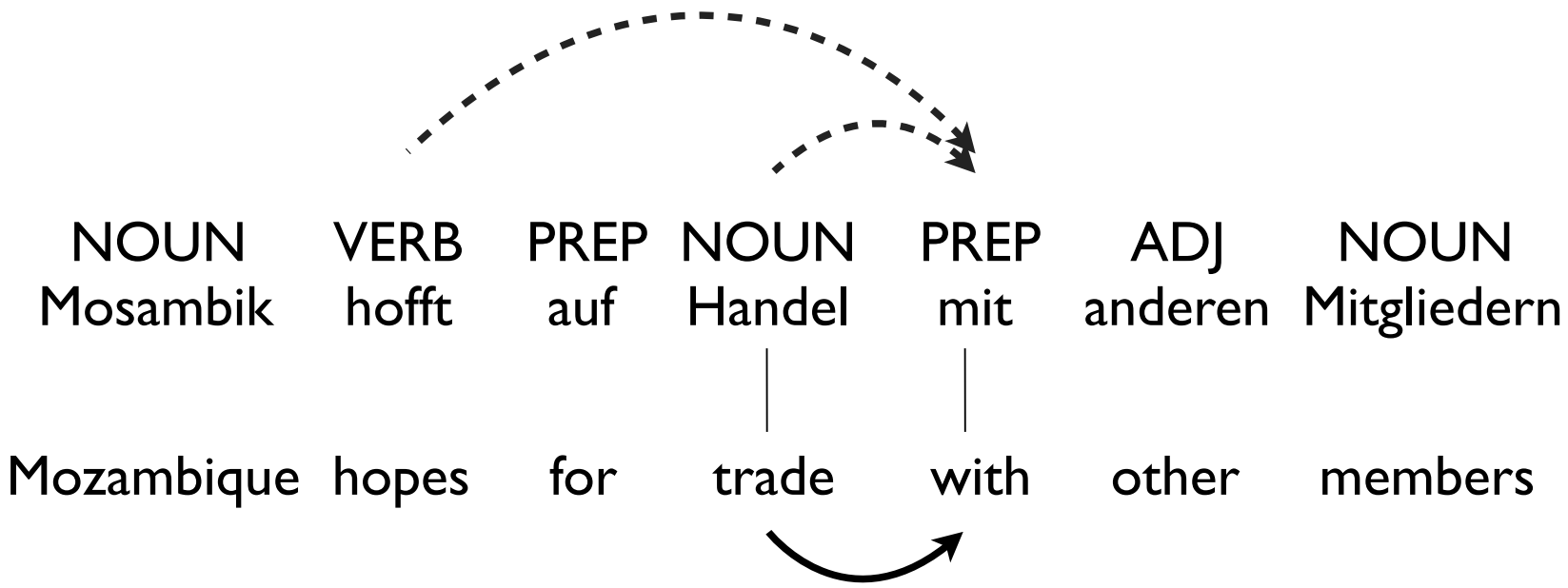
Parsing a New Language



[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

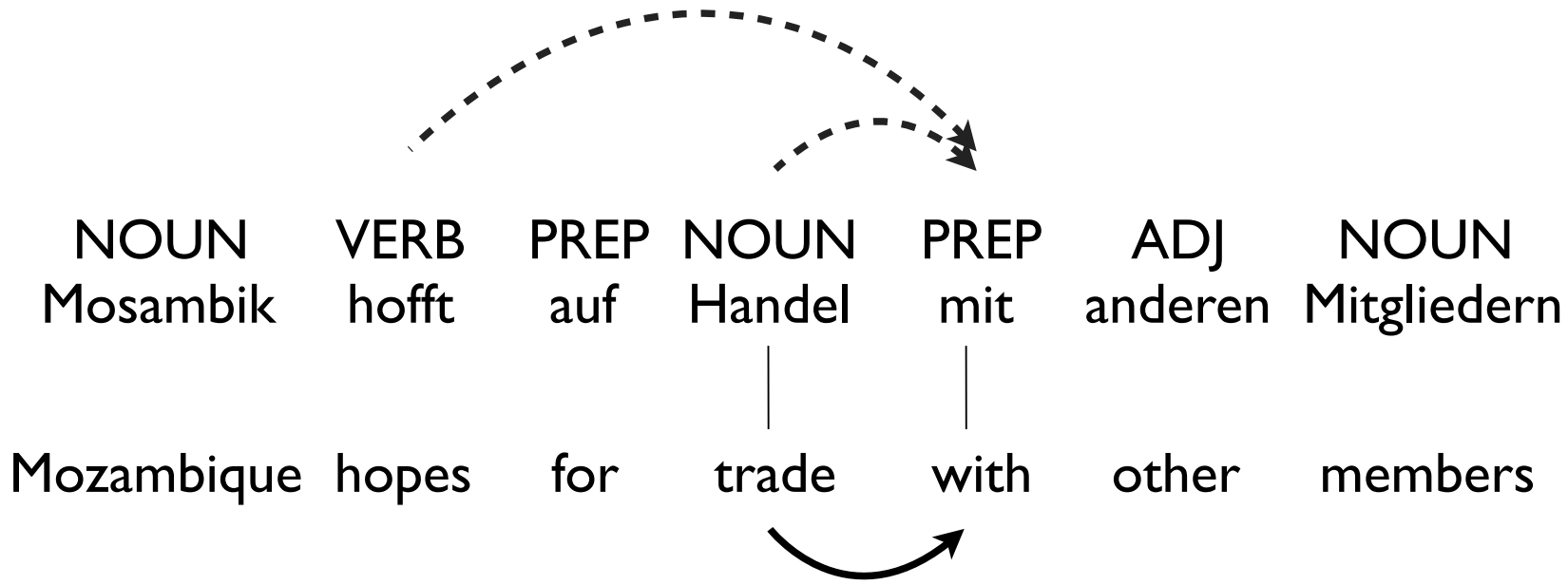
[McDonald et al. (2011), Cohen et al. (2011)]

Token-level Projection



[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]

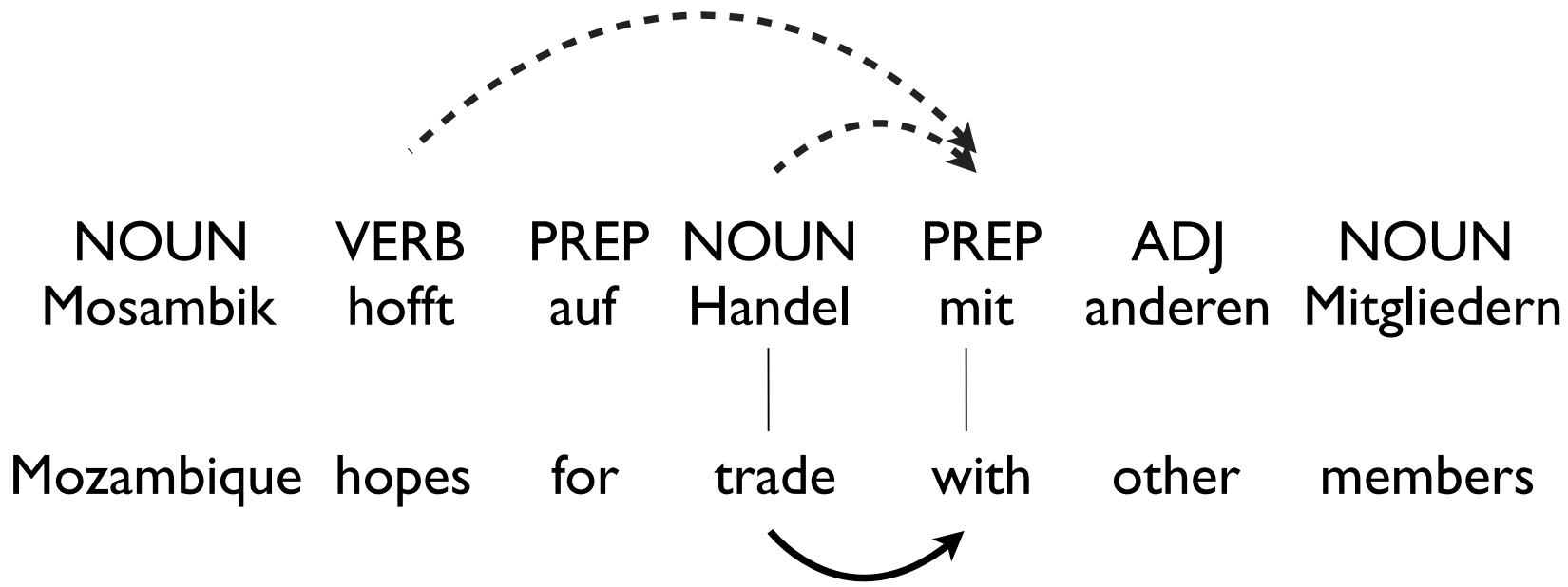
Token-level Projection



- Assumes access to parallel data

[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]

Token-level Projection



- Assumes access to parallel data
- Asymmetric train and test conditions

[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]

Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern

Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
			trade, deal	*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

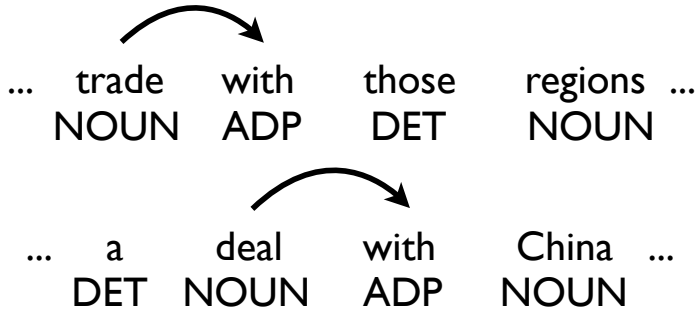
Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
			trade, deal	*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

EN trees



Type-level Projection

Feature: "Goodness in English"
Value: high

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
			trade, deal	*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

EN trees

...	trade	with	those	regions	...
	NOUN	ADP	DET	NOUN	

...	a	deal	with	China	...
	DET	NOUN	ADP	NOUN	

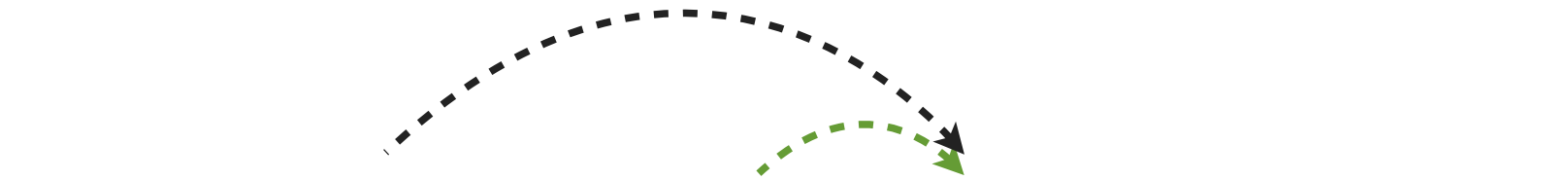
Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

Type-level Projection



NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
	hope			*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

Type-level Projection

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
	hope			*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

EN trees

no examples

Type-level Projection

Feature: “Goodness in English”
Value: low

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
	hope			*with*		

DE–EN dict

Handel	*trade, deal*
hofft	*hope*
mit	*with*

EN trees

no examples

Model

- Edge-factored discriminative parser, based on MSTParser (McDonald et al. 2005)
- DELEX features: universal POS (McDonald et al. 2011)
- PROJ features: "Goodness in English"

Model: DELEX Features

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

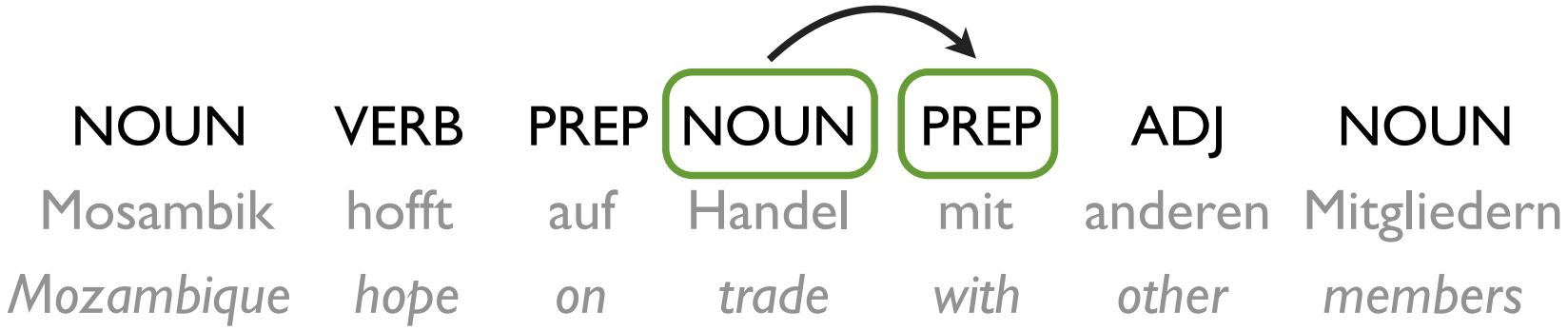
Model: DELEX Features

ATTACH

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

Model: DELEX Features

ATTACH



Model: DELEX Features

ATTACH

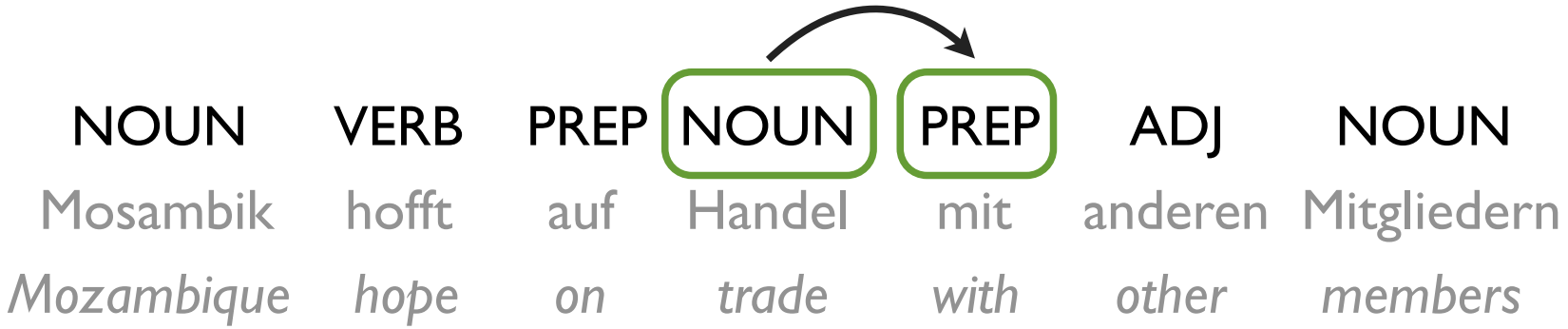
Feature	Value
NOUN → PREP	1

Model: DELEX Features

ATTACH

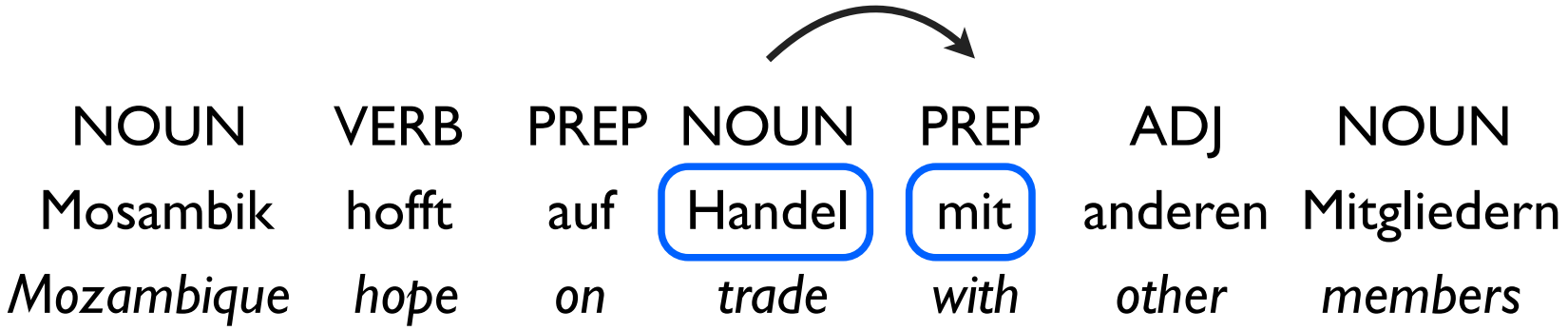
Feature	Value
NOUN → PREP	1

+ indicators including direction and distance



Model: PROJ Features

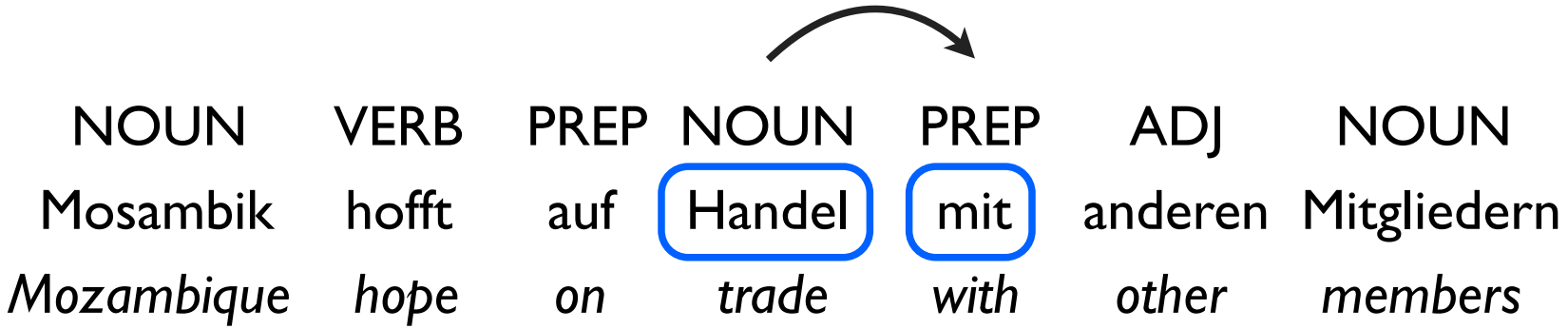
ATTACH



Model: PROJ Features

ATTACH

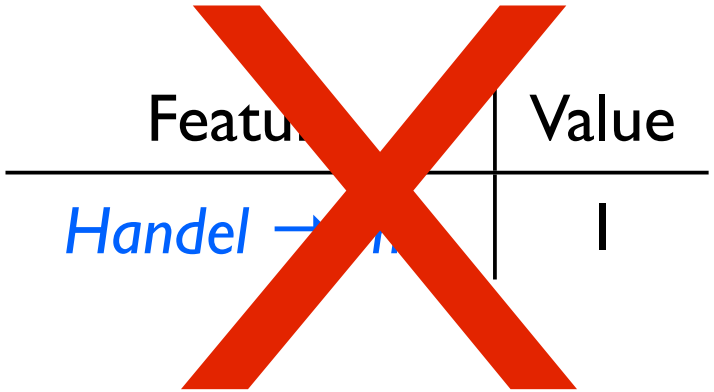
Feature	Value
Handel → mit	1



Model: PROJ Features

ATTACH

Insufficient target-language supervised data to use

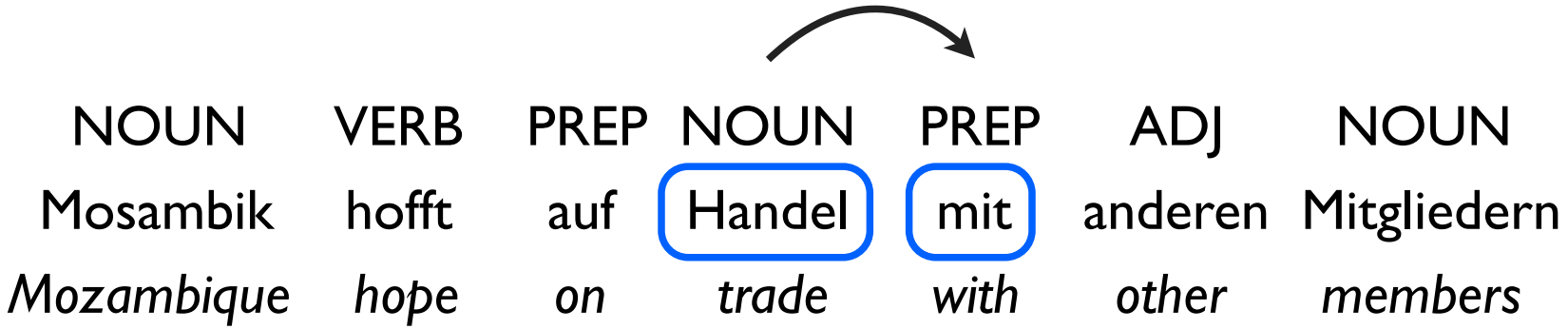


NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

Model: PROJ Features

ATTACH

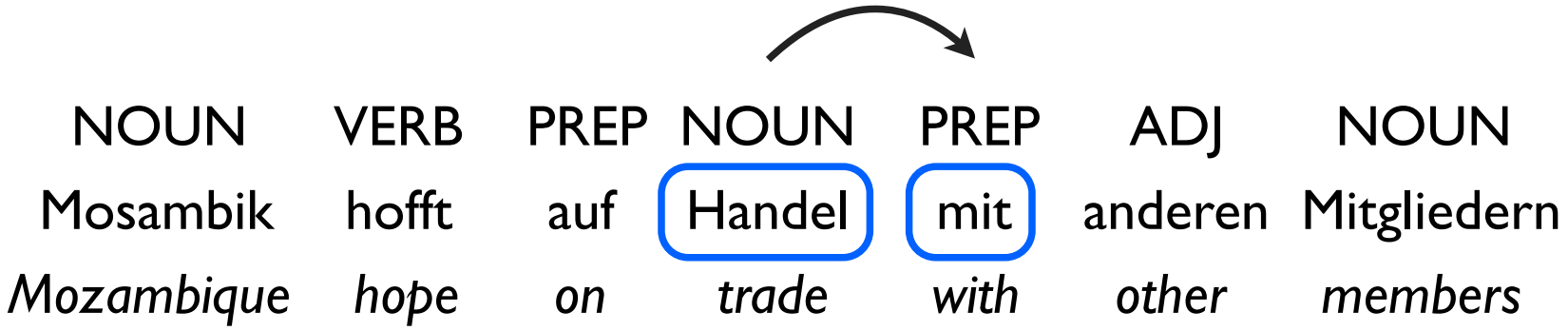
Feature	Value
Handel → mit	



Model: PROJ Features

ATTACH

Feature	Value
	Score(*Handel* → *mit*)



Model: PROJ Features

ATTACH

Feature	Value
	Score(*Handel* → *mit*)

Query

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

Model: PROJ Features

ATTACH

Feature	Value
Goodness in English	Score(*Handel* → *mit*)

Query

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

Query Score Computation

$$\text{Score}(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP} \mid \textit{Handel})$$

Query Score Computation

$$\text{Score}(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP}|\textit{Handel})$$

$$= \log \sum_{w_s} p(\text{PREP}|w_s)\ p(w_s|\textit{Handel})$$

Query Score Computation

$$\text{Score}(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP}|\textit{Handel})$$

$$= \log \sum_{w_s} p(\text{PREP}|w_s)\ p(w_s|\textit{Handel})$$

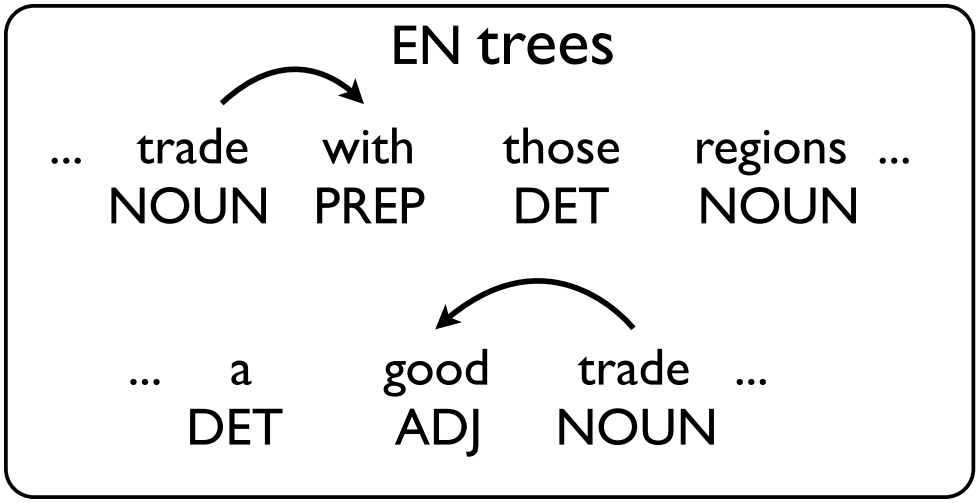
w_s	$p(w_s\|\textit{Handel})$
trade	0.5
deal	0.5

Query Score Computation

$$\text{Score}(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP}|\textit{Handel})$$

$$= \log \sum_{w_s} p(\text{PREP}|w_s)\ p(w_s|\textit{Handel})$$

$$w_s = \textit{trade}$$



w_s	$p(w_s\|\textit{Handel})$
trade	0.5
deal	0.5

Query Score Computation

$$\text{Score}(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP} | \textit{Handel})$$

$$= \log \sum_{w_s} p(\text{PREP} | w_s)\ p(w_s | \textit{Handel})$$

$$w_s = \textit{trade}$$

x	$p(x \mid \textit{trade})$
PREP	0.5
ADJ	0.5

w_s	$p(w_s \mid \textit{Handel})$
trade	0.5
deal	0.5

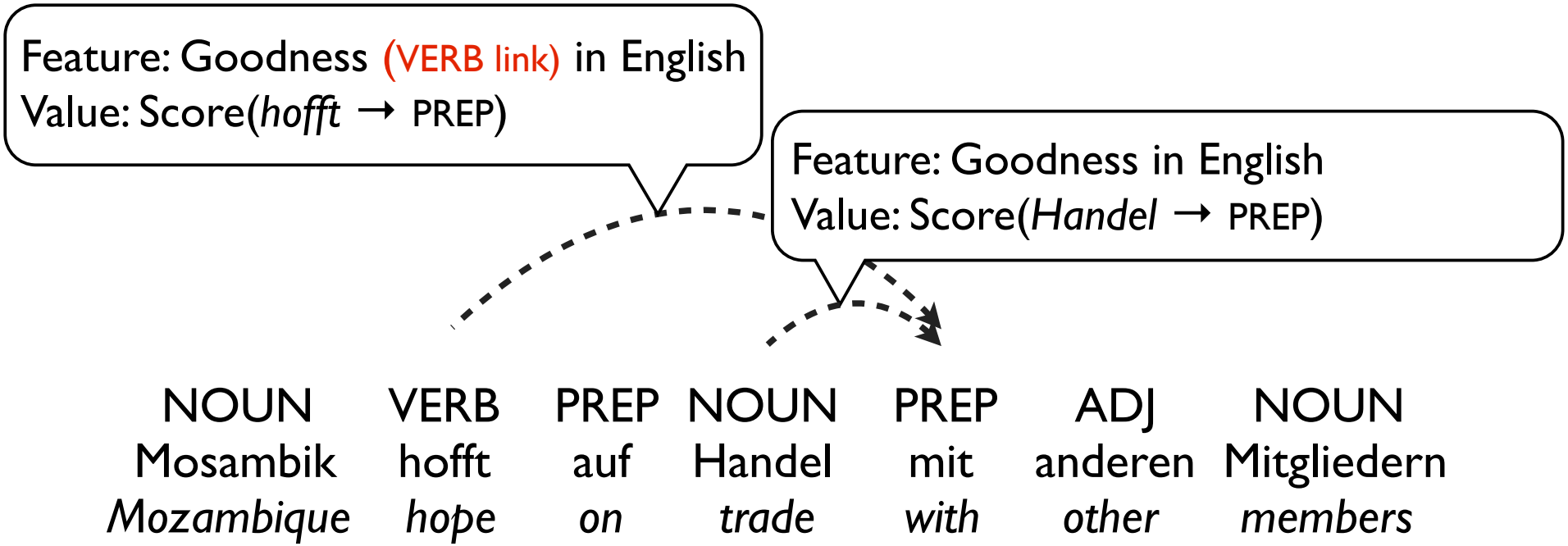
Feature Specialization

Feature: Goodness in English
Value: Score(*hofft* → PREP)

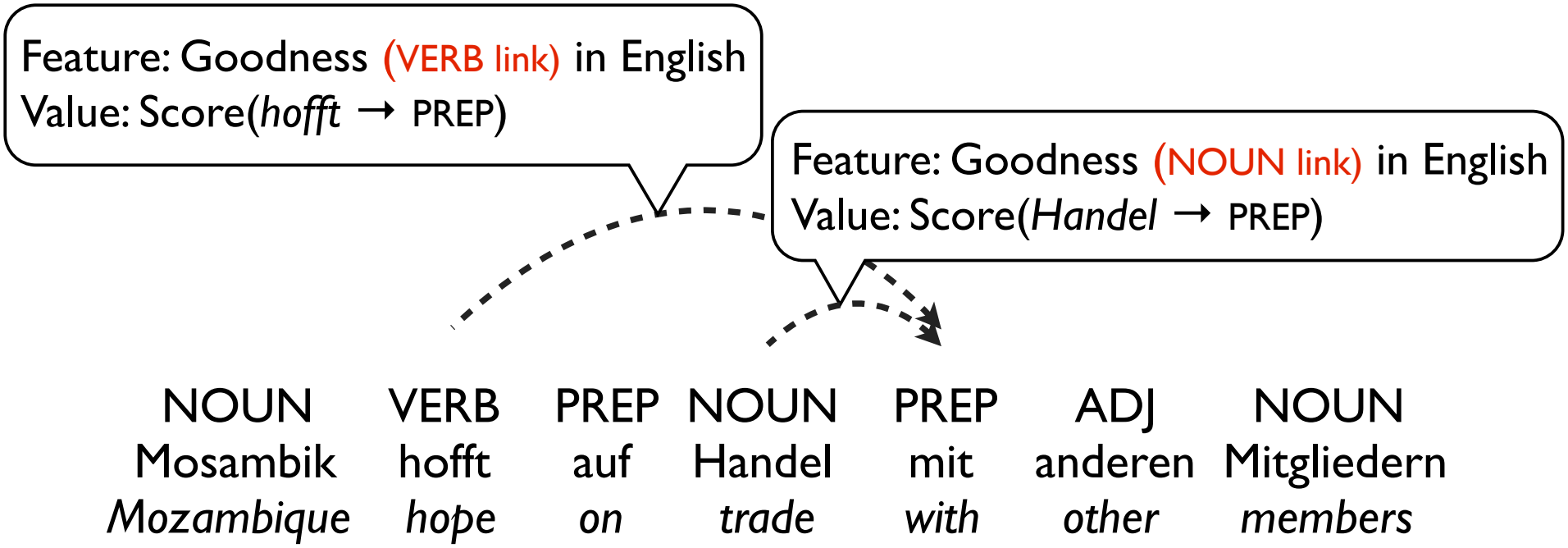
Feature: Goodness in English
Value: Score(*Handel* → PREP)

NOUN	VERB	PREP	NOUN	PREP	ADJ	NOUN
Mosambik	hofft	auf	Handel	mit	anderen	Mitgliedern
Mozambique	*hope*	*on*	*trade*	*with*	*other*	*members*

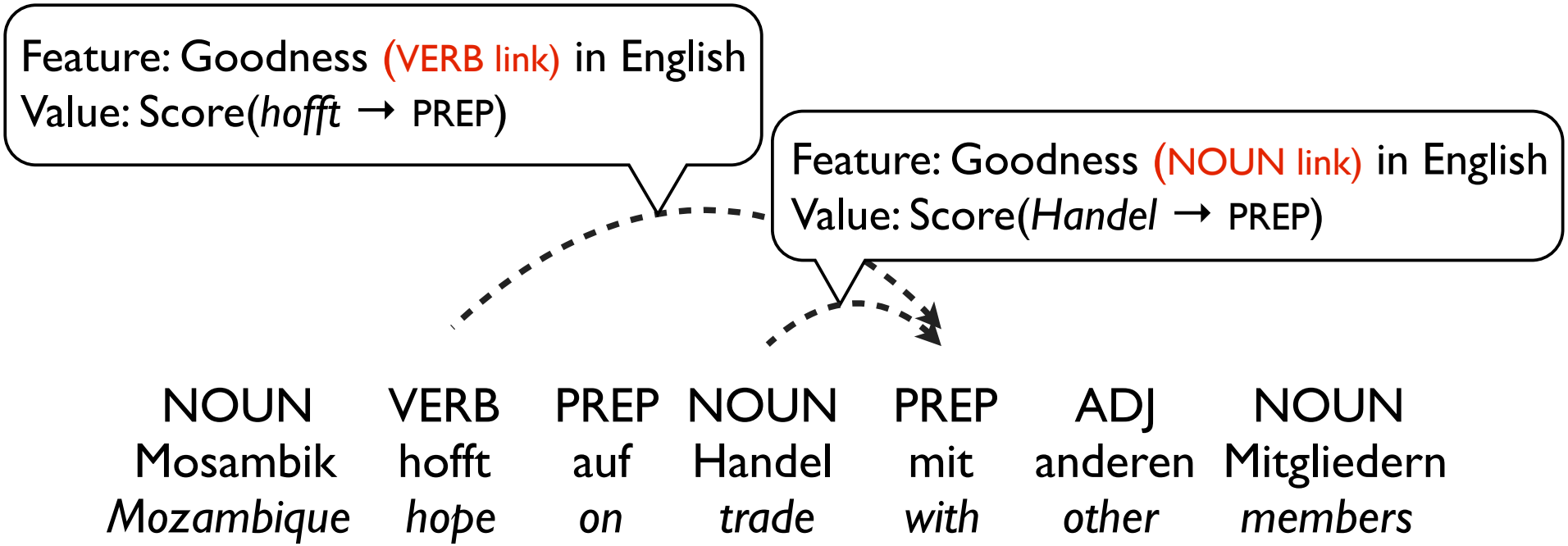
Feature Specialization



Feature Specialization



Feature Specialization



- Similar to selective sharing of Naseem et al. (2012)

Experiments

- Sources of bilingual dictionaries
- Training
- Results

MANUAL Lexicon

Handel

See also **handel**

Contents [hide]

German

Etymology

From handeln, later specialized to its current meaning.

Noun

Handel *m*

1. deal
2. trade, trading

MANUAL Lexicon

Handel

See also **handel**

Contents [hide]

German

Etymology

From handeln, later specialized to its current meaning.

Noun

Handel *m*

1. deal
2. trade, trading

MANUAL Lexicon

Handel

See also **handel**

Contents [hide]

German

Etymology

From handeln, later specialized to its current meaning.

Noun

Handel *m*

1. deal
2. trade, trading

w_s	$p(w_s \mid Handel)$
deal	0.333
trade	0.333
trading	0.333

AUTOMATIC Lexicon

Da ist zum Beispiel der Handel mit Drogen

Trafficking in drugs is one example

Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Ich will mit dem Handel anfangen

I will start with trade

AUTOMATIC Lexicon

Da ist zum Beispiel der Handel mit Drogen

Trafficking in drugs is one example

w_s	$p(w_s \mid Handel)$
trade	0.66
trafficking	0.33

Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Ich will mit dem Handel anfangen

I will start with trade

AUTOMATIC Lexicon

Da ist zum Beispiel der Handel mit Drogen

Trafficking in drugs is one example

Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Ich will mit dem Handel anfangen

I will start with trade

w_s	$p(w_s \mid Handel)$
trade	0.768
trafficking	0.057
commerce	0.05
trading	0.03
market	0.02
marketing	0.02
business	0.013
traffic	0.007
sales	0.002
deal	0.002
...	...

Small Target Training Set

- 8 target languages, CoNLL shared task treebanks
  - Train on 400 trees from CoNLL training set
  - Test on CoNLL test set

Small Target Training Set

- 8 target languages, CoNLL shared task treebanks
  - Train on 400 trees from CoNLL training set
  - Test on CoNLL test set

- Universal POS tags generated from gold tags using mapping of Petrov et al. (2011)

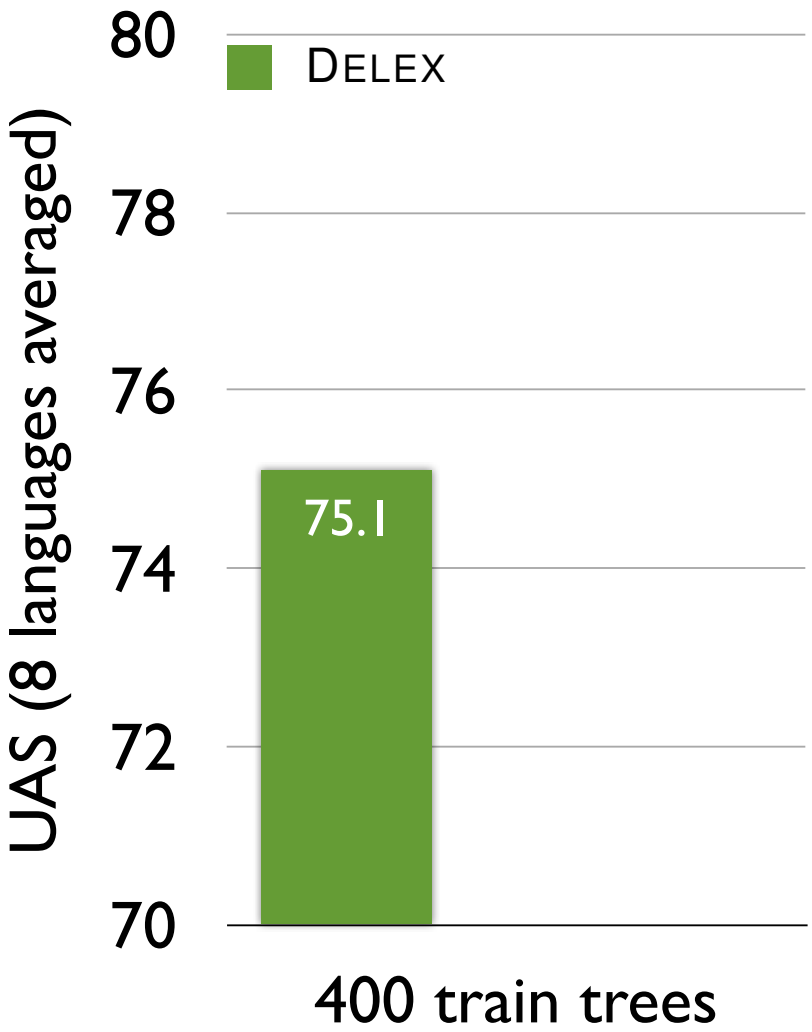
Small Target Training Set

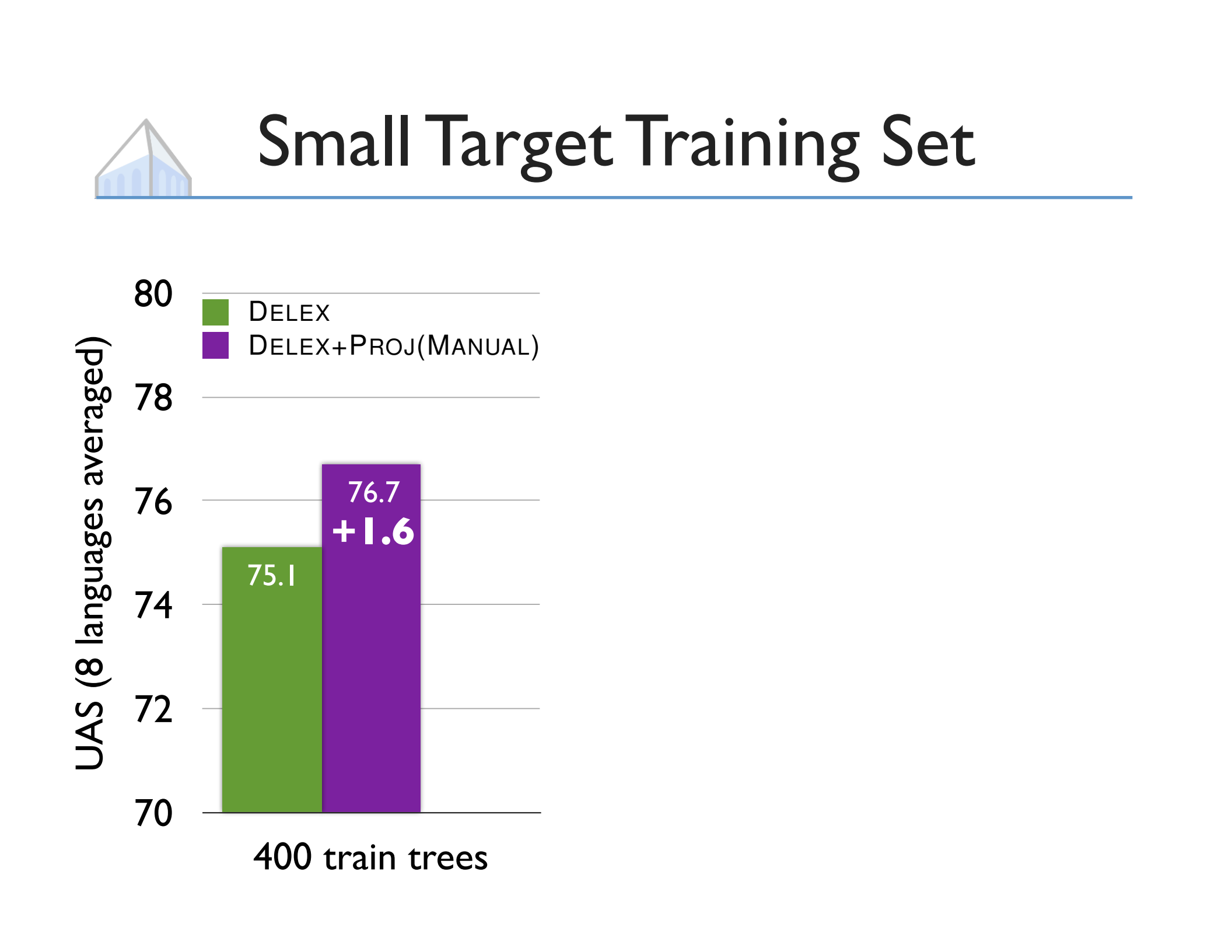
- 8 target languages, CoNLL shared task treebanks
  - Train on 400 trees from CoNLL training set
  - Test on CoNLL test set

- Universal POS tags generated from gold tags using mapping of Petrov et al. (2011)

- Source corpus: parsed English Gigaword

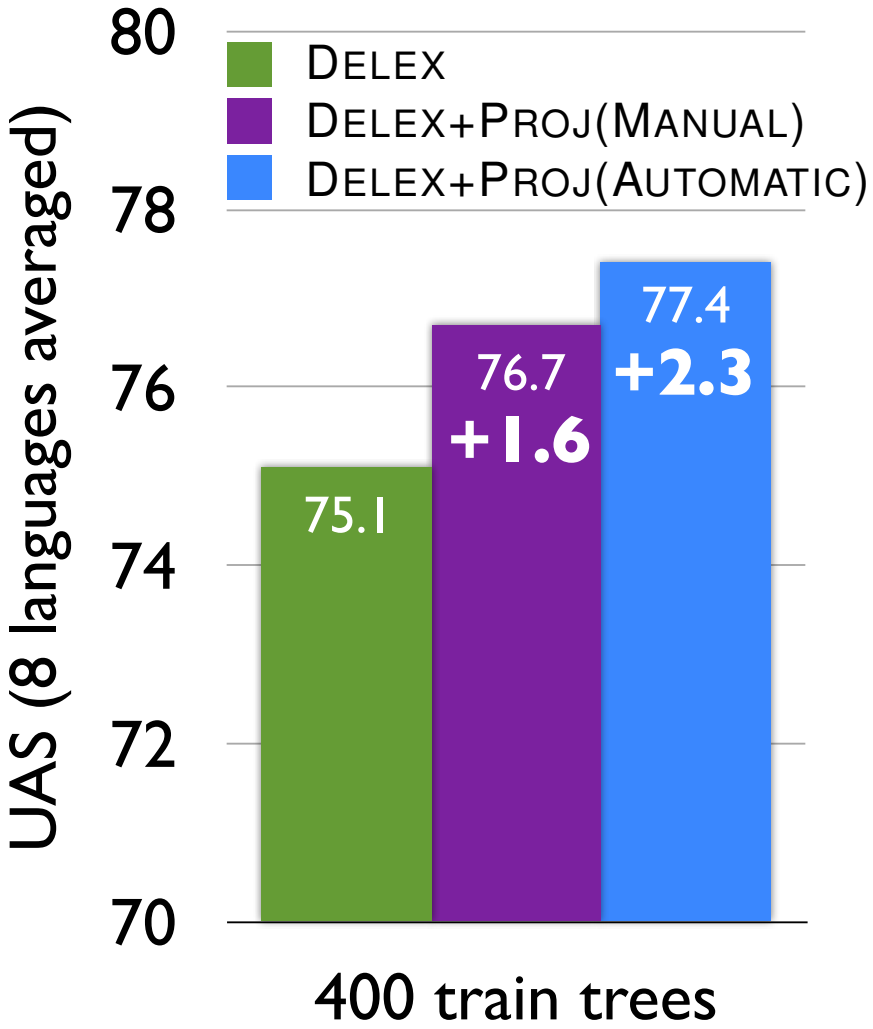
Small Target Training Set



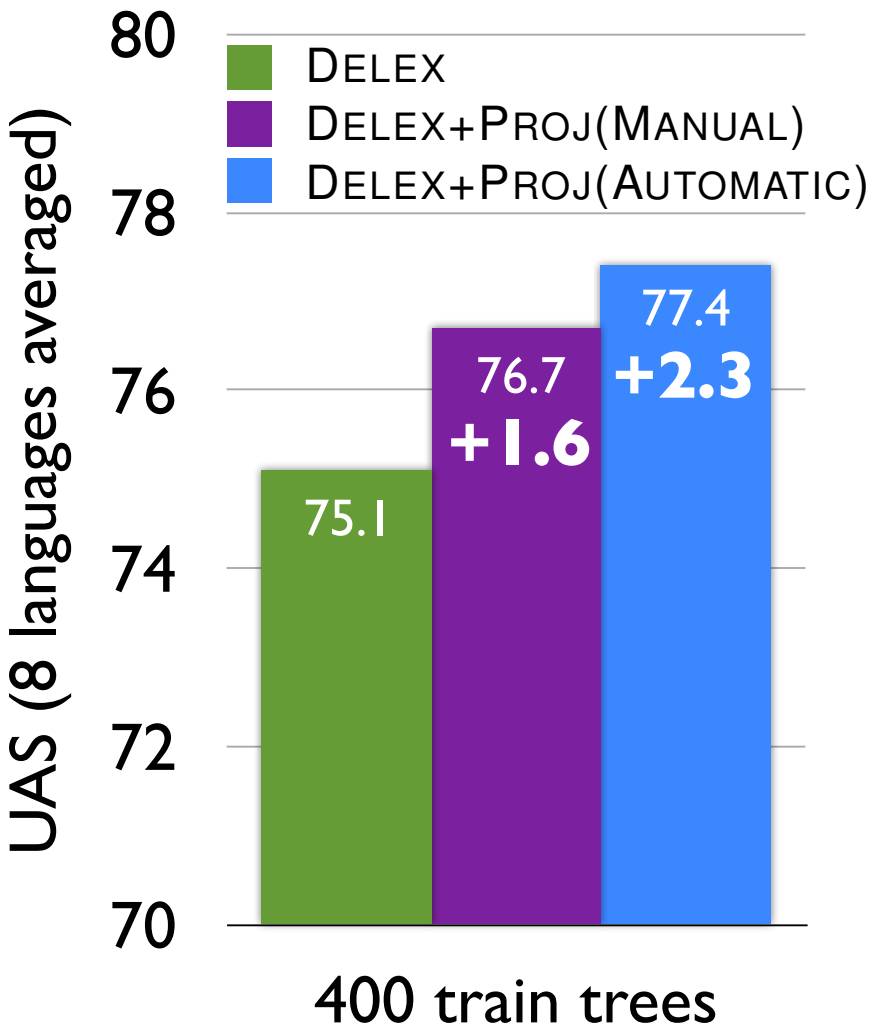


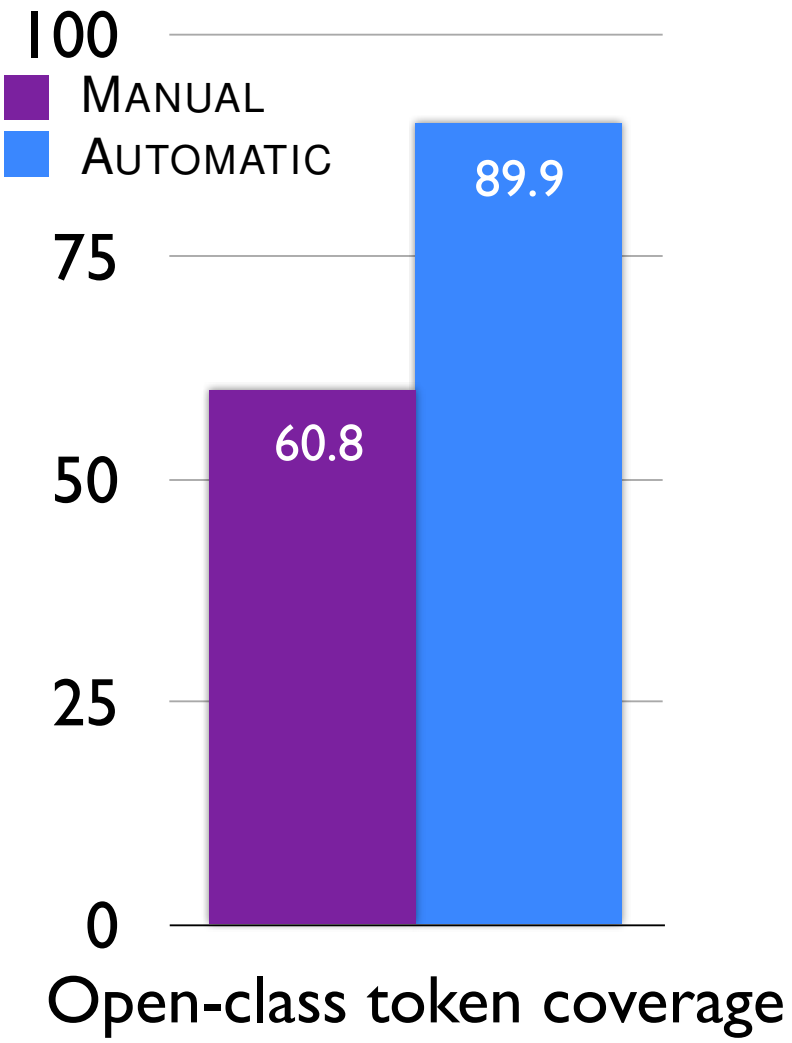
Small Target Training Set

Small Target Training Set

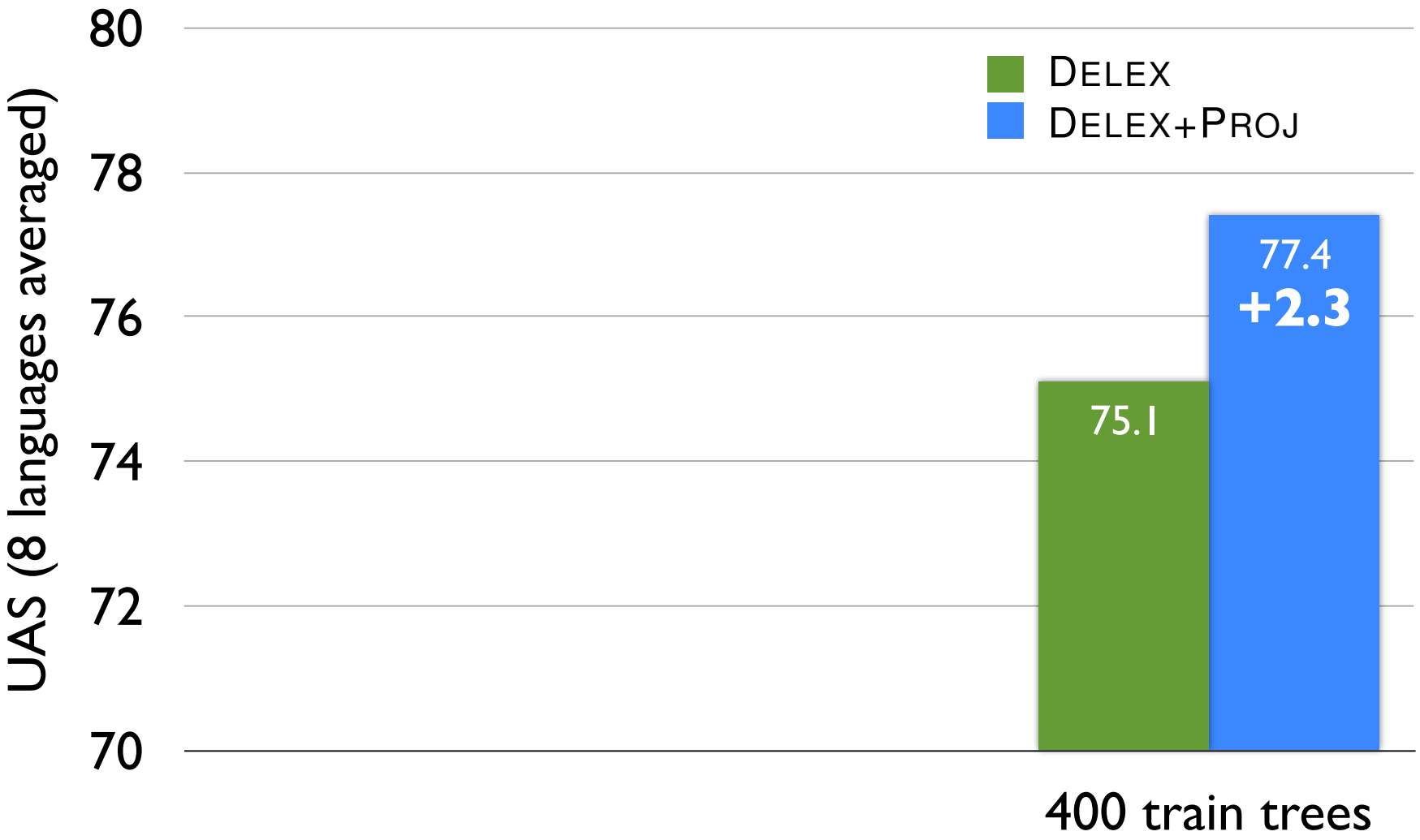


Small Target Training Set

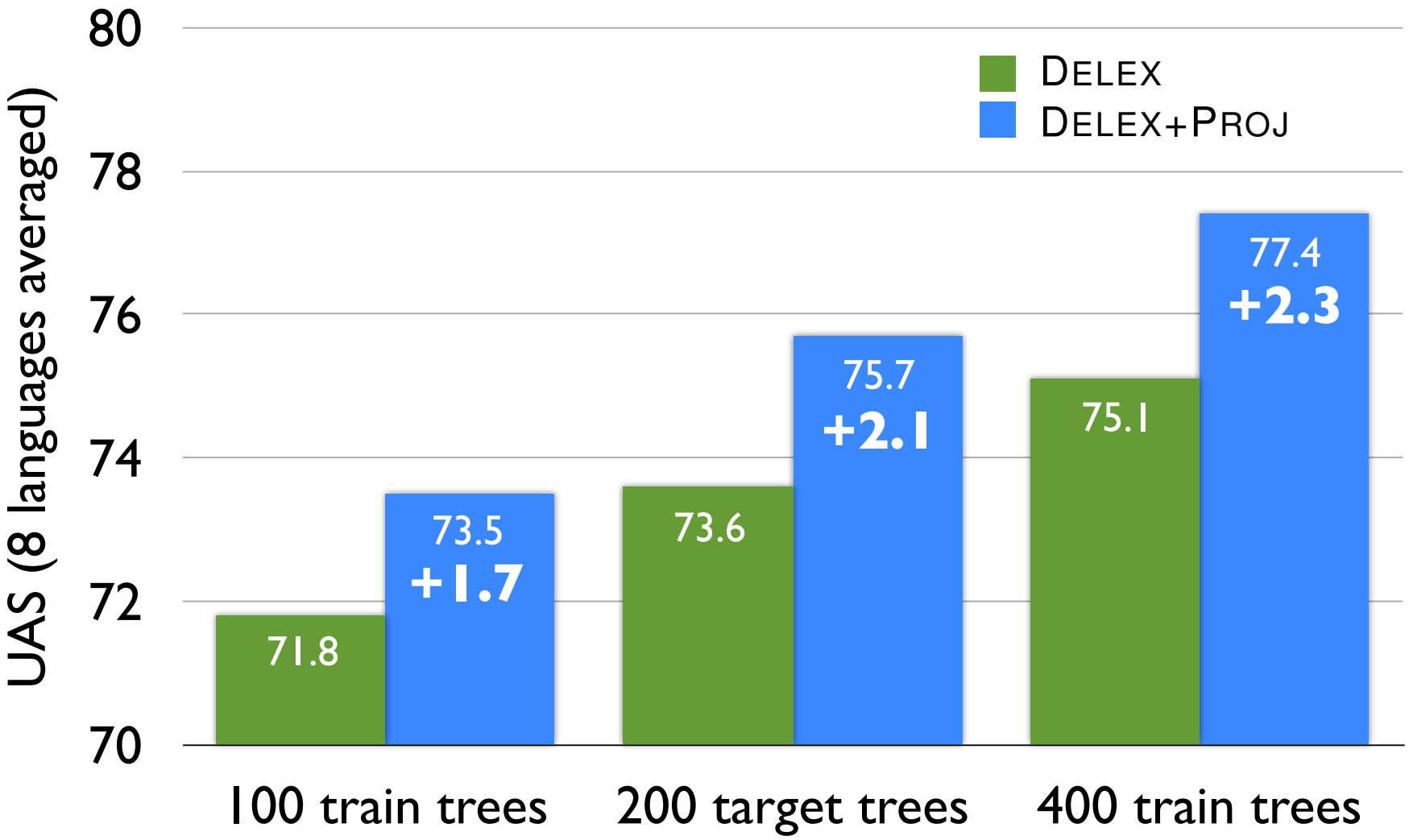




Small Target Training Set



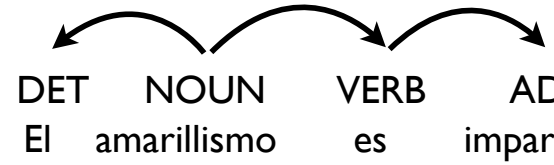
Small Target Training Set



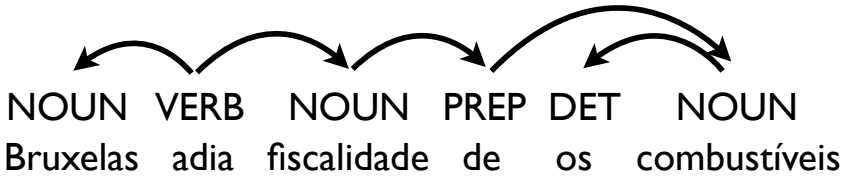
No Target Training Set

Train

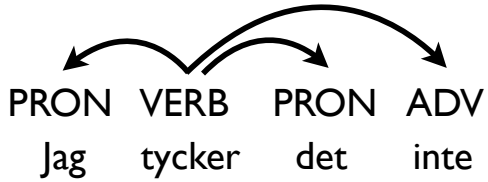
ES trees



PT trees



SV trees



...

Test

DE CoNLL test set

[McDonald et al. (2011)]

No Target Training Set

Train

ES trees

PT trees

SV trees

...

Test

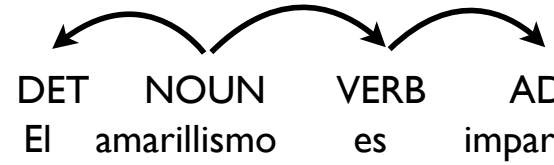
DE CoNLL test set

[McDonald et al. (2011)]

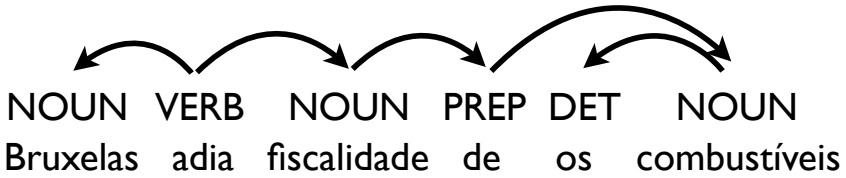
No Target Training Set

Train

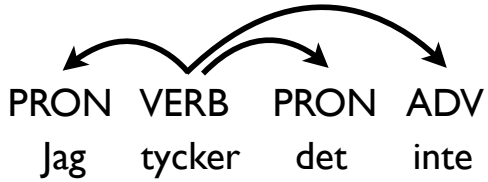
ES trees



PT trees



SV trees



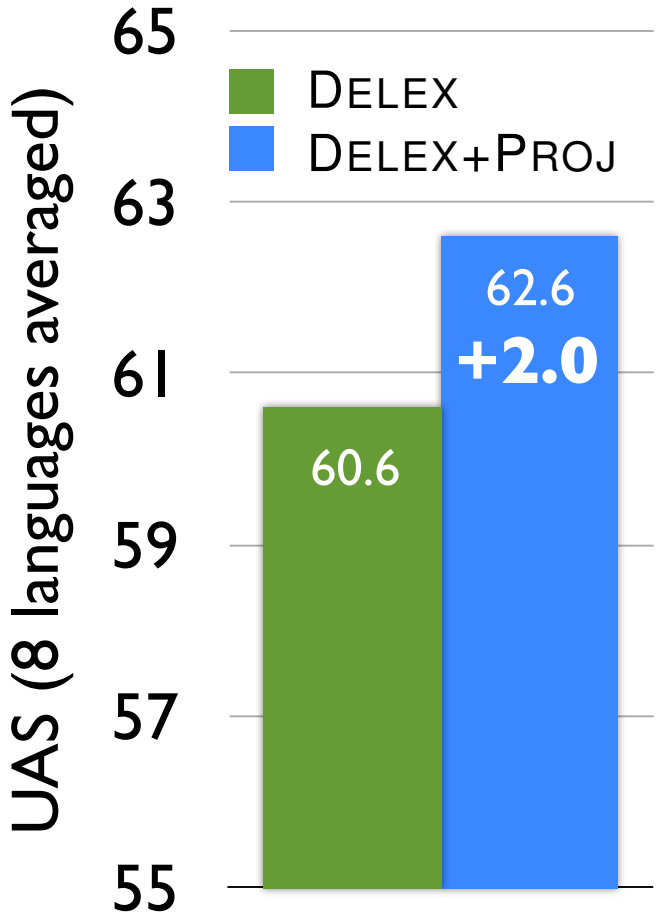
...

Test

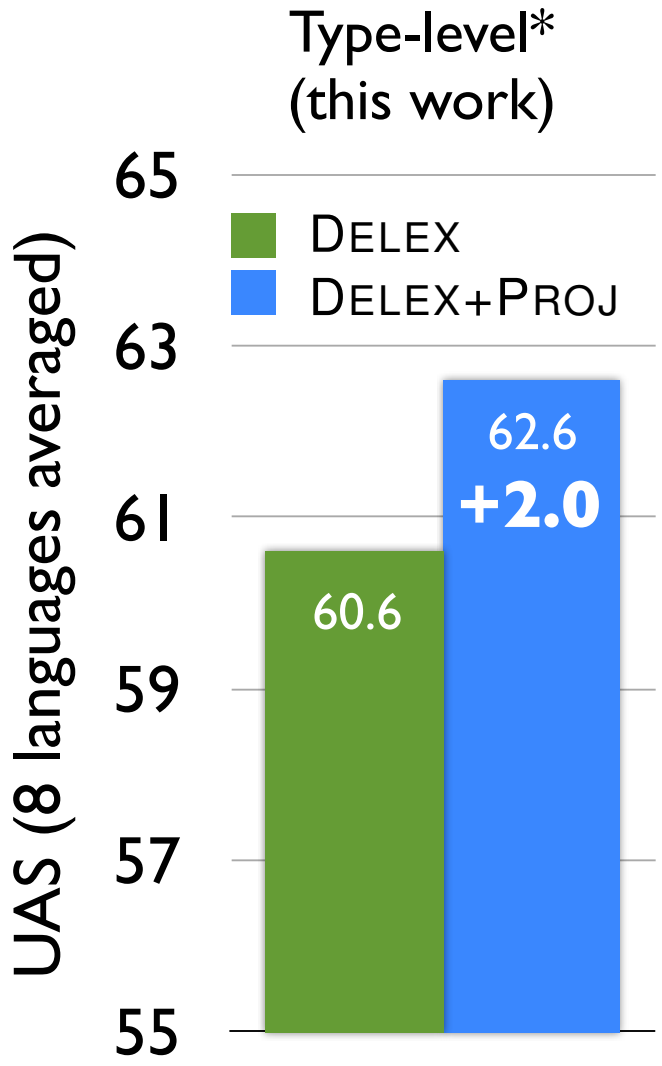
DE CoNLL test set

[McDonald et al. (2011)]

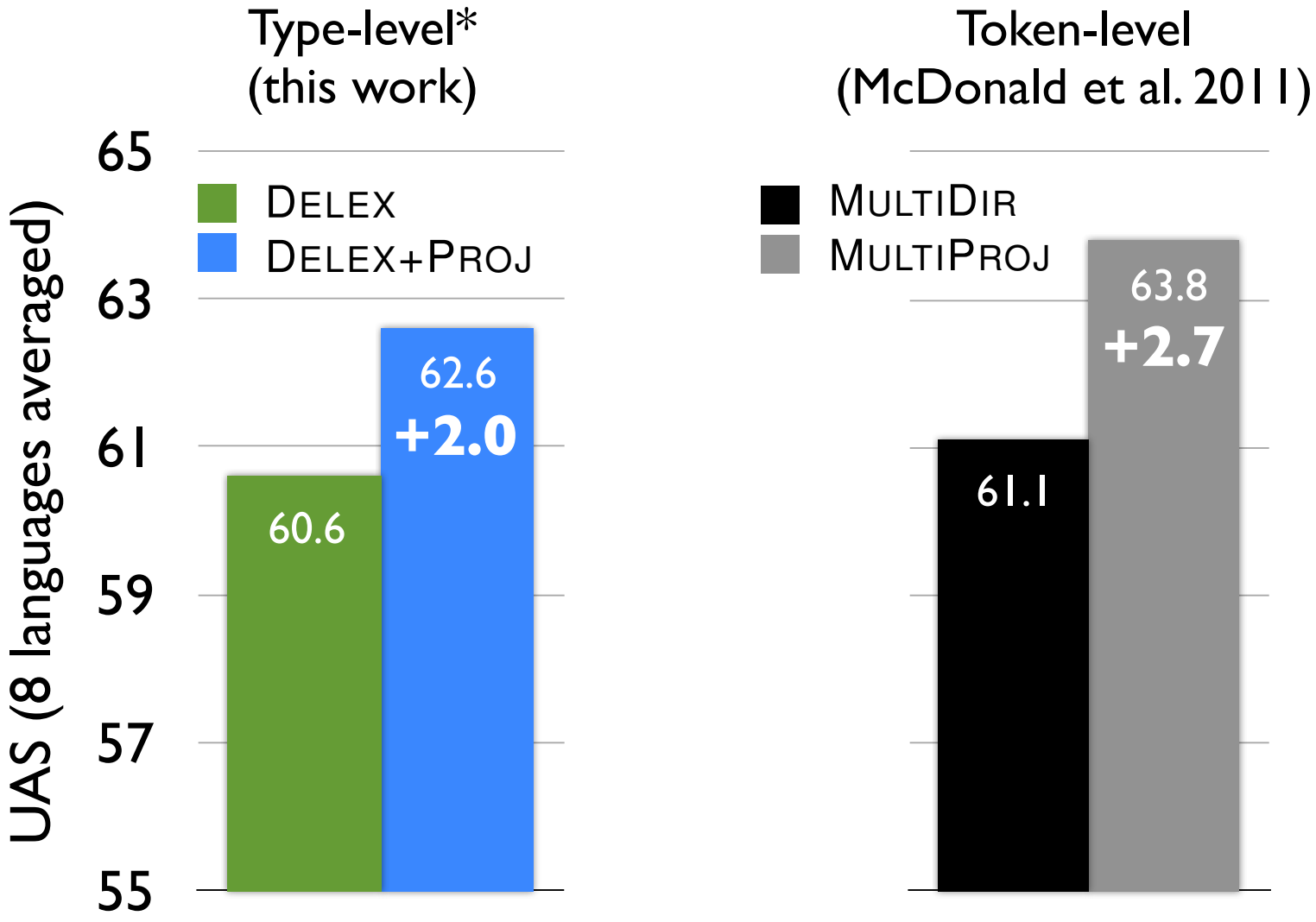
No Target Training Set



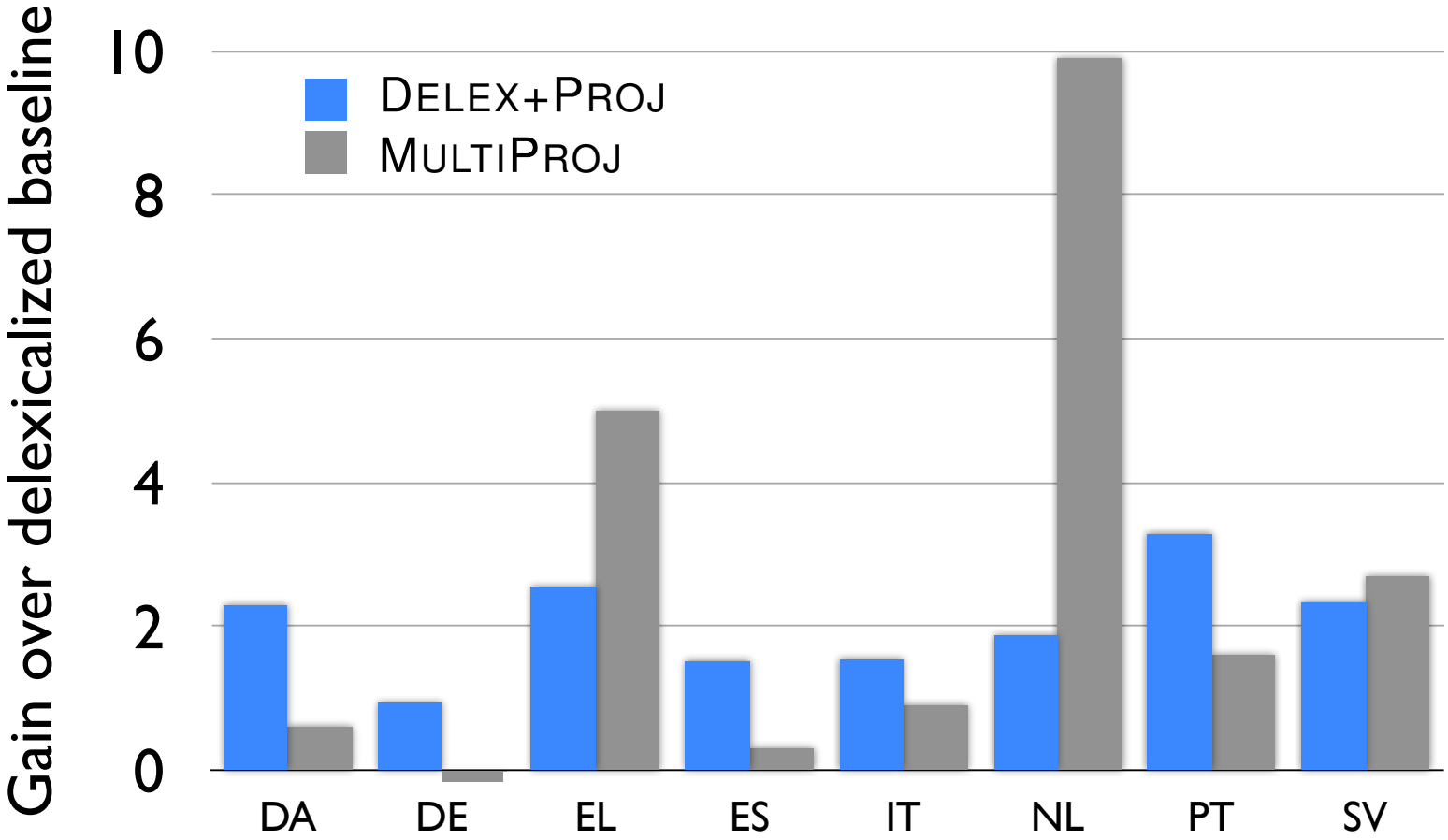
No Target Training Set



No Target Training Set



Gains


Gain over delexicalized baseline
10
8
6
4
2
0
DELEX+PROJ
MULTIPROJ
DA
DE
EL
ES
IT
NL
PT
SV

Related Work

- Täckström et al. (2012): type-level transfer using word clusters

Related Work

- Täckström et al. (2012): type-level transfer using word clusters

- Naseem et al. (2012): adaptation of universal POS transfer to the target language

Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level

Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level

- Type-level transfer:
  - Comparable to token-level transfer
  - Simpler to implement
  - Does not require bitext

Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level

- Type-level transfer:
  - Comparable to token-level transfer
  - Simpler to implement
  - Does not require bitext

Thank you!