

Parameter Sharing in Multilingual Dependency Parsing

MIRYAM DE LHONEUX



UPPSALA
UNIVERSITET

SLTC - 8 November 2018

Work in Collaboration with:



(a) Anders
Soegaard



(b) Johannes
Bjerva



(c) Isabelle
Augenstein

Work in Collaboration with:



(a) Anders
Soegaard



(b) Johannes
Bjerva



(c) Isabelle
Augenstein

Parameter sharing between dependency parsers for related languages

Miryam de Lhoneux^{1*} Johannes Bjerva² Isabelle Augenstein² Anders Søgaard²

¹Department of Linguistics and Philology
Uppsala University
Uppsala, Sweden

²Department of Computer Science
University of Copenhagen
Copenhagen, Denmark

Work in Collaboration with:



(a) Anders
Søgaard



(b) Johannes
Bjerva



(c) Isabelle
Augenstein

Parameter sharing between dependency parsers for related languages

Miryam de Lhoneux^{1*} Johannes Bjerva² Isabelle Augenstein² Anders Søgaard²

¹Department of Linguistics and Philology
Uppsala University
Uppsala, Sweden

²Department of Computer Science
University of Copenhagen
Copenhagen, Denmark

+ Joakim Nivre, Sara Stymne and Aaron Smith

Overview

1 Parsing Architecture

2 Parameter Sharing

3 Results

Outline for section 1

1 Parsing Architecture

2 Parameter Sharing

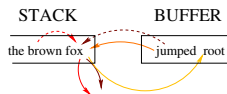
3 Results

Configuration:

STACK	BUFFER
the brown fox	jumped root

Kiperwasser and Goldberg (2016); de Lhoneux et al. (2017)

Configuration:



Transitions:

LEFT-ARC

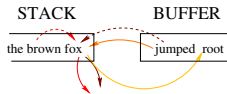
RIGHT-ARC

SHIFT

SWAP

Kiperwasser and Goldberg (2016); de Lhoneux et al. (2017)

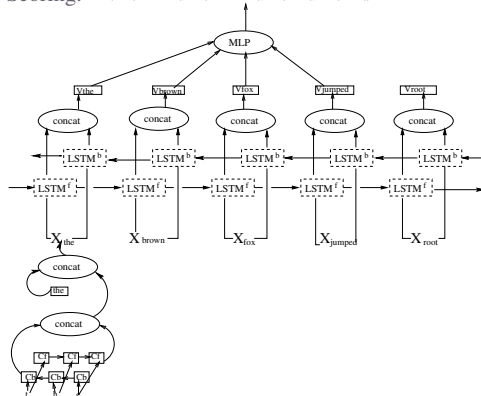
Configuration:



Transitions:

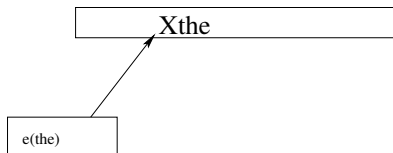
LEFT-ARC
RIGHT-ARC
SHIFT
SWAP

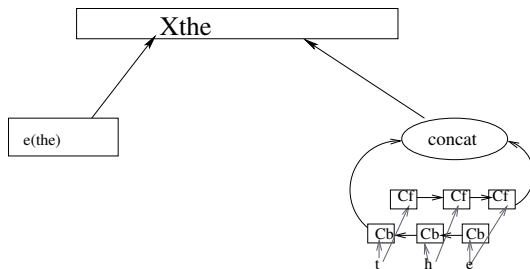
Scoring:

$$(\text{score}(\text{LEFT-ARC}), \text{score}(\text{RIGHT-ARC}), \text{score}(\text{SHIFT}), \text{score}(\text{SWAP}))$$


Kiperwasser and Goldberg (2016); de Lhoneux et al. (2017)

Xthe





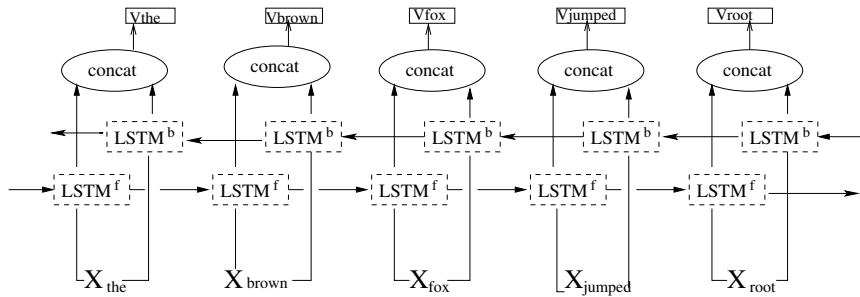
X_{the}

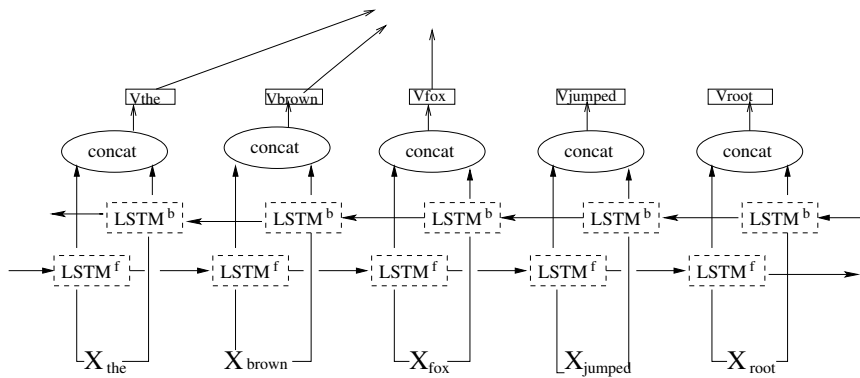
X_{brown}

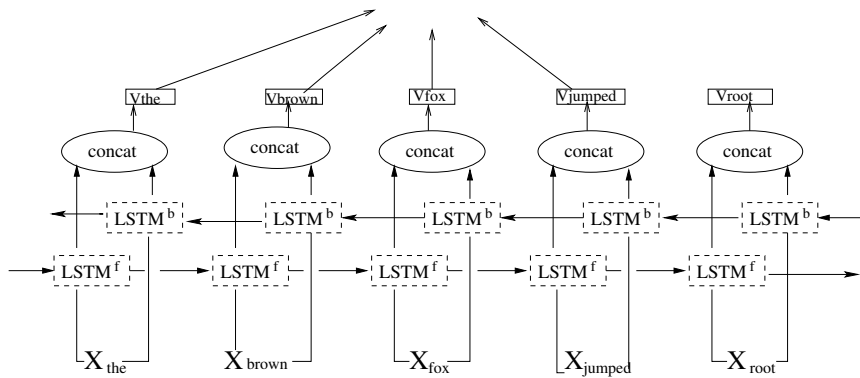
X_{fox}

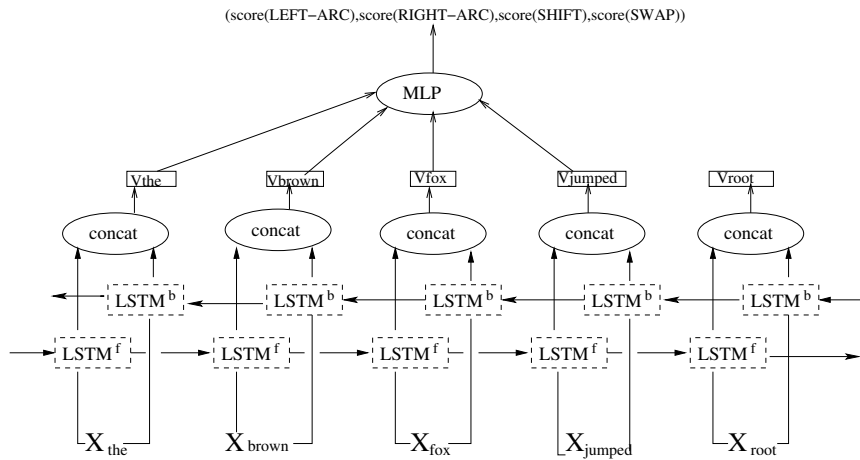
X_{jumped}

X_{root}



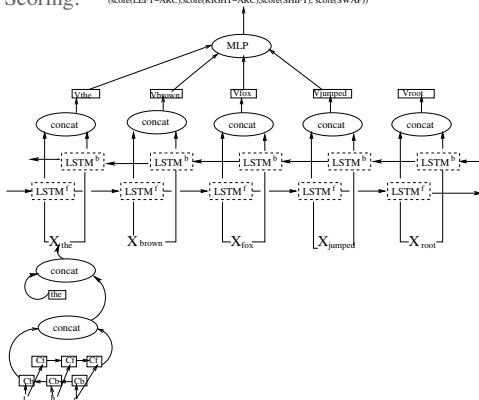






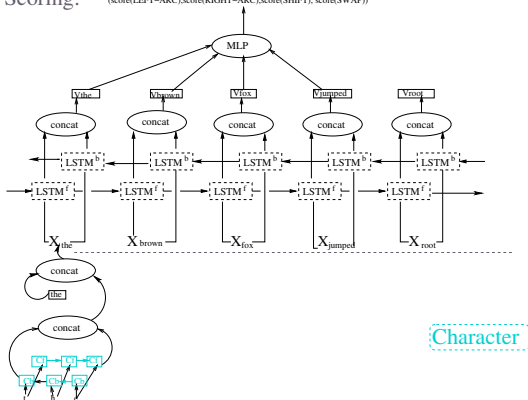
Scoring:

(score(LEFT-ARC), score(RIGHT-ARC), score(SHIFT), score(SWAP))



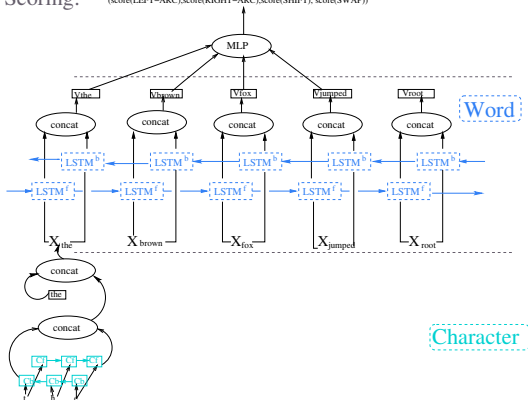
Scoring:

$(\text{score}(\text{LEFT-ARC}), \text{score}(\text{RIGHT-ARC}), \text{score}(\text{SHIFT}), \text{score}(\text{SWAP}))$



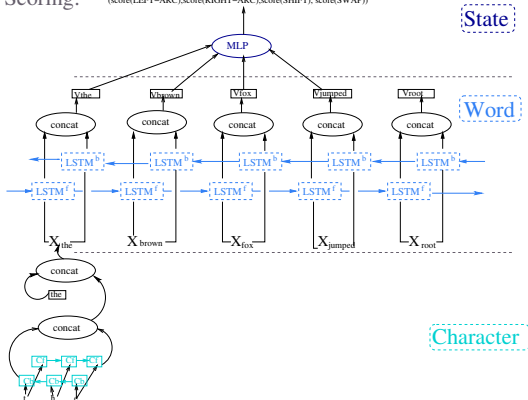
Scoring:

$(\text{score}(\text{LEFT-ARC}), \text{score}(\text{RIGHT-ARC}), \text{score}(\text{SHIFT}), \text{score}(\text{SWAP}))$



Scoring:

$(\text{score}(\text{LEFT}-\text{ARC}), \text{score}(\text{RIGHT}-\text{ARC}), \text{score}(\text{SHIFT}), \text{score}(\text{SWAP}))$



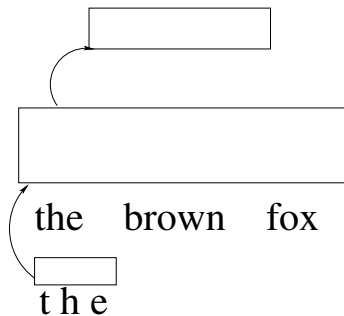
Outline for section 2

1 Parsing Architecture

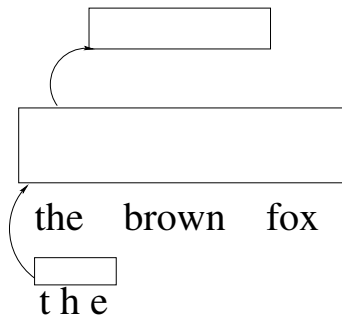
2 **Parameter Sharing**

3 Results

Parameter Sharing



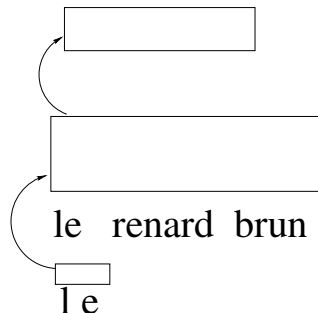
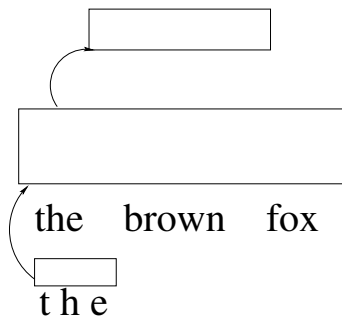
Parameter Sharing



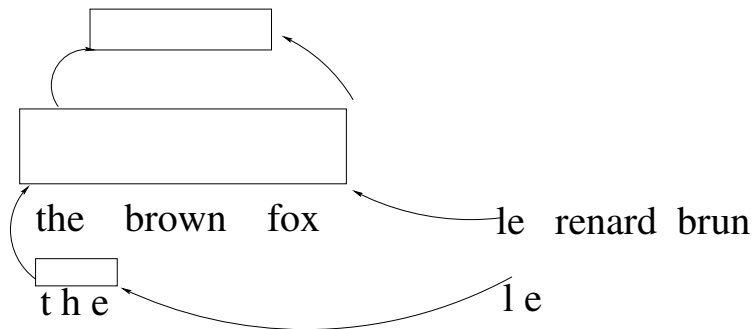
le renard brun

l e

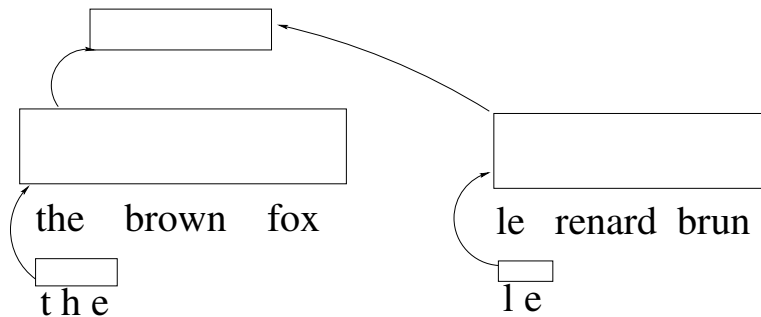
Parameter Sharing



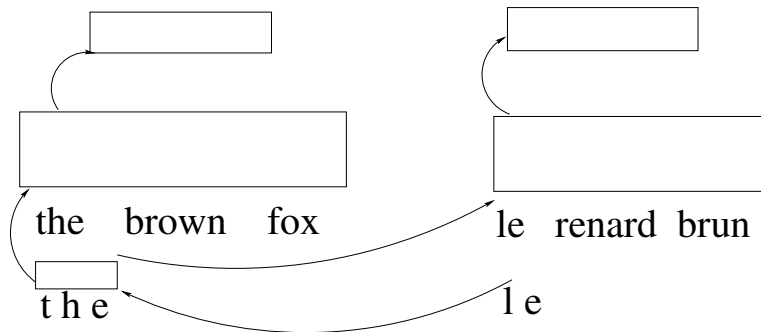
Parameter Sharing



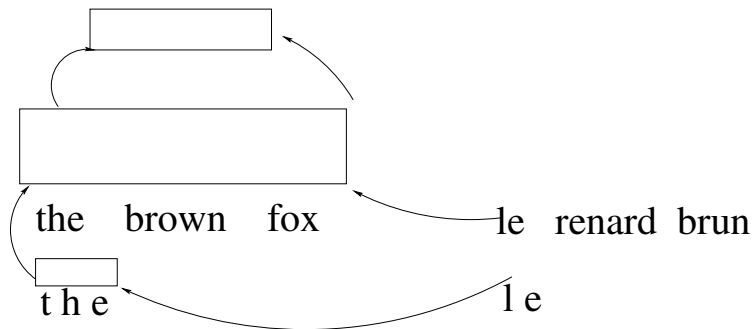
Parameter Sharing



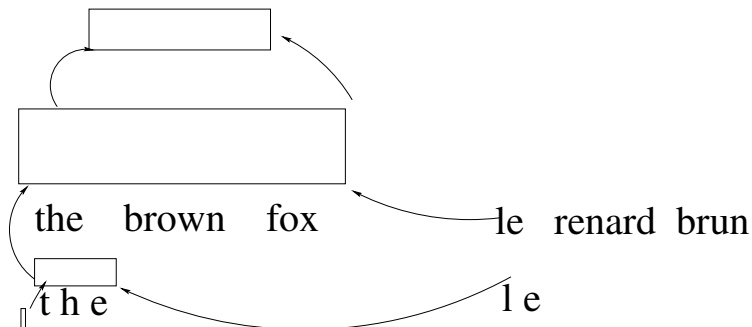
Parameter Sharing



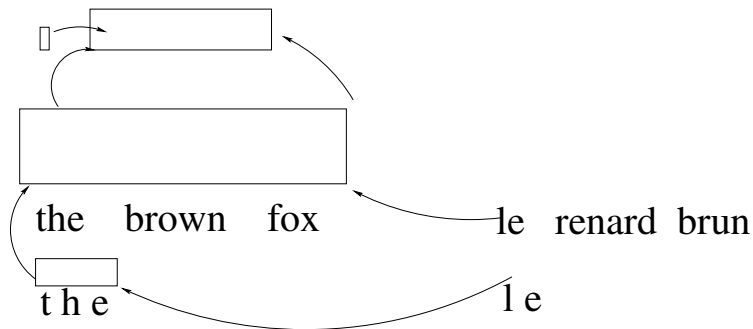
Parameter Sharing



Parameter Sharing



Parameter Sharing



Experimental setup

- 3 types of sharing: hard, soft, not

Experimental setup

- 3 types of sharing: hard, soft, not
- 3 sets of parameters: MLP, word, char

Experimental setup

- 3 types of sharing: hard, soft, not
- 3 sets of parameters: MLP, word, char
- $3^3 = 27$ combinations

Experimental setup

Lang	#sen	#tok	Group	Word order
Arabic	5,000	208,932	Semitic	VSO
Hebrew	5,000	161,685	Semitic	SVO
Estonian	5,000	60,393	Finnic	SVO
Finnish	5,000	67,258	Finnic	SVO
Croatian	5,000	109,965	Slavic	SVO
Russian	5,000	90,170	Slavic	SVO
Italian	5,000	113,825	Romance	SVO
Spanish	5,000	154,844	Romance	SVO
Dutch	5,000	75,796	Germanic	No dom. order
Norwegian	5,000	76,622	Germanic	SVO

Table: Dataset characteristics

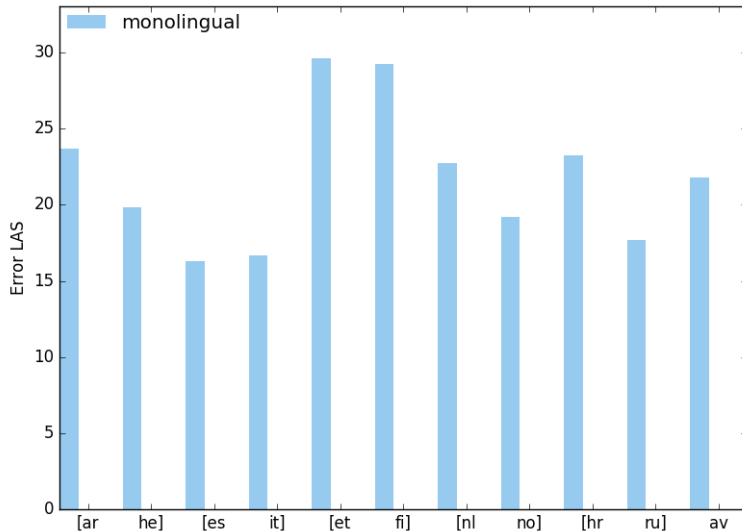
Outline for section 3

1 Parsing Architecture

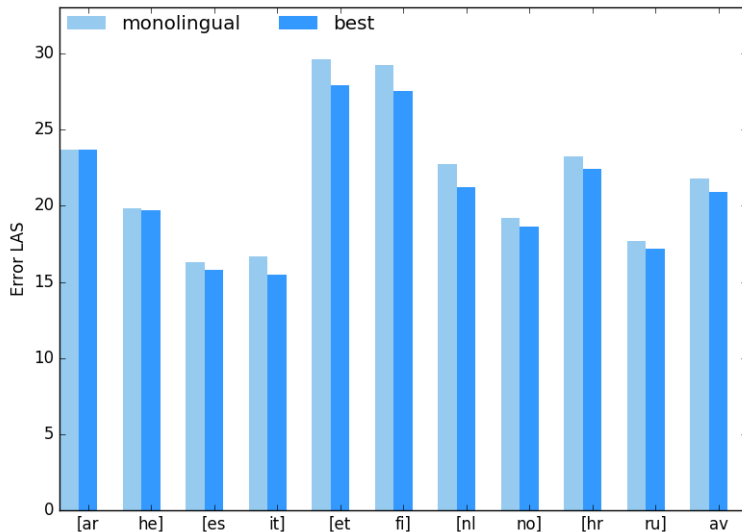
2 Parameter Sharing

3 Results

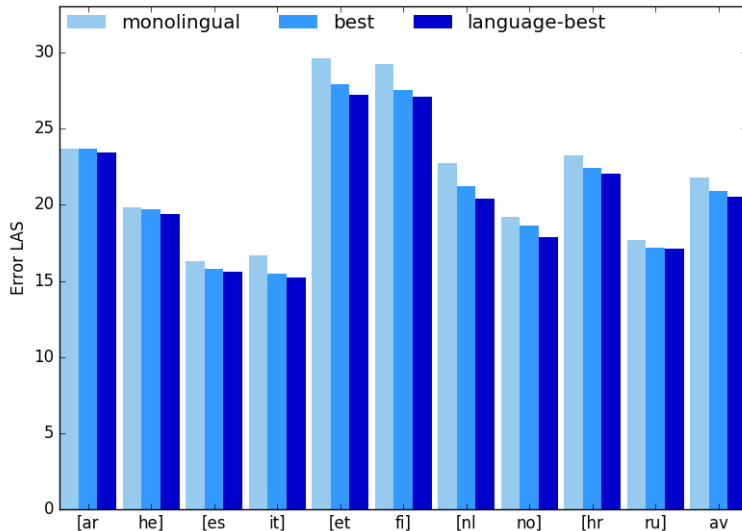
Results



Results



Results



- Sharing the MLP is helpful

Results

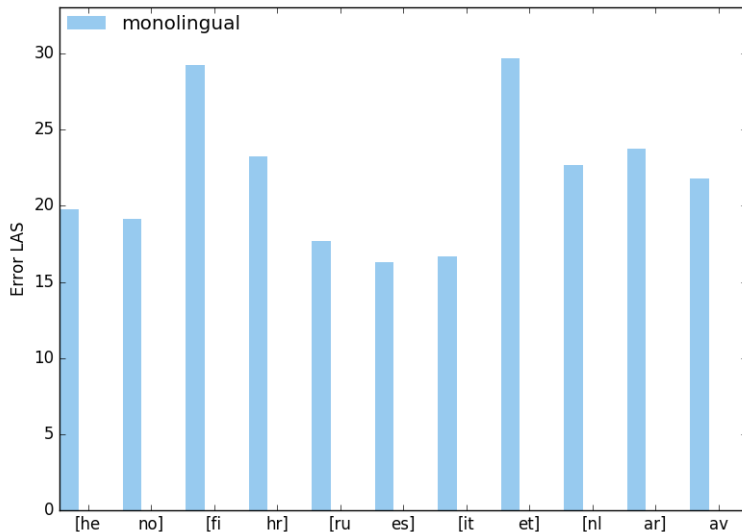
- Sharing the MLP is helpful
- Sharing Word and Characters depends on language pair

Results

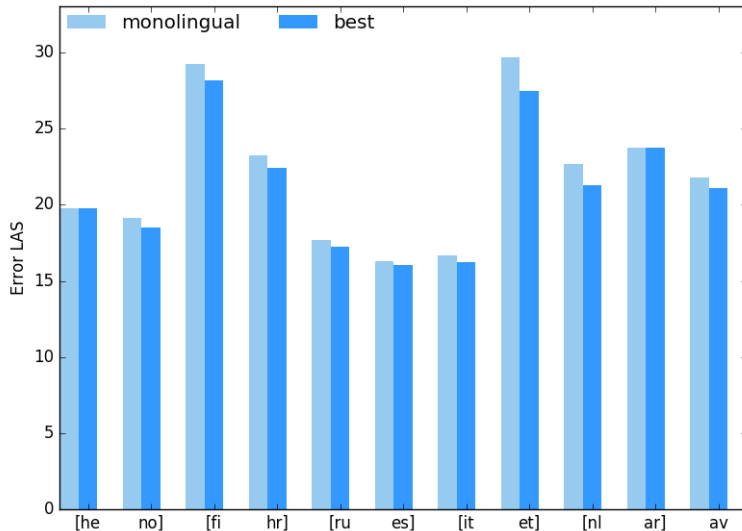
	W	C	OURS	MONO	δ
ar	X	X	77.2	77.1	0.1
he	✓	X	80.0	79.8	0.3
et	X	ID	71.4	70.5	0.8
fi	X	X	71.6	71.6	0.1
hr	✓	X	77.9	78.0	-0.1
ru	✓	X	83.5	82.7	0.8
it	ID	✓	85.0	84.0	1.0
es	ID	✓	84.3	83.8	0.5
nl	ID	✓	75.5	74.1	1.4
no	X	ID	81.1	80.1	1.0
av.			78.8	78.2	0.6

Table: LAS on the test sets of the best of 9 sharing strategies and the monolingual baseline.

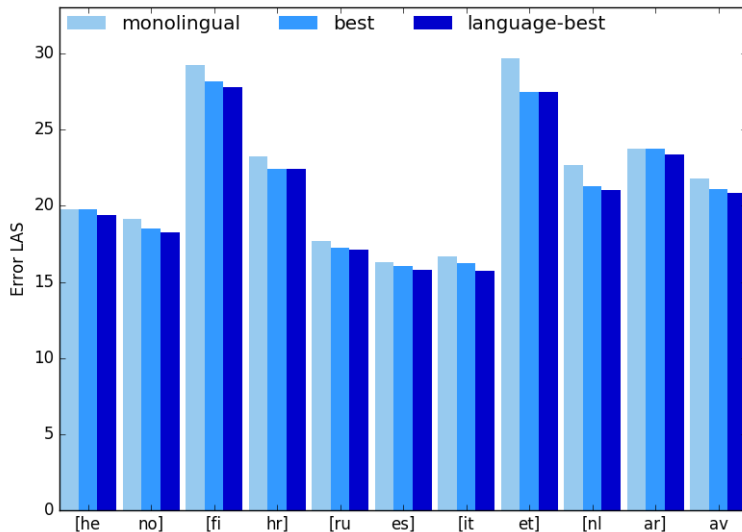
Unrelated languages



Unrelated languages



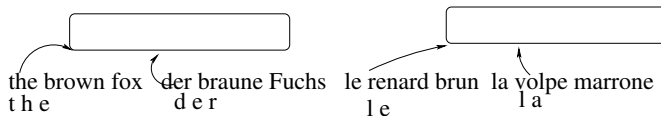
Unrelated languages



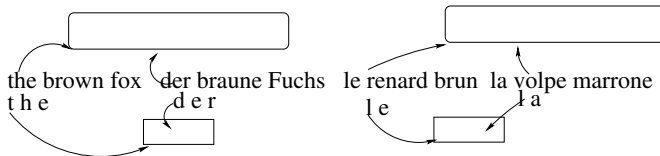
A Universal Parser?

the brown fox	der braune Fuchs	le renard brun	la volpe marrone
the	der	le	la

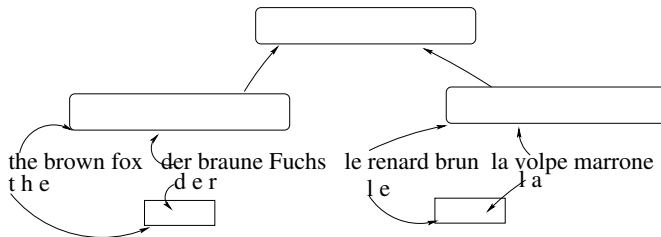
A Universal Parser?



A Universal Parser?



A Universal Parser?



Thanks!

Questions?

References

- Miryam de Lhoneux, Yan Shao, Ali Basirat, Eliyahu Kiperwasser, Sara Stymne, Yoav Goldberg, and Joakim Nivre. 2017. From raw text to universal dependencies - look, no tags! In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Vancouver, Canada, pages 207–217.
- Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and accurate dependency parsing using bidirectional LSTM feature representations. *Transactions of the Association for Computational Linguistics* 4:313–327.