Tree Transformations for Parsing

2 Case Studies

Miryam de Lhoneux

UPPSALA
UNIVERSITET

June 7, 2016

Overview

Outline for section 1

Introduction

- Common methodology - different motivation
- NLP engineering:
  Verb groups in UD parsing (de Lhoneux and Nivre, 2016)
- Linguistics (and NLP engineering):
  MWEs in CCG parsing (de Lhoneux, 2014)

Introduction

- Common methodology - different motivation
- NLP engineering:
  Verb groups in UD parsing (de Lhoneux and Nivre, 2016)
- Linguistics (and NLP engineering):
  MWEs in CCG parsing (de Lhoneux, 2014)

Introduction

- Common methodology - different motivation
- NLP engineering:
  Verb groups in UD parsing (de Lhoneux and Nivre, 2016)
- Linguistics (and NLP engineering):
  MWEs in CCG parsing (de Lhoneux, 2014)

Introduction

- Common methodology - different motivation
- NLP engineering:
  Verb groups in UD parsing (de Lhoneux and Nivre, 2016)
- Linguistics (and NLP engineering):
  MWEs in CCG parsing (de Lhoneux, 2014)

Outline for section 2

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups (Nilsson et al., 2006, 2007; Schwartz et al., 2012)

Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups (Nilsson et al., 2006, 2007; Schwartz et al., 2012)

Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups (Nilsson et al., 2006, 2007; Schwartz et al., 2012)

Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups
(Nilsson et al., 2006, 2007; Schwartz et al., 2012)

Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups
(Nilsson et al., 2006, 2007; Schwartz et al., 2012)

Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups (Nilsson et al., 2006, 2007; Schwartz et al., 2012)

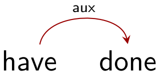


Figure : MS

Figure : UD

Verb Groups in UD parsing

Introduction

- problem: UD suboptimal for parsing (de Marneffe et al., 2014)
- solution: Create a parsing representation (de Marneffe et al., 2014)
- focus of the study: verb groups

Auxiliary as heads of verb groups
(Nilsson et al., 2006, 2007; Schwartz et al., 2012)

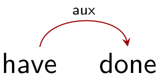


Figure : MS

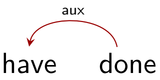
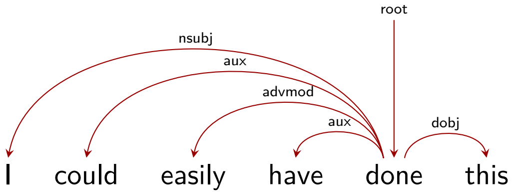


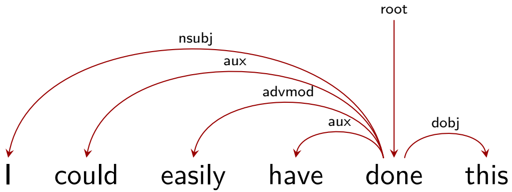
Figure : UD

Transformation Algorithm:
Modified from Nilsson et al. (2006)

Transformation Algorithm

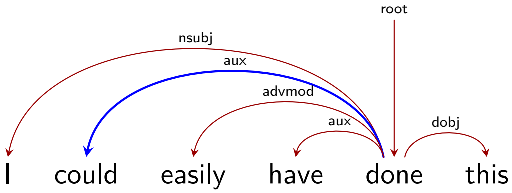


Transformation Algorithm



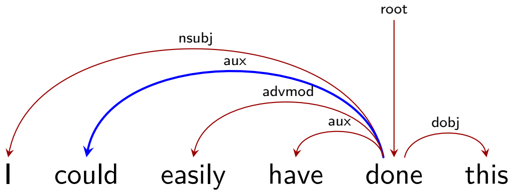
$V = []$

Transformation Algorithm



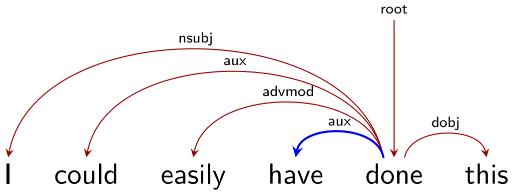
V = []

Transformation Algorithm



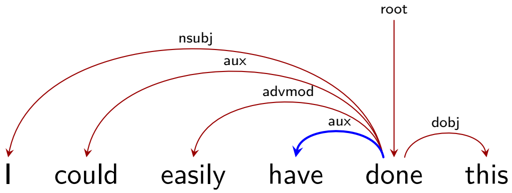
V = [V_i: main verb=done; aux = [could]]

Transformation Algorithm



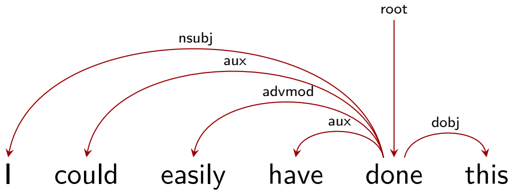
V = [V_i: main verb=done; aux = [could]]

Transformation Algorithm

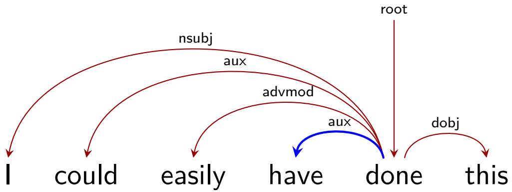


V = [V_i: main verb=done; aux = [could, have]]

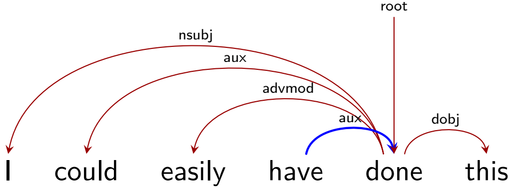
Transformation Algorithm



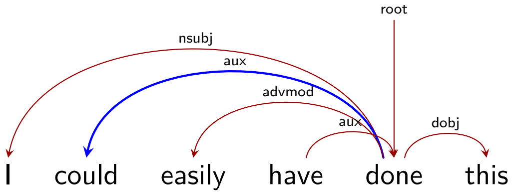
Transformation Algorithm



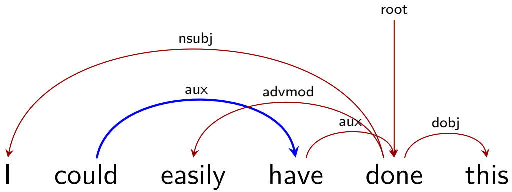
Transformation Algorithm



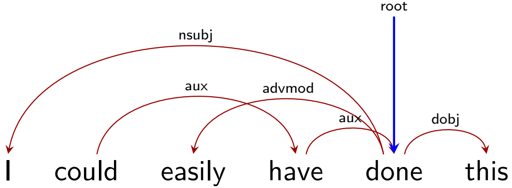
Transformation Algorithm



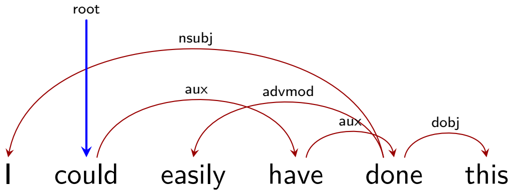
Transformation Algorithm



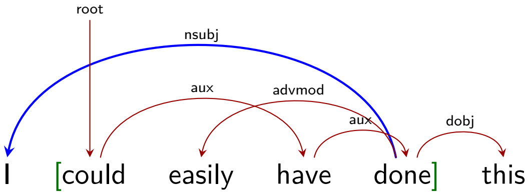
Transformation Algorithm



Transformation Algorithm

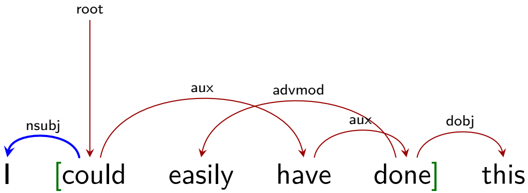


Transformation Algorithm



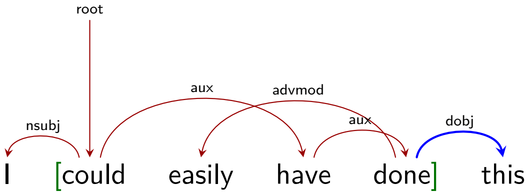
left dependents

Transformation Algorithm



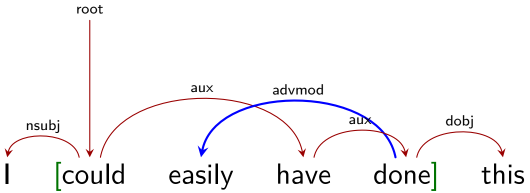
left dependents

Transformation Algorithm



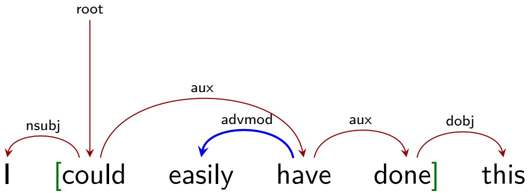
right dependents

Transformation Algorithm



middle dependents

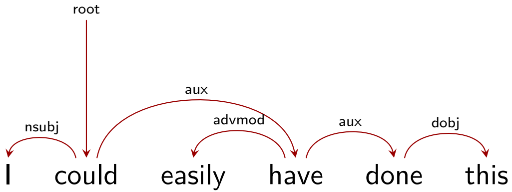
Transformation Algorithm



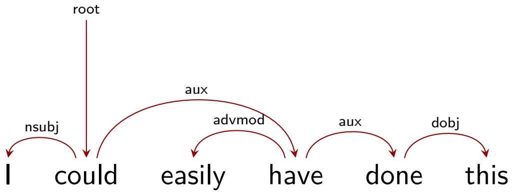
middle dependents

Back Transformation Algorithm
Modified from Nilsson et al. (2006)

Back Transformation Algorithm

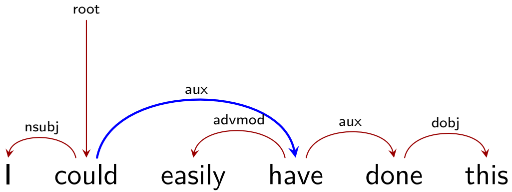


Back Transformation Algorithm



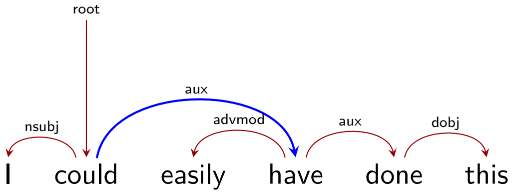
V = []

Back Transformation Algorithm



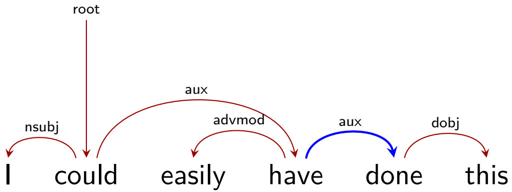
V = []

Back Transformation Algorithm



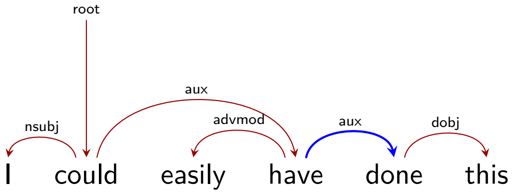
V = [V_i: aux=[could]]

Back Transformation Algorithm



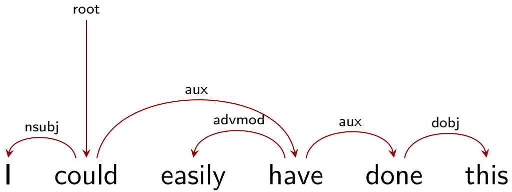
V = [V_i: aux=[could]]

Back Transformation Algorithm

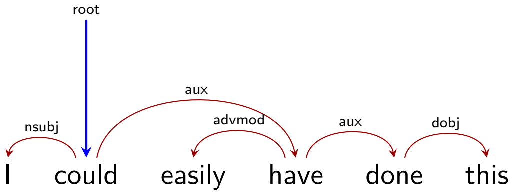


V = [V_i: aux=[could,have];main verb=done]

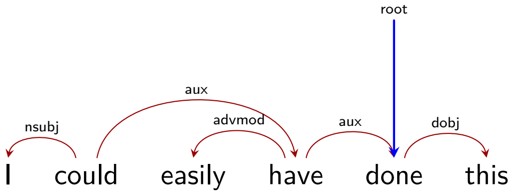
Back Transformation Algorithm



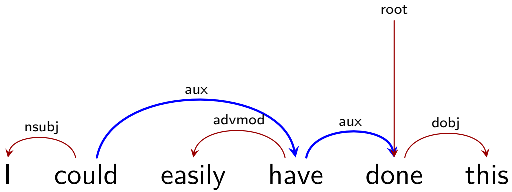
Back Transformation Algorithm



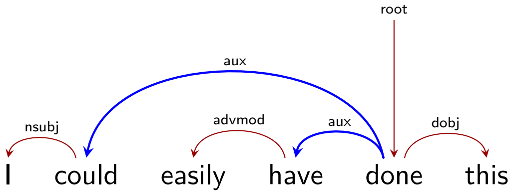
Back Transformation Algorithm



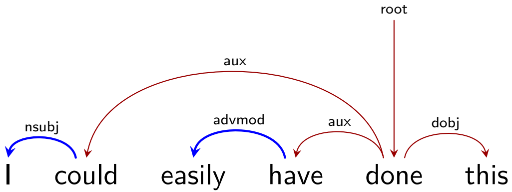
Back Transformation Algorithm



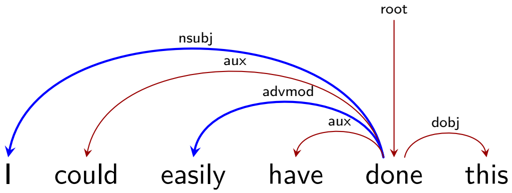
Back Transformation Algorithm



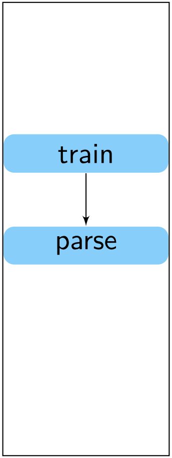
Back Transformation Algorithm



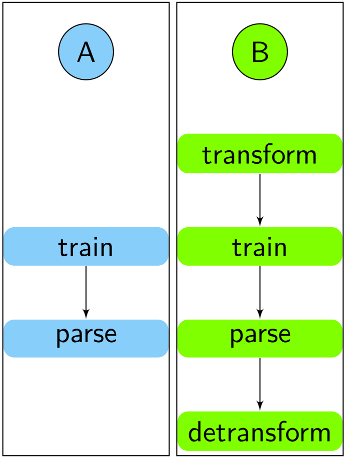
Back Transformation Algorithm



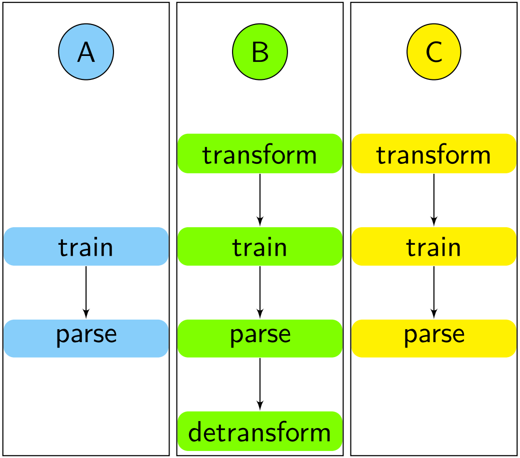
Pipeline



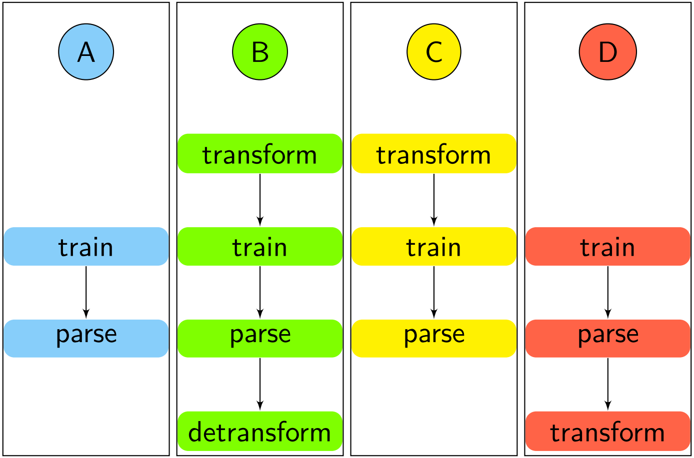
Pipeline



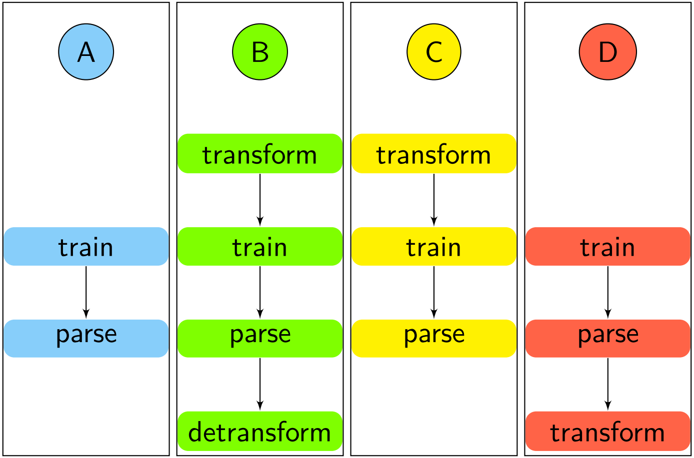
Pipeline


A
train
parse
B
transform
train
parse
detransform
C
transform
train
parse

Pipeline


A
train
parse
B
transform
train
parse
detransform
C
transform
train
parse
D
train
parse
transform

Pipeline



Model/Gold	UD	MS
UD	A	B
MS	D	C

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing

B > A

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing

B > A

MS is easier to learn than UD

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing — B > A

MS is easier to learn than UD — C > A

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing B > A

MS is easier to learn than UD C > A

Symmetry in differences

Hypotheses

Model/Gold	UD	MS
UD	A	B
MS	D	C

MS is better than UD for parsing $B > A$

MS is easier to learn than UD $C > A$

Symmetry in differences $A - B = C - D$

Resources

- Parser: MaltParser (Nivre et al., 2006)
  - settings: default
  - POS tag used: UD coarse POS tag
- Data: 25 UD treebanks (version 1.2) + SDT and PDT

Resources

- Parser: MaltParser (Nivre et al., 2006)
  - settings: default
  - POS tag used: UD coarse POS tag
- Data: 25 UD treebanks (version 1.2) + SDT and PDT

Resources

- Parser: MaltParser (Nivre et al., 2006)
  - settings: default
  - POS tag used: UD coarse POS tag
- Data: 25 UD treebanks (version 1.2) + SDT and PDT

Resources

- Parser: MaltParser (Nivre et al., 2006)
  - settings: default
  - POS tag used: UD coarse POS tag
- Data: 25 UD treebanks (version 1.2) + SDT and PDT

Resources

- Parser: MaltParser (Nivre et al., 2006)
  - settings: default
  - POS tag used: UD coarse POS tag
- Data: 25 UD treebanks (version 1.2) + SDT and PDT

Effect of VG transformation on UD

UD lang.	A	B
Basque	**64.4**	63.8**
Bulgarian	**83.4**	83.2*
Croatian	**75.9**	74.6**
Czech	**80.0**	76.5**
Danish	**75.9**	75.2**
English	**81.7**	80.4**
Estonian	77.1	**77.8**
Finnish	**66.9**	66.4*
Finnish-FTB	**71.3**	70.4**
French	**82.1**	81.6**
German	**76.6**	76.0**
Greek	75.2	**75.3**
Hebrew	**78.4**	77.9**
Hindi	**85.4**	84.2**
Italian	**83.8**	83.6
Norwegian	**84.5**	82.0**
Old_Church_Slavonic	**68.8**	68.7
Persian	**81.1**	79.8**
Polish	**79.4**	79.1
Portuguese	81.3	**81.5**
Romanian	**64.2**	62.5*
Slovenian	**80.8**	79.7**
Spanish	**81.5**	81.2**
Swedish	**76.8**	75.7**
Tamil	**67.2**	67.1

Results

UD language	A	B
Basque	**64.4**	63.8
Bulgarian	**83.4**	83.2
Croatian	**75.9**	74.6
Czech	**80.0**	76.5
Danish	**75.9**	75.2
English	**81.7**	80.4
Estonian	77.1	**77.8**
Finnish	**66.9**	66.4
Finnish-FTB	**71.3**	70.4
French	**82.1**	81.6
German	**76.6**	76.0
Greek	75.2	**75.3**
Hebrew	**78.4**	77.9
Hindi	**85.4**	84.2
Italian	**83.8**	83.6
Norwegian	**84.5**	82.0
OC_Slavonic	**68.8**	68.7
Persian	**81.1**	79.8
Polish	**79.4**	79.1
Portuguese	81.3	**81.5**
Romanian	**64.2**	62.5
Slovenian	**80.8**	79.7
Spanish	**81.5**	81.2
Swedish	**76.8**	75.7
Tamil	**67.2**	67.1

Results

UD language	A	B
Basque	**64.4**	63.8
Bulgarian	**83.4**	83.2
Croatian	**75.9**	74.6
Czech	**80.0**	76.5
Danish	**75.9**	75.2
English	**81.7**	80.4
Estonian	77.1	**77.8**
Finnish	**66.9**	66.4
Finnish-FTB	**71.3**	70.4
French	**82.1**	81.6
German	**76.6**	76.0
Greek	75.2	**75.3**
Hebrew	**78.4**	77.9
Hindi	**85.4**	84.2
Italian	**83.8**	83.6
Norwegian	**84.5**	82.0
OC_Slavonic	**68.8**	68.7
Persian	**81.1**	79.8
Polish	**79.4**	79.1
Portuguese	81.3	**81.5**
Romanian	**64.2**	62.5
Slovenian	**80.8**	79.7
Spanish	**81.5**	81.2
Swedish	**76.8**	75.7
Tamil	**67.2**	67.1

MS is better than UD for parsing B > A

Results

UD language	A	B
Basque	**64.4**	63.8
Bulgarian	**83.4**	83.2
Croatian	**75.9**	74.6
Czech	**80.0**	76.5
Danish	**75.9**	75.2
English	**81.7**	80.4
Estonian	77.1	**77.8**
Finnish	**66.9**	66.4
Finnish-FTB	**71.3**	70.4
French	**82.1**	81.6
German	**76.6**	76.0
Greek	75.2	**75.3**
Hebrew	**78.4**	77.9
Hindi	**85.4**	84.2
Italian	**83.8**	83.6
Norwegian	**84.5**	82.0
OC_Slavonic	**68.8**	68.7
Persian	**81.1**	79.8
Polish	**79.4**	79.1
Portuguese	81.3	**81.5**
Romanian	**64.2**	62.5
Slovenian	**80.8**	79.7
Spanish	**81.5**	81.2
Swedish	**76.8**	75.7
Tamil	**67.2**	67.1

MS is better than UD for parsing B > A ✗

Results

UD language	A	B	C
Basque	**64.4**	63.8	64.0
Bulgarian	**83.4**	83.2	82.5
Croatian	**75.9**	74.6	73.7
Czech	**80.0**	76.5	76.4
Danish	**75.9**	75.2	74.8
English	**81.7**	80.4	80.2
Estonian	77.1	77.8	**77.6**
Finnish	**66.9**	66.4	65.9
Finnish-FTB	71.3	70.4	**72.1**
French	**82.1**	81.6	81.3
German	**76.6**	76.0	75.4
Greek	**75.2**	75.3	75.1
Hebrew	**78.4**	77.9	77.9
Hindi	**85.4**	84.2	84.9
Italian	**83.8**	83.6	83.3
Norwegian	**84.5**	82.0	81.7
OC_Slavonic	**68.8**	68.7	68.7
Persian	**81.1**	79.8	79.8
Polish	**79.4**	79.1	79.0
Portuguese	81.3	81.5	**81.6**
Romanian	**64.2**	62.5	64.0
Slovenian	**80.8**	79.7	79.8
Spanish	**81.5**	81.2	81.2
Swedish	**76.8**	75.7	75.6
Tamil	67.2	67.1	**67.4**

MS is better than UD for parsing   B > A   ✗

Results

UD language	A	B	C
Basque	**64.4**	63.8	64.0
Bulgarian	**83.4**	83.2	82.5
Croatian	**75.9**	74.6	73.7
Czech	**80.0**	76.5	76.4
Danish	**75.9**	75.2	74.8
English	**81.7**	80.4	80.2
Estonian	77.1	77.8	**77.6**
Finnish	**66.9**	66.4	65.9
Finnish-FTB	71.3	70.4	**72.1**
French	**82.1**	81.6	81.3
German	**76.6**	76.0	75.4
Greek	**75.2**	75.3	75.1
Hebrew	**78.4**	77.9	77.9
Hindi	**85.4**	84.2	84.9
Italian	**83.8**	83.6	83.3
Norwegian	**84.5**	82.0	81.7
OC_Slavonic	**68.8**	68.7	68.7
Persian	**81.1**	79.8	79.8
Polish	**79.4**	79.1	79.0
Portuguese	81.3	81.5	**81.6**
Romanian	**64.2**	62.5	64.0
Slovenian	**80.8**	79.7	79.8
Spanish	**81.5**	81.2	81.2
Swedish	**76.8**	75.7	75.6
Tamil	67.2	67.1	**67.4**

MS is better than UD for parsing — B > A ✗

MS is easier to learn than UD — C > A

Results

UD language	A	B	C
Basque	**64.4**	63.8	64.0
Bulgarian	**83.4**	83.2	82.5
Croatian	**75.9**	74.6	73.7
Czech	**80.0**	76.5	76.4
Danish	**75.9**	75.2	74.8
English	**81.7**	80.4	80.2
Estonian	77.1	77.8	**77.6**
Finnish	**66.9**	66.4	65.9
Finnish-FTB	71.3	70.4	**72.1**
French	**82.1**	81.6	81.3
German	**76.6**	76.0	75.4
Greek	**75.2**	75.3	75.1
Hebrew	**78.4**	77.9	77.9
Hindi	**85.4**	84.2	84.9
Italian	**83.8**	83.6	83.3
Norwegian	**84.5**	82.0	81.7
OC_Slavonic	**68.8**	68.7	68.7
Persian	**81.1**	79.8	79.8
Polish	**79.4**	79.1	79.0
Portuguese	81.3	81.5	**81.6**
Romanian	**64.2**	62.5	64.0
Slovenian	**80.8**	79.7	79.8
Spanish	**81.5**	81.2	81.2
Swedish	**76.8**	75.7	75.6
Tamil	67.2	67.1	**67.4**

MS is better than UD for parsing — B > A ✗

MS is easier to learn than UD — C > A ✗

Results

UD language	A	B	C	D
Basque	**64.4**	63.8	64.0	**64.4**
Bulgarian	**83.4**	83.2	82.5	**82.9**
Croatian	**75.9**	74.6	73.7	**75.9**
Czech	**80.0**	76.5	76.4	**79.9**
Danish	**75.9**	75.2	74.8	**75.8**
English	**81.7**	80.4	80.2	**81.5**
Estonian	77.1	**77.8**	**77.6**	77.0
Finnish	**66.9**	66.4	65.9	**66.4**
Finnish-FTB	**71.3**	70.4	72.1	**72.5**
French	**82.1**	81.6	81.3	**81.8**
German	**76.6**	76.0	75.4	**76.1**
Greek	75.2	**75.3**	75.1	**75.2**
Hebrew	**78.4**	77.9	77.9	**78.5**
Hindi	**85.4**	84.2	84.9	**85.2**
Italian	**83.8**	83.6	83.3	**83.6**
Norwegian	**84.5**	82.0	81.7	**84.5**
OC_Slavonic	**68.8**	68.7	68.7	**68.9**
Persian	**81.1**	79.8	79.8	**81.1**
Polish	**79.4**	79.1	79.0	**79.3**
Portuguese	81.3	**81.5**	**81.6**	81.3
Romanian	**64.2**	62.5	64.0	**64.6**
Slovenian	**80.8**	79.7	79.8	**80.8**
Spanish	**81.5**	81.2	81.2	**81.4**
Swedish	**76.8**	75.7	75.6	**76.7**
Tamil	67.2	67.1	67.4	**67.5**

MS is better than UD for parsing B > A ✗

MS is easier to learn than UD C > A ✗

Results

UD language	A	B	C	D
Basque	**64.4**	63.8	64.0	**64.4**
Bulgarian	**83.4**	83.2	82.5	**82.9**
Croatian	**75.9**	74.6	73.7	**75.9**
Czech	**80.0**	76.5	76.4	**79.9**
Danish	**75.9**	75.2	74.8	**75.8**
English	**81.7**	80.4	80.2	**81.5**
Estonian	77.1	**77.8**	**77.6**	77.0
Finnish	**66.9**	66.4	65.9	**66.4**
Finnish-FTB	**71.3**	70.4	72.1	**72.5**
French	**82.1**	81.6	81.3	**81.8**
German	**76.6**	76.0	75.4	**76.1**
Greek	75.2	**75.3**	75.1	**75.2**
Hebrew	**78.4**	77.9	77.9	**78.5**
Hindi	**85.4**	84.2	84.9	**85.2**
Italian	**83.8**	83.6	83.3	**83.6**
Norwegian	**84.5**	82.0	81.7	**84.5**
OC_Slavonic	**68.8**	68.7	68.7	**68.9**
Persian	**81.1**	79.8	79.8	**81.1**
Polish	**79.4**	79.1	79.0	**79.3**
Portuguese	81.3	**81.5**	**81.6**	81.3
Romanian	**64.2**	62.5	64.0	**64.6**
Slovenian	**80.8**	79.7	79.8	**80.8**
Spanish	**81.5**	81.2	81.2	**81.4**
Swedish	**76.8**	75.7	75.6	**76.7**
Tamil	67.2	67.1	67.4	**67.5**

MS is better than UD for parsing — B > A ✗

MS is easier to learn than UD — C > A ✗

Symmetry in differences — A - B = C - D

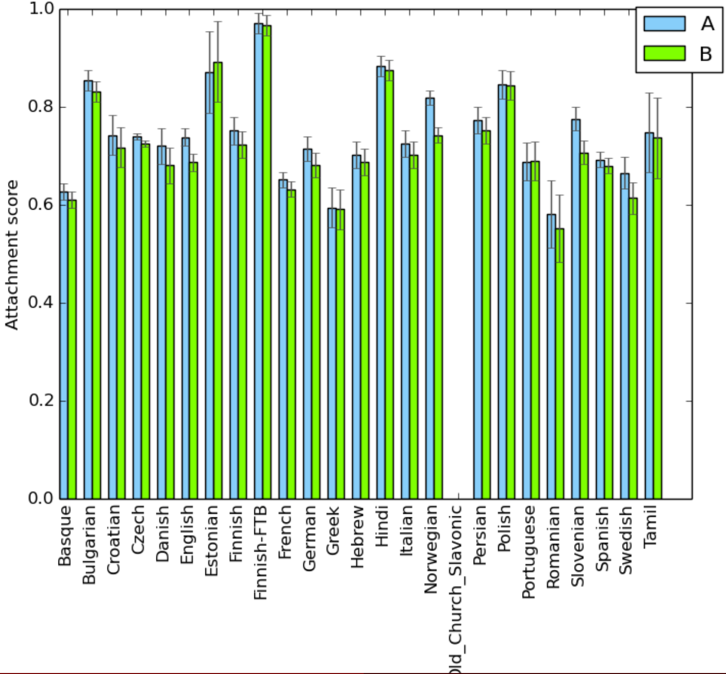
Results

UD language	A	B	C	D
Basque	**64.4**	63.8	64.0	**64.4**
Bulgarian	**83.4**	83.2	82.5	**82.9**
Croatian	**75.9**	74.6	73.7	**75.9**
Czech	**80.0**	76.5	76.4	**79.9**
Danish	**75.9**	75.2	74.8	**75.8**
English	**81.7**	80.4	80.2	**81.5**
Estonian	77.1	**77.8**	**77.6**	77.0
Finnish	**66.9**	66.4	65.9	**66.4**
Finnish-FTB	**71.3**	70.4	72.1	**72.5**
French	**82.1**	81.6	81.3	**81.8**
German	**76.6**	76.0	75.4	**76.1**
Greek	75.2	**75.3**	75.1	**75.2**
Hebrew	**78.4**	77.9	77.9	**78.5**
Hindi	**85.4**	84.2	84.9	**85.2**
Italian	**83.8**	83.6	83.3	**83.6**
Norwegian	**84.5**	82.0	81.7	**84.5**
OC_Slavonic	**68.8**	68.7	68.7	**68.9**
Persian	**81.1**	79.8	79.8	**81.1**
Polish	**79.4**	79.1	79.0	**79.3**
Portuguese	81.3	**81.5**	**81.6**	81.3
Romanian	**64.2**	62.5	64.0	**64.6**
Slovenian	**80.8**	79.7	79.8	**80.8**
Spanish	**81.5**	81.2	81.2	**81.4**
Swedish	**76.8**	75.7	75.6	**76.7**
Tamil	67.2	67.1	67.4	**67.5**

MS is better than UD for parsing B > A ✗

MS is easier to learn than UD C > A ✗

Symmetry in differences A - B = C - D ✗


1.0
0.8
0.6
0.4
0.2
0.0
Attachment score
A
B
Basque
Bulgarian
Croatian
Czech
Danish
English
Estonian
Finnish
Finnish-FTB
French
German
Greek
Hebrew
Hindi
Italian
Norwegian
Old_Church_Slavonic
Persian
Polish
Portuguese
Romanian
Slovenian
Spanish
Swedish
Tamil

Comparison with SDT and PDT

	Orig	**Transf**
UD_Czech	**80.0**	76.5**
PDT	68.5	**68.8****
UD_Slovenian	**80.5**	79.1**
SDT	65.7	**66.2**

Table : LAS with the original and transformed treebanks.

Ambiguity of POS tags in SDT

POS	main verb	aux
Verb-main	72.81	0.22
Verb-copula	22.30	95.95

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Were improvements in SDT/PDT the result of disambiguation?

Create different versions of the treebanks:

- Ambiguous: all verbal POS tags are the same (VERB)
- Disambiguated: main verb POS tag is Verb-main and auxiliary is AUX

Compare impact of VG transformation on:

- τ_o : original treebank
- τ_d : disambiguated treebank
- τ_a : ambiguous treebank

Results of the ambiguous experiment

	A	B	Δ
SDT τ_d	**67.8**	67.4	-0.4
SDT τ_o	65.7	**66.2**	0.5
SDT τ_a	64.2	**65.4***	1.2
PDT τ_d	69.2	69.2	0.0
PDT τ_o	68.5	**68.8****	0.3
PDT τ_a	68.2	**68.4***	0.2

Table : LAS on A and B with different levels of POS tag ambiguity.

Δ = B - A

Predicted vs gold POS tags

Can UD benefit from the transformation when using predicted POS tags?

✗ It seems not.

POS tag	A	B	Δ
gold	76.8	75.7**	-1.1
predicted	76.4	75.6**	-0.8

Table : LAS on UD_Swedish. Δ = B - A

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Conclusion and Future Work

Conclusions

- Keep verb groups as is in UD
- Benefits of error analysis
- Previous results were obtained through POS tag disambiguation

Future Work

- More in-depth error analysis
- Other representations (e.g. PPs)
- Other parsing models

Outline for section 3

Multiword Expressions and Syntactic Parsing

MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

Simple approach: representing MWEs as words-with-spaces

Multiword Expressions and Syntactic Parsing

MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

Simple approach: representing MWEs as words-with-spaces

Multiword Expressions and Syntactic Parsing

MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

Simple approach: representing MWEs as words-with-spaces

Multiword Expressions and Syntactic Parsing

MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

Simple approach: representing MWEs as words-with-spaces

Multiword Expressions and Syntactic Parsing

MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

Simple approach: representing MWEs as words-with-spaces

Multiword Expressions and Syntactic Parsing

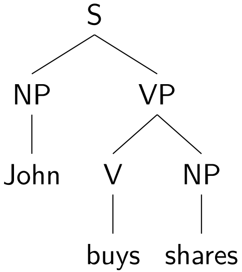
MWE: A group of multiple lexemes which have some level of idiomaticity or irregularity.

Motivation

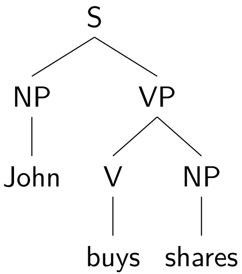
- Linguistics research: Construction Grammar (Hoffmann and Trousdale, 2013)
- MWE identification and syntactic parsing benefit from one another (Nivre and Nilsson, 2004; Korkontzelos and Manandhar, 2010)

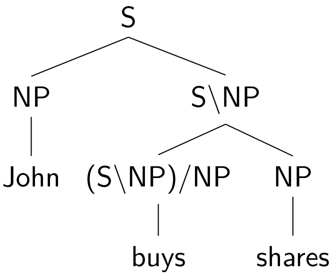
Simple approach: representing MWEs as words-with-spaces

CCG: a strongly lexicalized formalism



CCG: a strongly lexicalized formalism





Research Questions

- Can CCG parsing benefit from MWE information?

Research Questions

- Can CCG parsing benefit from MWE information?
- Can automatic MWE recognition improve parsing?

Research Questions

- Can CCG parsing benefit from MWE information?
- Can automatic MWE recognition improve parsing?
- Can we obtain improvements on different types of MWEs?

Research Questions

- Can CCG parsing benefit from MWE information?
- Can automatic MWE recognition improve parsing?
- Can we obtain improvements on different types of MWEs?

1 and 2

- Can we obtain a *training effect*?
- Can we obtain a *parsing effect*?

Research Questions

- Can CCG parsing benefit from MWE information?
- Can automatic MWE recognition improve parsing?
- Can we obtain improvements on different types of MWEs?

1 and 2

- Can we obtain a *training effect*?
- Can we obtain a *parsing effect*?

An example

It gives part of speech and lemma information.
It gives part+of+speech and lemma information.

MWE Recognition

Library jMWE (Finlayson and Kulkarni, 2011)

- Input: sentence.
  e.g: *Mr. Spoon said the plan is not an attempt to shore up a decline in ad pages in the first nine months of 1989 ; Newsweek 's ad pages totaled 1,620 , a drop of 3.2 % from last year , according to Publishers Information Bureau .*

- Output: list of multiword expressions from left to right.

  mr._spoon, shore_up, according_to, publishers_information_bureau

MWE Recognition (2)

3 components: detector, filter, resolver

MWE Recognition (2)

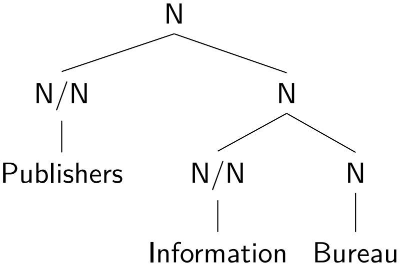
3 components: detector, filter, resolver

detector	**filter**	**resolver**
Exhaustive	MoreFrequentAsMWE	Longest
Exhaustive	MoreFrequentAsMWE	Leftmost
Proper Nouns	no filter	Longest
Exhaustive	ConstrainLength	Leftmost
Stop words	no filter	Longest

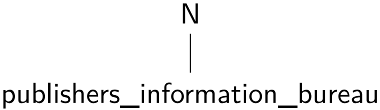
Transformation

MWEs with units that are siblings.

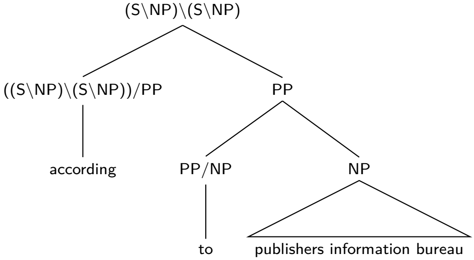
Input:



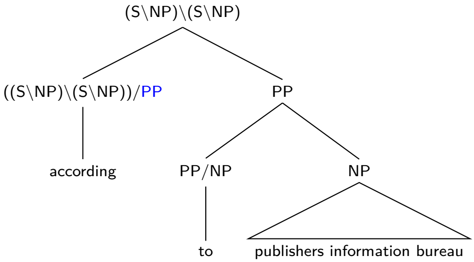
Output:



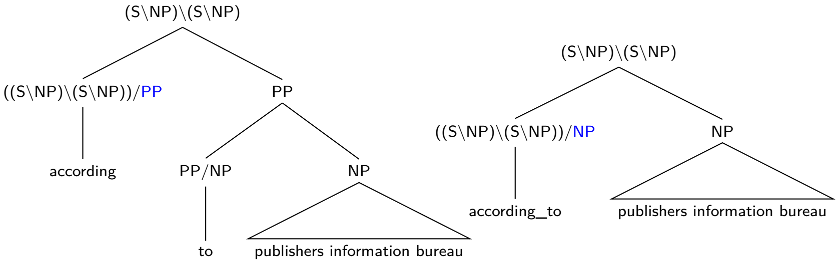
MWEs with units that are not sibling: less straightforward



MWEs with units that are not sibling: less straightforward

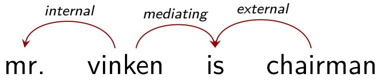


MWEs with units that are not sibling: less straightforward

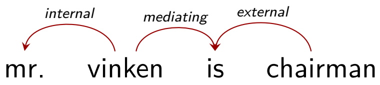


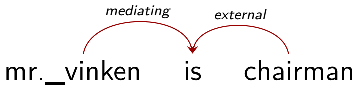
detector	**filter**	**resolver**	**MWE count**	**Sibling %**
Exhaustive	MoreFrequentAsMWE	Longest	53,208	79.51
Exhaustive	MoreFrequentAsMWE	Leftmost	51,543	41.85
Proper Nouns	no filter	Longest	32,583	86.14
Exhaustive	ConstrainLength	Leftmost	49,587	40.30
Stop words	no filter	Longest	13,623	2.09

Transformation (2)

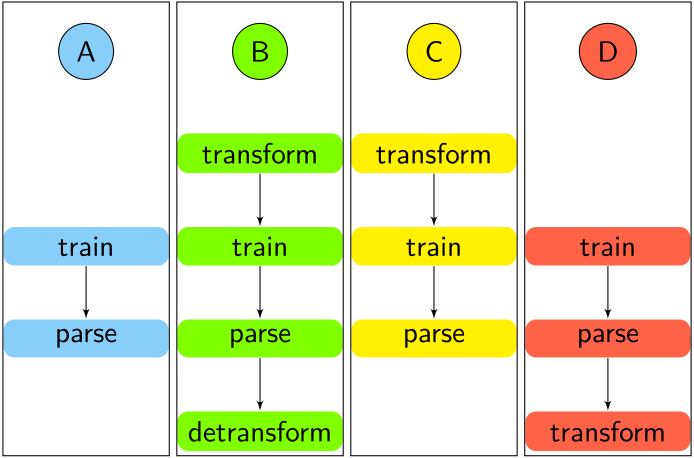


Transformation (2)



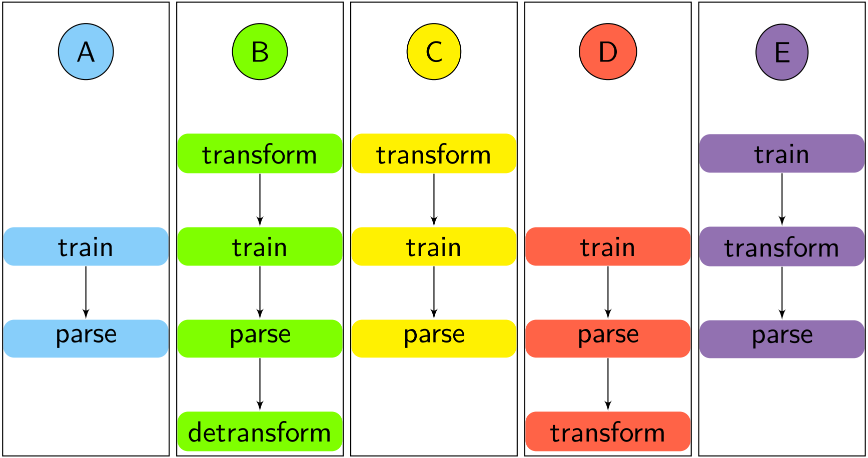


Pipeline



A
train
parse
B
transform
train
parse
detransform
C
transform
train
parse
D
train
parse
transform

Pipeline


A
train
parse
B
transform
train
parse
detransform
C
transform
train
parse
D
train
parse
transform
E
train
transform
parse

Research Questions

Research Questions

Is there a training effect?

Research Questions

Is there a training effect? C > E

Research Questions

Is there a training effect? C > E

Is there a parsing effect?

Research Questions

Is there a training effect? C > E

Is there a parsing effect? E > D

Research Questions

Is there a training effect? C > E

Is there a parsing effect? E > D

Can we improve results on the original gold standard?

Research Questions

Is there a training effect?	C > E
Is there a parsing effect?	E > D
Can we improve results on the original gold standard?	B > A

Research Questions

Is there a training effect? C > E

Is there a parsing effect? E > D

Can we improve results on the original gold standard? B > A

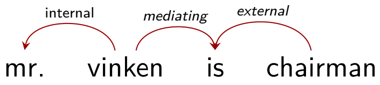
Can we obtain improvements on different types of MWEs?

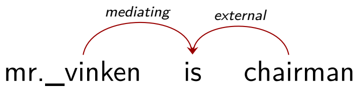
Research Questions

Is there a training effect?	C > E
Is there a parsing effect?	E > D
Can we improve results on the original gold standard?	B > A
Can we obtain improvements on different types of MWEs?	repeat and compare B

How to detransform?

How to detransform?

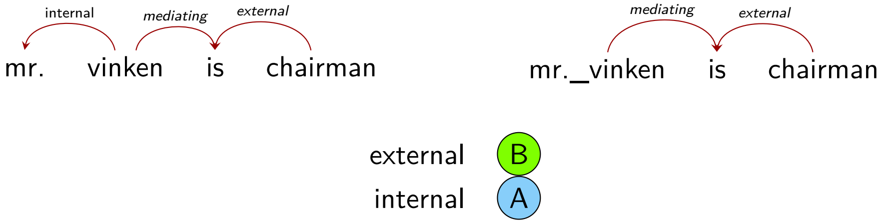




How to detransform?

external B

How to detransform?



How to detransform?

external B
internal A
mediating A

Results

Results

	E	C
only sibling MWEs	84.64	**84.88**
all MWEs	72.77	**72.92**

Results

	E	C
only sibling MWEs	84.64	**84.88**
all MWEs	72.77	**72.92**

Is there a training effect? C > E

Results

	E	C
only sibling MWEs	84.64	**84.88**
all MWEs	72.77	**72.92**

Is there a training effect? C > E ✓

Results

	D	E
only sibling MWEs	79.17	**84.06**
all MWEs	79.49	**79.69**

Is there a training effect? C > E ✓

Results

	D	E
only sibling MWEs	79.17	**84.06**
all MWEs	79.49	**79.69**

Is there a training effect? C > E ✓

Is there a parsing effect? E > D

Results

	D	E
only sibling MWEs	79.17	**84.06**
all MWEs	79.49	**79.69**

Is there a training effect? C > E ✓

Is there a parsing effect? E > D ✓

Results

A	B
85.15	**85.28**

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓

Results

A	B
85.15	**85.28**

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓
Can we improve results on the original gold standard?	B > A	

Results

A B

85.15 **85.28**

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓
Can we improve results on the original gold standard?	B > A	✓

Results

model	detector type	resolver type	F_1
A			85.15
B_1	exhaustive	longest	85.18
B_2	exhaustive	leftmost	85.02
B_3	Proper Nouns	longest	**85.28**
B_4	Length 2	leftmost	85.07
B_5	Stop words	longest	85.19

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓
Can we improve results on the original gold standard?	B > A	✓

Results

model	detector type	resolver type	F_1
A			85.15
B_1	exhaustive	longest	85.18
B_2	exhaustive	leftmost	85.02
B_3	Proper Nouns	longest	**85.28**
B_4	Length 2	leftmost	85.07
B_5	Stop words	longest	85.19

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓
Can we improve results on the original gold standard?	B > A	✓
Can we obtain improvements on different types of MWEs?	repeat and compare B	

Results

model	detector type	resolver type	F_1
A			85.15
B_1	exhaustive	longest	85.18
B_2	exhaustive	leftmost	85.02
B_3	Proper Nouns	longest	**85.28**
B_4	Length 2	leftmost	85.07
B_5	Stop words	longest	85.19

Is there a training effect?	C > E	✓
Is there a parsing effect?	E > D	✓
Can we improve results on the original gold standard?	B > A	✓
Can we obtain improvements on different types of MWEs?	repeat and compare B	✓

Conclusion

Contributions

- Improvements on CCG parsing with automatic MWE recognition
- Significant results despite limited settings
- Techniques for distinguishing training from parsing effects
- Empirical support that there is both training and parsing effects
- Differences in results when using different recognizers

Conclusion

Future Work

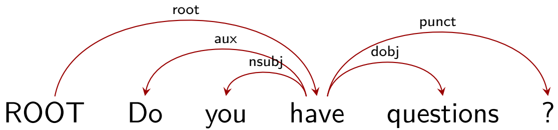
- Extending the transformation algorithm to the non-sibling case
- Testing more MWE recognition methods
- Conducting error analysis

Outline for section 4

General Conclusion

Principled methods for studying the interaction between syntactic representations and parsing accuracy.

Thanks!



References

Miryam de Lhoneux. 2014. *CCG Parsing and Multiword Expressions*. Master's thesis, The University of Edinburgh.

Miryam de Lhoneux and Joakim Nivre. 2016. Should have, would have, could have. investigating verb group representations for parsing with universal dependencies. In *Proceedings of the Workshop on Multilingual and Cross-lingual Methods in NLP*. Association for Computational Linguistics, San Diego, California, pages 10–19.

Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher D. Manning. 2014. Universal stanford dependencies: A cross-linguistic typology. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014.*. pages 4585–4592.

Mark Alan Finlayson and Nidhi Kulkarni. 2011. Detecting Multi-word Expressions Improves Word Sense Disambiguation. In *Proceedings of the Workshop on Multiword Expressions: From Parsing and Generation to the Real World*. Association for Computational Linguistics, Stroudsburg, PA, USA, MWE '11, pages 20–24.

Thomas Hoffmann and Graeme Trousdale. 2013. The Handbook of Construction Grammar. In Thomas Hoffmann and Graeme Trousdale, editors, *The Oxford Handbook of Construction Grammar*, Oxford University Press, Oxford Handbooks in Linguistics, chapter Constructi, pages 1–14.

Ioannis Korkontzelos and Suresh Manandhar. 2010. Can recognising multiword expressions improve shallow parsing? In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 636–644.

Jens Nilsson, Joakim Nivre, and Johan Hall. 2006. Graph transformations in data-driven dependency parsing. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Stroudsburg, PA, USA, ACL-44, pages 257–264.

Jens Nilsson, Joakim Nivre, and Johan Hall. 2007. Generalizing tree transformations for inductive dependency parsing. In *ACL 2007, Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, June 23-30, 2007, Prague, Czech Republic*.

Joakim Nivre, Johan Hall, and Jens Nilsson. 2006. MaltParser: A data-driven parser-generator for dependency parsing. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*. pages 2216–2219.

Joakim Nivre and Jens Nilsson. 2004. Multiword units in syntactic parsing. In *MEMURA 2004 – Methodologies and Evaluation of Multiword Units in Real-World Applications (LREC Workshop)*. pages 39–46.

Roy Schwartz, Omri Abend, and Ari Rappoport. 2012. Learnability-based syntactic annotation design. In *COLING*. volume 24, pages 2405–2422.