Perspectives on Universal Dependencies

Joakim Nivre
Uppsala University
Department of Linguistics and Philology

Based on collaborative work with Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher Manning, Ryan McDonald, Natalia Silveira, Slav Petrov, Sampo Pyysalo, Sebastian Schuster, Reut Tsarfaty, Francis Tyers, Daniel Zeman and many others
Introduction
Introduction

Growing interest in multilingual and cross-lingual NLP

- Multilingual evaluation campaigns to test generality of approaches
- Cross-lingual learning to support low-resource languages
Introduction

Growing interest in multilingual and cross-lingual NLP
  • Multilingual evaluation campaigns to test generality of approaches
  • Cross-lingual learning to support low-resource languages

Growing awareness of methodological problems
  • Current NLP relies heavily on linguistic annotation
  • Annotation guidelines vary across languages
A cat chases rats and mice.
A cat chases rats and mice

En katt jagar råttor och möss
A cat chases rats and mice.

En katt jagar råttor och möss.

En kat jager rotter og mus.
A cat chases rats and mice.
A cat chases rats and mice.

En katt jagar råttor och möss.

En kat jager rotter og mus.
Why is this a problem?
Why is this a problem?

- Hard to compare empirical results across languages
Why is this a problem?

- Hard to compare empirical results across languages
- Hard to usefully do cross-lingual structure transfer
Why is this a problem?

• Hard to compare empirical results across languages
• Hard to usefully do cross-lingual structure transfer
• Hard to evaluate cross-lingual learning
Why is this a problem?

• Hard to compare empirical results across languages
• Hard to usefully do cross-lingual structure transfer
• Hard to evaluate cross-lingual learning
• Hard to build and maintain multilingual systems
Why is this a problem?

• Hard to compare empirical results across languages
• Hard to usefully do cross-lingual structure transfer
• Hard to evaluate cross-lingual learning
• Hard to build and maintain multilingual systems
• Hard to make comparative linguistic studies
Why is this a problem?

- Hard to compare empirical results across languages
- Hard to usefully do cross-lingual structure transfer
- Hard to evaluate cross-lingual learning
- Hard to build and maintain multilingual systems
- Hard to make comparative linguistic studies
- Hard to validate linguistic typology
Why is this a problem?

- Hard to compare empirical results across languages
- Hard to usefully do cross-lingual structure transfer
- Hard to evaluate cross-lingual learning
- Hard to build and maintain multilingual systems
- Hard to make comparative linguistic studies
- Hard to validate linguistic typology
- Hard to make progress towards a universal parser
Toutefois, les filles adorent les desserts.

toutefois, le fille adorer les dessert.
Part-of-speech tags

Toutefois, les filles adorent les desserts.

toutefois, le fille adorer les dessert.
Universal Dependencies

http://universaldependencies.org

Toutefois, les filles adorent les desserts.

toutefois, le fille adorer les dessert.

ADV PUNCT DET NOUN VERB DET NOUN PUNCT

Definite=Def Gender=Fem Number=Plur Number=Plur Person=3 Tense=Pres

Part-of-speech tags

Morphological features
Universal Dependencies
http://universaldependencies.org

Brief history of UD:

- Kick-off meeting at EACL in Gothenburg, April 2014
- Guidelines v1, October 2014
- Treebank releases every 6 months (v1.0–v1.4)
- Guidelines v2, December 2016
- Treebank release v2.0, March 2017

Open community effort – anyone can contribute!
Universal Dependencies

http://universaldependencies.org

Brief history of UD:

• Kick-off meeting at EACL in Gothenburg, April 2014
• Guidelines v1, October 2014
• Treebank releases every 6 months (v1.0–v1.4)
• Guidelines v2, December 2016
• Treebank release v2.0, March 2017

Open community effort – anyone can contribute!
June, 2017:

- 56 languages
- 95 treebanks
- 13.9 million words
- 202 contributors
- 10791 downloads
June, 2017:
- 56 languages
- 95 treebanks
- 13.9 million words
- 202 contributors
- 10791 downloads
The UD Philosophy
The UD Philosophy

Maximize parallelism – but don’t overdo it

- Don’t annotate the same thing in different ways
- Don’t make different things look the same
- Don’t annotate things that are not there
The UD Philosophy

Maximize parallelism – but don’t overdo it
  • Don’t annotate the same thing in different ways
  • Don’t make different things look the same
  • Don’t annotate things that are not there

Universal taxonomy with language-specific elaboration
  • Languages select from a universal pool of categories
  • Allow language-specific extensions
Design Principles
Design Principles

Dependency

• Widely used in practical NLP systems
• Available in treebanks for many languages
Design Principles

Dependency
- Widely used in practical NLP systems
- Available in treebanks for many languages

Lexicalism
- Basic annotation units are words – syntactic words
- Words have morphological properties
- Words enter into syntactic relations
Morphology

Le chat chasse les chiens.
Morphology

• Lemma representing the semantic content of the word
Morphology

- Lemma representing the semantic content of the word
- Part-of-speech tag representing its grammatical class
Morphology

- Lemma representing the semantic content of the word
- Part-of-speech tag representing its grammatical class

<table>
<thead>
<tr>
<th>Open</th>
<th>Closed</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADJ</td>
<td>ADP</td>
<td>PUNCT</td>
</tr>
<tr>
<td>ADV</td>
<td>AUX</td>
<td>SYM</td>
</tr>
<tr>
<td>INTJ</td>
<td>CCONJ</td>
<td>X</td>
</tr>
<tr>
<td>NOUN</td>
<td>DET</td>
<td></td>
</tr>
<tr>
<td>PROPN</td>
<td>NUM</td>
<td></td>
</tr>
<tr>
<td>VERB</td>
<td>PART</td>
<td></td>
</tr>
<tr>
<td>PRON</td>
<td>SCONJ</td>
<td></td>
</tr>
</tbody>
</table>
Morphology

- Lemma representing the semantic content of the word
- Part-of-speech tag representing its grammatical class
- Features representing lexical and grammatical properties of the lemma or the particular word form
Morphology

- **Lexical**: Lemma representing the semantic content of the word.
- **Part-of-speech tag**: Representing its grammatical class.
- **Features**: Representing lexical and grammatical properties of the lemma or the particular word form.

<table>
<thead>
<tr>
<th>PronType</th>
<th>Gender</th>
<th>VerbForm</th>
</tr>
</thead>
<tbody>
<tr>
<td>NumType</td>
<td>Animacy</td>
<td>Mood</td>
</tr>
<tr>
<td>Poss</td>
<td>Number</td>
<td>Tense</td>
</tr>
<tr>
<td>Reflex</td>
<td>Case</td>
<td>Aspect</td>
</tr>
<tr>
<td>Foreign</td>
<td>Definite</td>
<td>Voice</td>
</tr>
<tr>
<td>Abbr</td>
<td>Degree</td>
<td>Evident</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polarity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Person</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polite</td>
</tr>
</tbody>
</table>
The cat could have chased all the dogs down the street.
The cat could have chased all the dogs down the street.

- Content words are related by dependency relations.
The cat could have chased all the dogs down the street.

- Content words are related by dependency relations
- Function words attach to the content word they modify
Content words are related by dependency relations

Function words attach to the content word they modify

Punctuation attach to head of phrase or clause
Syntactic Relations
Syntactic Relations

Taxonomy of 37 universal syntactic relations

• Three types of structures: nominals, clauses, modifiers
• Core arguments vs. other dependents (not complements vs. adjuncts)
Syntactic Relations

Taxonomy of 37 universal syntactic relations

- Three types of structures: nominals, clauses, modifiers
- Core arguments vs. other dependents (not complements vs. adjuncts)

A two-level architecture

- Broad universal categories to allow cross-linguistic comparison
- Subtypes to capture language-specific phenomena
## Syntactic Relations

<table>
<thead>
<tr>
<th>Core Predicate Dep</th>
<th>Nominal</th>
<th>Clause</th>
<th>Modifier Word</th>
<th>Function Word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nsubj</td>
<td>csubj</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>obj</td>
<td>ccomp</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iobj</td>
<td>xcomp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-Core Predicate Dep</th>
<th>Nominal</th>
<th>Clause</th>
<th>Modifier Word</th>
<th>Function Word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obl</td>
<td>advcl</td>
<td>advmod*</td>
<td>aux</td>
</tr>
<tr>
<td></td>
<td>vocative</td>
<td></td>
<td>discourse</td>
<td>cop</td>
</tr>
<tr>
<td></td>
<td>expl</td>
<td></td>
<td></td>
<td>mark</td>
</tr>
<tr>
<td></td>
<td>dislocated</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nominal Dep</th>
<th>Nominal</th>
<th>Clause</th>
<th>Modifier Word</th>
<th>Function Word</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nmod</td>
<td>acl</td>
<td>amod</td>
<td>det</td>
</tr>
<tr>
<td></td>
<td>appos</td>
<td></td>
<td></td>
<td>clf</td>
</tr>
<tr>
<td></td>
<td>nummod</td>
<td></td>
<td></td>
<td>case</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coordination</th>
<th>MWE</th>
<th>Loose</th>
<th>Special</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>conj cc</td>
<td>fixed flat compound</td>
<td>parataxis list</td>
<td>orphan goes with reparandum</td>
<td>punct root dep</td>
</tr>
</tbody>
</table>
The Primacy of Content Words

The cat could have chased all the dogs down the street.
The Primacy of Content Words

Dubious Linguistics?

“Such an approach to the syntax of natural languages is contrary to most work in theoretical syntax in the past 35 years, regardless of whether this work is constituency- or dependency-based.” (Groß and Osborne, 2015)
The Primacy of Content Words

“It is now fairly well known that, while dependency representations in which content words are made heads tend to help semantically oriented downstream applications, dependency parsing numbers are higher if you make auxiliary verbs heads […] and if you make prepositions the head of prepositional phrases.” (De Marneffe et al., 2014)
Manning’s Law

The secret to understanding the design of UD is to realize that it is a very subtle compromise between approximately 6 things:

1. UD needs to be satisfactory on linguistic analysis grounds for individual languages.
2. UD needs to be good for linguistic typology, i.e., providing a suitable basis for bringing out cross-linguistic parallelism across languages and language families.
3. UD must be suitable for rapid, consistent annotation by a human annotator.
4. UD must be suitable for computer parsing with high accuracy.
5. UD must be easily comprehended and used by a non-linguist, whether a language learner or an engineer with prosaic needs for language processing.
6. UD must support well downstream language understanding tasks (relation extraction, reading comprehension, machine translation, …).

It’s easy to come up with a proposal that improves UD on one of these dimensions. The interesting and difficult part is to improve UD while remaining sensitive to all these dimensions.
Manning’s Law

The secret to understanding the design of UD is to realize that it is a very subtle compromise between approximately 6 things:

1. UD needs to be satisfactory on linguistic analysis grounds for individual languages.
2. UD needs to be good for linguistic typology, i.e., providing a suitable basis for bringing out cross-linguistic parallelism across languages and language families.
3. UD must be suitable for rapid, consistent annotation by a human annotator.
4. UD must be suitable for computer parsing with high accuracy.
5. UD must be easily comprehended and used by a non-linguist, whether a language learner or an engineer with prosaic needs for language processing.
6. UD must support well downstream language understanding tasks (relation extraction, reading comprehension, machine translation, …).

It’s easy to come up with a proposal that improves UD on one of these dimensions. The interesting and difficult part is to improve UD while remaining sensitive to all these dimensions.
Manning’s Law

The secret to understanding the design of UD is to realize that it is a very subtle compromise between approximately 6 things:

1. UD needs to be satisfactory on linguistic analysis grounds for individual languages.
2. UD needs to be good for linguistic typology, i.e., providing a suitable basis for bringing out cross-linguistic parallelism across languages and language families.
3. UD must be suitable for rapid, consistent annotation by a human annotator.
4. UD must be suitable for computer parsing with high accuracy.
5. UD must be easily comprehended and used by a non-linguist, whether a language learner or an engineer with prosaic needs for language processing.
6. UD must support well downstream language understanding tasks (relation extraction, reading comprehension, machine translation, …).

It’s easy to come up with a proposal that improves UD on one of these dimensions. The interesting and difficult part is to improve UD while remaining sensitive to all these dimensions.
Manning’s Law

The secret to understanding the design of UD is to realize that it is a very subtle compromise between approximately 6 things:

1. UD needs to be satisfactory on linguistic analysis grounds for individual languages.
2. UD needs to be good for linguistic typology, i.e., providing a suitable basis for bringing out cross-linguistic parallelism across languages and language families.
3. UD must be suitable for rapid, consistent annotation by a human annotator.
4. UD must be suitable for computer parsing with high accuracy.
5. UD must be easily comprehended and used by a non-linguist, whether a language learner or an engineer with prosaic needs for language processing.
6. UD must support well downstream language understanding tasks (relation extraction, reading comprehension, machine translation, …).

It’s easy to come up with a proposal that improves UD on one of these dimensions. The interesting and difficult part is to improve UD while remaining sensitive to all these dimensions.
Linguistic Typology

With contributions by William Croft
Similarities and Differences

The dog was chased by the cat.

Hunden jagades av katten.

Pes byl honěn kočkou.
Similarities and Differences

The dog was chased by the cat.

Hunden jagades av katten.

Pes byl honěn kočkou.
Similarities and Differences

The dog was chased by the cat

Hunden jagades av katten

Pes byl honěn kočkou

Definite=Def Voice=Pass Definite=Def

Voice=Pass
Similarities and Differences

The dog was chased by the cat.

Hunden jagades av katten.

Pes byl honěn kočkou.
What is universal?
What is universal?

- Descriptive categories – language-specific, cannot be equated across languages
- Comparative concepts – created by typologists for cross-linguistic comparison

Haspelmath (2010) Comparative Concepts and Descriptive Categories in Crosslinguistic Studies
What is universal?

- Descriptive categories – language-specific, cannot be equated across languages
- Comparative concepts – created by typologists for cross-linguistic comparison

Haspelmath (2010) Comparative Concepts and Descriptive Categories in Crosslinguistic Studies

- Construction – whatever structure is used to express a function, universal by definition
- Strategy – a specific, cross-linguistically definable structure used to express a function

Croft et al. (2017) Linguistic Typology Meets Universal Dependencies
Constructions vs. strategies

**Construction**

- predication of object concept

**English:**
Ivan is the best dancer.

**Russian:**
Ivan lučšij tancor

**Strategies**

- inflected copula
  - zero copula/zero inflection
Croft et al. (2017) Linguistic Typology Meets Universal Dependencies
• A universal annotation scheme should have a classification of constructions, not strategies, as its universal foundational layer.
• A universal annotation scheme should have a classification of constructions, not strategies, as its universal foundational layer

• However, common recurrent strategies can and should be included as well
• A universal annotation scheme should have a classification of constructions, not strategies, as its universal foundational layer

• However, common recurrent strategies can and should be included as well

• The primacy of content words in UD is good, because function words such as copulas are strategies and not found in every language
• A universal annotation scheme should have a classification of constructions, not strategies, as its universal foundational layer

• However, common recurrent strategies can and should be included as well

• The primacy of content words in UD is good, because function words such as copulas are strategies and not found in every language

• This means that the topology of dependency trees in UD are basically fine from a typological-universal point of view
we have come to Osaka

dependency | nucleus
we have come to Osaka

dependency

karaka

nucleus

vibhakti
we have come to Osaka
Conclusion – Part I

The primacy of content words

• Perhaps incompatible with some approaches to theoretical syntax
• Compatible with current approaches to linguistic typology
• Consistent with older traditions of dependency grammar
Natural Language Understanding

With contributions (and slides) by Siva Reddy et al.
From Syntax to Semantics

- Direct relations from predicates to arguments/modifiers
- Cross-linguistically consistent patterns
- Universal syntax-semantics interface?
UDepLambda
UDepLambda

Goals

• From dependencies to logical forms
• Compositional, language-agnostic conversion
• Dependency tree dictates the semantics
UDepLambda

Goals
• From dependencies to logical forms
• Compositional, language-agnostic conversion
• Dependency tree dictates the semantics

Compositionality
The semantics of a complex expression is determined by the semantics of its constituent expressions and the rules used to combine them
• Complex expressions are dependency trees
• Constituent expressions are subtrees
• Rules are dependency labels
Why Dependencies?

- Easy to annotate (with UD annotation scheme)
- Treebanks exist in many languages
- Robust, efficient, accurate parsers

Dependency Tree to Semantics

CCG

\[
\begin{align*}
\text{Disney} & \quad \text{acquired} & \quad \text{Pixar} \\
NP & \quad S \setminus NP/ NP & \quad NP \\
\lambda y \lambda x \lambda e. \ \text{acquired}(e) & \quad \land \ \text{arg}_1(e, x) & \quad \land \ \text{arg}_2(e, y) \\
& \quad (S \setminus NP) & \quad (S \setminus NP) \\
\lambda x \lambda e. \ \text{acquired}(e) & \quad \land \ \text{arg}_1(e, x) & \quad \land \ \text{arg}_2(e, \text{Pixar}) \\
& \quad (S) & \quad (S) \\
\lambda e. \ \text{acquired}(e) & \quad \land \ \text{arg}_1(e, \text{Disney}) & \quad \land \ \text{arg}_2(e, \text{Pixar})
\end{align*}
\]

Typing and Combinator Rules allow a Synchronous Syntax-Semantics interface.

Dependencies lack a formal theory of semantics.
Disney acquired Pixar

Dependency labels drive the composition
Disney \text{nsubj} \text{acquired} \text{root} \text{dobj} Pixar

\ldots \text{dobj} \ldots \text{nsubj} \ldots

\text{(dobj acquired Pixar)}
Disney acquired Pixar

(nsubj (dobj acquired Pixar) Disney)
Single Type System

Disney \textit{acquired} Pixar

Lambda Expression for words

acquired \(\Rightarrow \lambda x_a x_e. \text{acquired}(x_e) \Rightarrow \text{TYPE} = \text{Ind} \times \text{Event} \rightarrow \text{Bool}\)

Pixar \(\Rightarrow \lambda x_a x_e. \text{Pixar}(x_a) \Rightarrow \text{TYPE} = \text{Ind} \times \text{Event} \rightarrow \text{Bool}\)

All constituents are of the same lambda expression type

\text{TYPE}[\text{acquired}] = \text{TYPE}[\text{Pixar}] = \text{TYPE}[(\text{dobj acquired Pixar})]
Lambda Expression for dependency labels

\[ \text{dobj} \Rightarrow \lambda f \lambda g \lambda z . \exists x . f(z) \land g(x) \land \text{arg}_2(z_e, x_a) \]
Disney acquired Pixar

\[ \lambda z. \exists x y. \text{acquired}(z_e) \land \text{Pixar}(y_a) \land \text{Disney}(x_a) \land \text{arg}_1(z_e, x_a) \land \text{arg}_2(z_e, y_a) \]
## Comparison with CCG

<table>
<thead>
<tr>
<th><strong>CCG</strong></th>
<th><strong>UDepLambda</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexicalized semantics</td>
<td>Simple lexical semantics</td>
</tr>
<tr>
<td>Words drive composition</td>
<td>Dependencies drive composition</td>
</tr>
<tr>
<td>Argument and adjunct distinction</td>
<td>Every dependent is an adjunct</td>
</tr>
<tr>
<td>Complex types are powerful</td>
<td>Simplicity gives robustness</td>
</tr>
</tbody>
</table>
Freebase Semantic Parsing

![English WebQuestions](image)

- **Yao15**: 44.3
- **BAST15**: 49.4
- **BERANT15**: 49.7
- **SIMPLEGRAPH**: 48.4
- **CCGGRAPH**: 48.6
- **DEPTREE**: 45.5
- **UDELPAMBDA**: 49.5
Freebase Semantic Parsing

Results: English WebQuestions

Yao15
Bast15
Berant15
SimpleGraph
CCGGraph
DepTree
UDepLambda
Yih15
Xu16

Average F1

44.3
49.4
49.7
48.8
48.6
49.5

Results: German WebQuestions

SimpleGraph
DepTree
UDepLambda

45.6
45.9
46.1

Results: Spanish WebQuestions

SimpleGraph
DepTree
UDepLambda

46.3
46.4
47.5
Conclusion – Part II

The primacy of content words

• Supports dependency-driven compositional semantics
• Generalizes across languages – universal semantic parsing?
• Useful for downstream natural language understanding tasks
Syntactic Parsing
Conventional wisdom

- Dependency trees with “function heads” are easier to parse
- Possibly related to dependency length minimisation
- Based on scattered empirical (and anecdotal) evidence
the cat for her has put is nice to go and that
the cat for her has put is red to put mark and him cc
<table>
<thead>
<tr>
<th>Paper</th>
<th>Language</th>
<th>det</th>
<th>case</th>
<th>aux</th>
<th>cop</th>
<th>mark</th>
<th>cc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td>the</td>
<td>for</td>
<td>has</td>
<td>is</td>
<td>nice</td>
<td>to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>has</td>
<td>her</td>
<td>put</td>
<td>nice</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>Language</td>
<td>det</td>
<td>case</td>
<td>aux</td>
<td>cop</td>
<td>mark</td>
<td>cc</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td><img src="green" alt="Det" /></td>
<td><img src="red" alt="Case" /></td>
<td><img src="yellow" alt="Aux" /></td>
<td><img src="pink" alt="Cop" /></td>
<td><img src="gray" alt="Mark" /></td>
<td><img src="green" alt="CC" /></td>
</tr>
<tr>
<td>Silveira and Manning (2015)</td>
<td>English</td>
<td><img src="gray" alt="Det" /></td>
<td><img src="gray" alt="Case" /></td>
<td><img src="gray" alt="Aux" /></td>
<td><img src="gray" alt="Cop" /></td>
<td><img src="gray" alt="Mark" /></td>
<td><img src="green" alt="CC" /></td>
</tr>
<tr>
<td>Paper</td>
<td>Language</td>
<td>det</td>
<td>case</td>
<td>aux</td>
<td>cop</td>
<td>mark</td>
<td>cc</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silveira and Manning (2015)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attardi et al. (2015)</td>
<td>Italian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>Language</td>
<td>det</td>
<td>case</td>
<td>aux</td>
<td>cop</td>
<td>mark</td>
<td>cc</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silveira and Manning (2015)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attardi et al. (2015)</td>
<td>Italian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosa (2015)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>Language</td>
<td>det</td>
<td>case</td>
<td>aux</td>
<td>cop</td>
<td>mark</td>
<td>cc</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silveira and Manning (2015)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attardi et al. (2015)</td>
<td>Italian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosa (2015)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Lhoneux and Nivre (2016)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>Language</td>
<td>det</td>
<td>case</td>
<td>aux</td>
<td>cop</td>
<td>mark</td>
<td>cc</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>----------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Schwartz et al. (2012)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silveira and Manning (2015)</td>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attardi et al. (2015)</td>
<td>Italian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosa (2015)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Lhoneux and Nivre (2016)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wisniewski and Lacroix (2017)</td>
<td>Multi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion – Part III

The primacy of content words

- Not universally harmful for monolingual syntactic parsing
- Arguably superior for cross-lingual parsing
- Neural dependency parsers less sensitive to representations?
Universal Dependencies

Designed to fulfill multiple purposes/requirements

- Meaningful linguistic analysis within and across languages
- Syntactic parsing in monolingual and cross-lingual settings
- Useful information for downstream language understanding tasks

Gives priority to dependencies between content words
Thanks to all UD contributors!