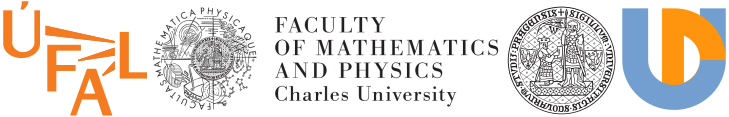
Universal Dependencies:

A Search for Harmonized Morphological and Syntactic Annotation

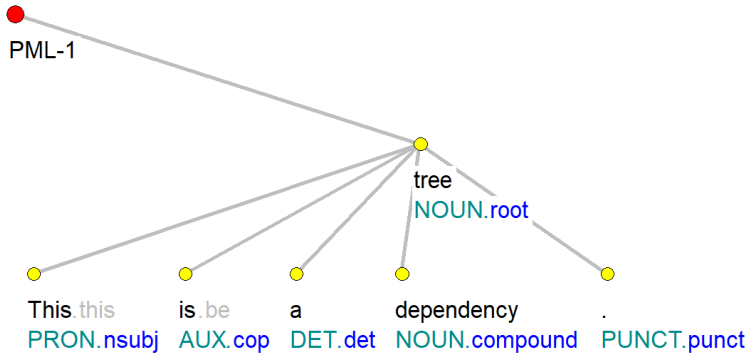


Daniel Zeman

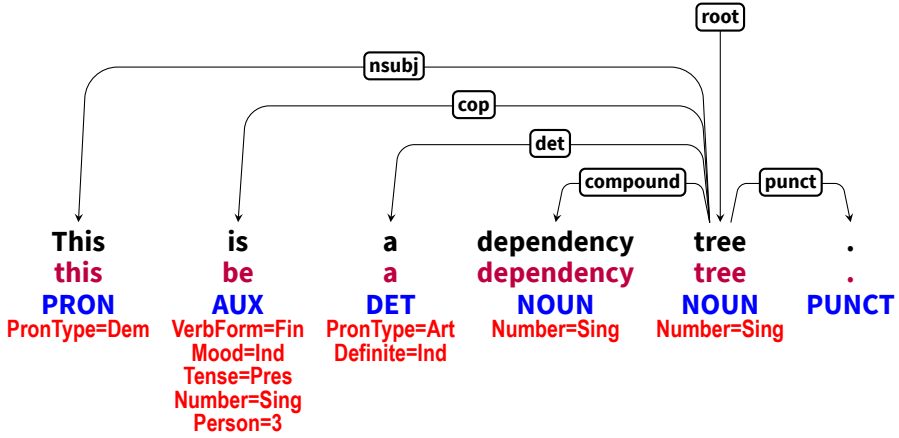
zeman@ufal.mff.cuni.cz
http://universaldependencies.org/

Universal Dependencies

Syntactic Trees



Syntactic Trees



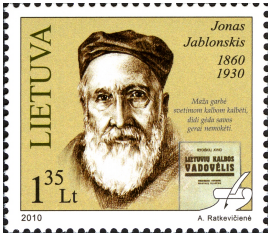
Users (Simplified)

LT engineers

Users (Simplified)

LT engineers

& linguists



Outline

1. Introduction to Universal Dependencies
2. Tricky part: core vs. oblique dependents
3. Closer to meaning: Enhanced and Deep UD

Basic Universal Dependencies: 92 (91) Languages and Growing

- I.-E.: Armenian, Ancient Greek, Greek, Albanian, Breton, Irish, Scottish, Welsh, Afrikaans, Danish, Dutch, English, Faroese, German, Gothic, Icelandic, Norwegian, Swedish, Swiss German, Catalan, French, Galician, Italian, Latin, Old French, Portuguese, Romanian, Spanish, Belarusian, Bulgarian, Church Slavonic, Croatian, Czech, Old Russian, Polish, Russian, Serbian, Slovak, Slovenian, Ukrainian, Upper Sorbian, Latvian, Lithuanian, Kurmanji, Persian, Hindi, Bhojpuri, Marathi, Sanskrit, Urdu
- Uralic: Erzya, Estonian, Finnish, Hungarian, Karelian, Livvi, Komi Permyak+Zyrian, Moksha, Sámi North+Skolt
- Dravidian: Tamil, Telugu; Turkic: Kazakh, Turkish, Uyghur
- Af.-As.: Akkadian, Amharic, Arabic, Assyrian, Coptic, Hebrew, Maltese
- Sino-Tibetan: Cantonese, Classical Chinese, Chinese; Aus.-As.: Vietnamese
- Tai-Kadai: Thai; Austronesian: Indonesian, Tagalog
- Other: Buryat, Japanese, Korean, Basque, Sw. Sign, Naija, Bambara, Wolof, Yoruba, Warlpiri, Mbyá Guaraní

370 Contributors

Daniel Zeman, Joakim Nivre, Mitchell Abrams, Elia Ackermann, Noëmi Aepli, Željko Agić, Lars Ahrenberg, Chika Kennedy Ajede, Gabrielė Aleksandravičiūtė, Lene Antonsen, Katya Aplonova, Angelina Aquino, Maria Jesus Aranzabe, Gashaw Arutie, Masayuki Asahara, Luma Ateyah, Furkan Atmaca, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Elena Badmaeva, Miguel Ballesteros, Esha Banerjee, Sebastian Bank, Verginica Barbu Mititelu, Victoria Basmov, Colin Batchelor, John Bauer, Kepa Bengoetxea, Yevgeni Berzak, Irshad Ahmad Bhat, Riyaz Ahmad Bhat, Erica Biagetti, Eckhard Bick, Agnė Bielinskienė, Rogier Blokland, Victoria Bobicev, Loïc Boizou, Emanuel Borges Völker, Carl Börstell, Cristina Bosco, Gosse Bouma, Sam Bowman, Adriane Boyd, Kristina Brokaitė, Aljoscha Burchardt, Marie Candito, Bernard Caron, Gauthier Caron, Tatiana Cavalcanti, Gülşen Cebiroğlu Eryiğit, Flavio Massimiliano Cecchini, Giuseppe G. A. Celano, Slavomír Čéplö, Savas Cetin, Fabricio Chalub, Ethan Chi, Jinho Choi, Yongseok Cho, Jayeol Chun, Alessandra T. Cignarella, Silvie Cinková, Aurélie Collomb, Çağrı Çöltekin, Miriam Connor, Marine Courtin, Elizabeth Davidson, Marie-Catherine de Marneffe, Valeria de Paiva, Elvis de Souza, Arantza Diaz de Ilarraza, Carly Dickerson, Bamba Dione, Peter Dirix, Kaja Dobrovoljc, Timothy Dozat, Kira Droganova, Puneet Dwivedi, Hanne Eckhoff, Marhaba Eli, Ali Elkahky, Binyam Ephrem, Olga Erina, Tomaž Erjavec, Aline Etienne, Wograine Evelyn, Richárd Farkas, Hector Fernandez Alcalde, Jennifer Foster, Cláudia Freitas, Kazunori Fujita, Katarína Gajdošová, Daniel Galbraith, Marcos Garcia, Moa Gärdenfors, Sebastian Garza, Kim Gerdes, Filip Ginter, Iakes Goenaga, Koldo Gojenola, Memduh Gökırmak, Yoav Goldberg, Xavier Gómez Guinovart, Berta González Saavedra, Bernadeta Griciūtė, Matias Grioni, Loïc Grobol, Normunds Grūzītis, Bruno Guillaume, Céline Guillot-Barbance, Tunga Güngör, Nizar Habash, Jan Hajič, Jan Hajič jr., Mika Hämäläinen, Linh Hà Mỹ, Na-Rae Han, Kim Harris, Dag Haug, Johannes Heinecke, Oliver Hellwig, Felix Hennig, Barbora Hladká, Jaroslava Hlaváčová, Florinel Hociung, Petter Hohle, Jena Hwang, Takumi Ikeda, Radu Ion, Elena Irimia, Ọlájídé Ishola, Tomáš Jelínek, Anders Johannsen, Hildur Jónsdóttir, Fredrik Jørgensen, Markus Juutinen, Hüner Kaşıkara, Andre Kaasen, Nadezhda Kabaeva, Sylvain Kahane, Hiroshi Kanayama, Jenna Kanerva, Boris Katz, Tolga Kayadelen, Jessica Kenney, Václava Kettnerová, Jesse Kirchner, Elena Klementieva, Arne Köhn, Abdullatif Köksal, Kamil Kopacewicz, Timo Korkiakangas, Natalia Kotsyba, Jolanta Kovalevskaitė, Simon Krek, Sookyoung Kwak, Veronika Laippala, Lorenzo Lambertino, Lucia Lam, Tatiana Lando, Septina Dian Larasati, Alexei Lavrentiev, John Lee, Phương Lê Hồng, Alessandro Lenci, Saran Lertpradit, Herman Leung, Maria Levina, Cheuk Ying Li, Josie Li, Keying Li, KyungTae Lim, Yuan Li, Nikola Ljubešić, Olga Loginova, Olga Lyashevskaya, Teresa Lynn, Vivien Macketanz, Aibek Makazhanov, Michael Mandl, Christopher Manning, Ruli Manurung, Cătălina Mărănduc, David Mareček, Katrin Marheinecke, Héctor Martínez Alonso, André Martins, Jan Mašek, Hiroshi Matsuda, Yuji Matsumoto, Ryan McDonald, Sarah McGuinness, Gustavo Mendonça, Niko Miekka, Margarita Misirpashayeva, Anna Missilä, Cătălin Mititelu, Maria Mitrofan, Yusuke Miyao, Simonetta Montemagni, Amir More, Laura Moreno Romero, Keiko Sophie Mori, Tomohiko Morioka, Shinsuke Mori, Shigeki Moro, Bjartur Mortensen, Bohdan Moskalevskyi, Kadri Muischnek, Robert Munro, Yugo Murawaki, Kaili Müürisep, Pinkey Nainwani, Juan Ignacio Navarro Horñiacek, Anna Nedoluzhko, Gunta Nešpore-Bērzkalne, Lương Nguyễn Thị, Huyền Nguyễn Thị Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Hanna Nurmi, Stina Ojala, Atul Kr. Ojha, Adédayọ̀ Olúòkun, Mai Omura, Emeka Onwuegbuzia, Petya Osenova, Robert Östling, Lilja Øvrelid, Şaziye Betül Özateş, Arzucan Özgür, Balkız Öztürk Başaran, Niko Partanen, Elena Pascual, Marco Passarotti, Agnieszka Patejuk, Guilherme Paulino-Passos, Angelika Peljak-Łapińska, Siyao Peng, Cenel-Augusto Perez, Guy Perrier, Daria Petrova, Slav Petrov, Jason Phelan, Jussi Piitulainen, Tommi A Pirinen, Emily Pitler, Barbara Plank, Thierry Poibeau, Larisa Ponomareva, Martin Popel, Lauma Pretkalniņa, Sophie Prévost, Prokopis Prokopidis, Adam Przepiórkowski, Tiina Puolakainen, Sampo Pyysalo, Peng Qi, Andriela Rääbis, Alexandre Rademaker, Loganathan Ramasamy, Taraka Rama, Carlos Ramisch, Vinit Ravishankar, Livy Real, Petru Rebeja, Siva Reddy, Georg Rehm, Ivan Riabov, Michael Rießler, Erika Rimkutė, Larissa Rinaldi, Laura Rituma, Luisa Rocha, Mykhailo Romanenko, Rudolf Rosa, Valentin Roșca, Davide Rovati, Olga Rudina, Jack Rueter, Shoval Sadde, Benoît Sagot, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Stephanie Samson, Manuela Sanguinetti, Dage Särg, Baiba Saulīte, Yanin Sawanakunanon, Salvatore Scarlata, Nathan Schneider, Sebastian Schuster, Djamé Seddah, Wolfgang Seeker, Mojgan Seraji, Mo Shen, Atsuko Shimada, Hiroyuki Shirasu, Muh Shohibussirri, Dmitry Sichinava, Aline Silveira, Natalia Silveira, Maria Simi, Radu Simionescu, Katalin Simkó, Mária Šimková, Kiril Simov, Maria Skachedubova, Aaron Smith, Isabela Soares-Bastos, Carolyn Spadine, Antonio Stella, Milan Straka, Jana Strnadová, Alane Suhr, Umut Sulubacak, Shingo Suzuki, Zsolt Szántó,

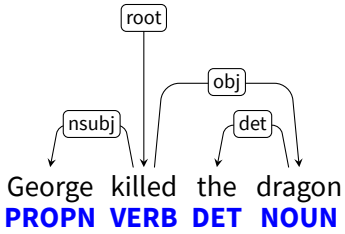
Manning's Law

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

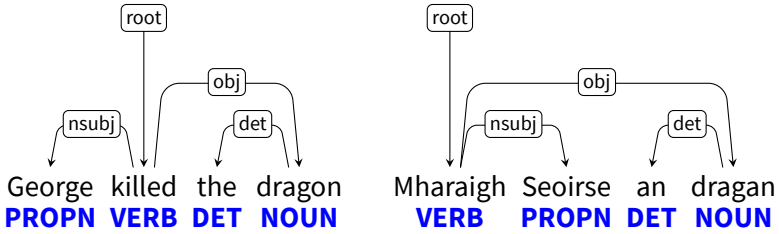
1. UD must be satisfactory on linguistic analysis grounds for individual languages.
2. UD must be good for linguistic typology, i.e., providing a suitable basis for bringing out cross-linguistic parallelism across languages and language families.
3. UD must be suitable for rapid, consistent annotation by a human annotator.
4. UD must be easily comprehended and used by a non-linguist, whether a language learner or an engineer with prosaic needs for language processing. … it leads us to favor traditional grammar notions and terminology.
5. UD must be suitable for computer parsing with high accuracy.
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.

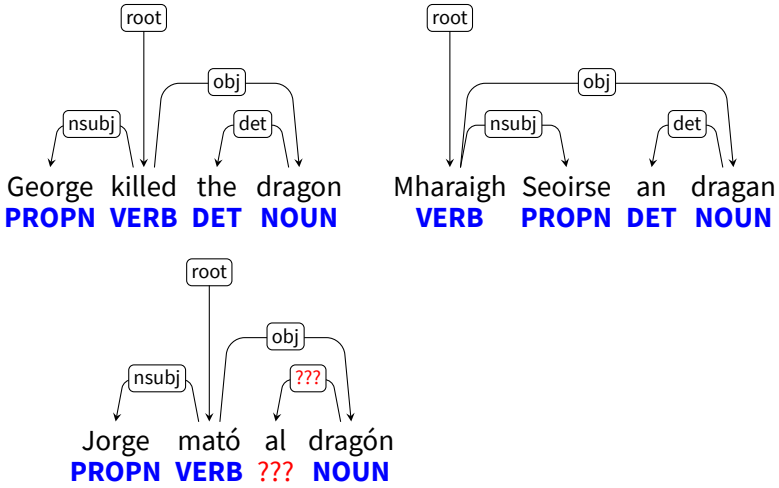
Same Thing Same Way



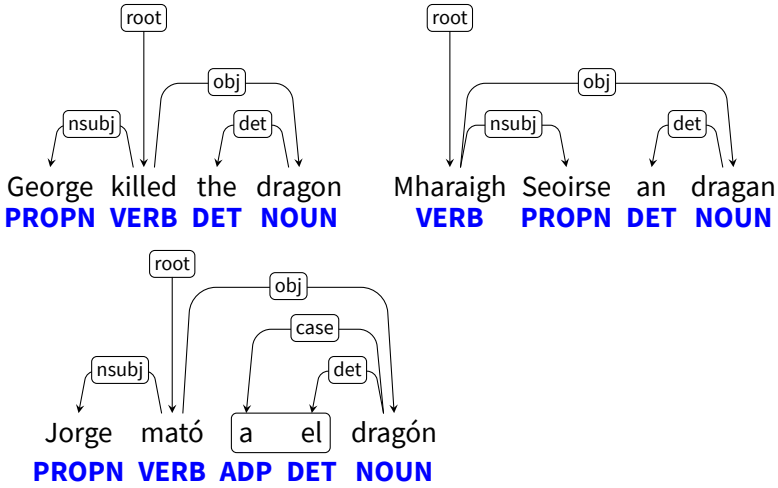
Same Thing Same Way



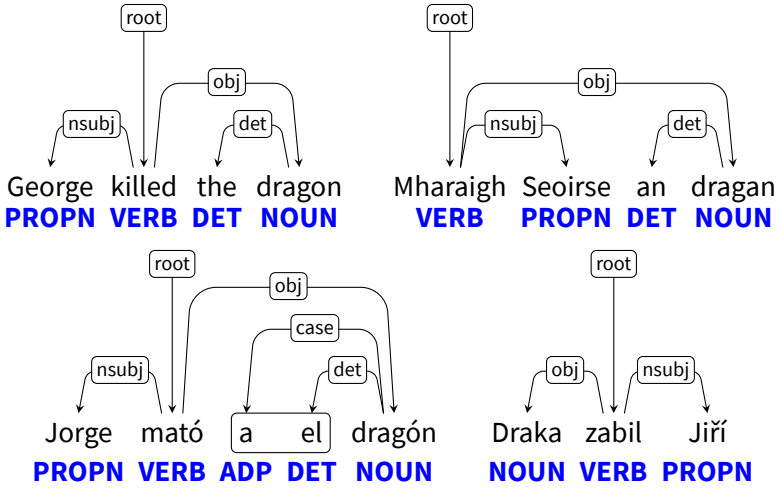
Same Thing Same Way



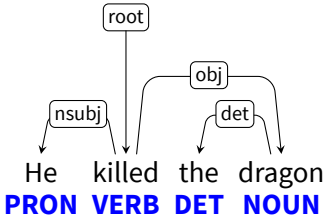
Same Thing Same Way



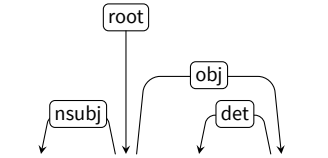
Same Thing Same Way

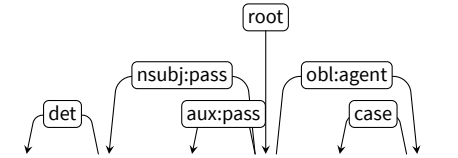


Same Meaning ≠ Same Construction!

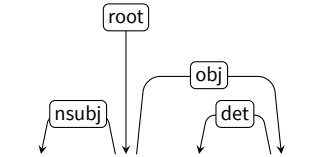


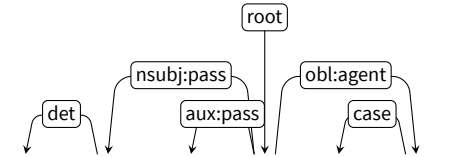
Same Meaning ≠ Same Construction!

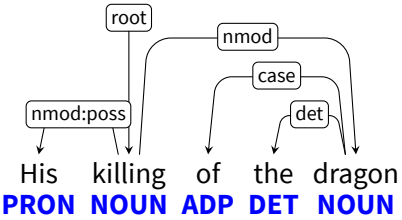




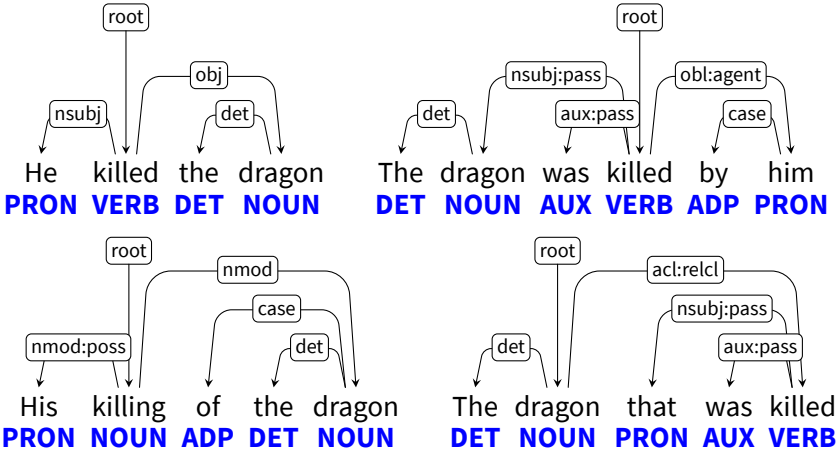
Same Meaning ≠ Same Construction!



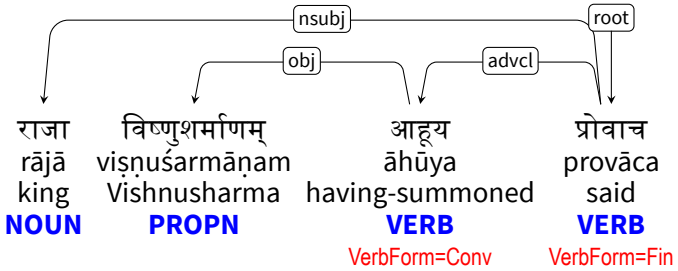




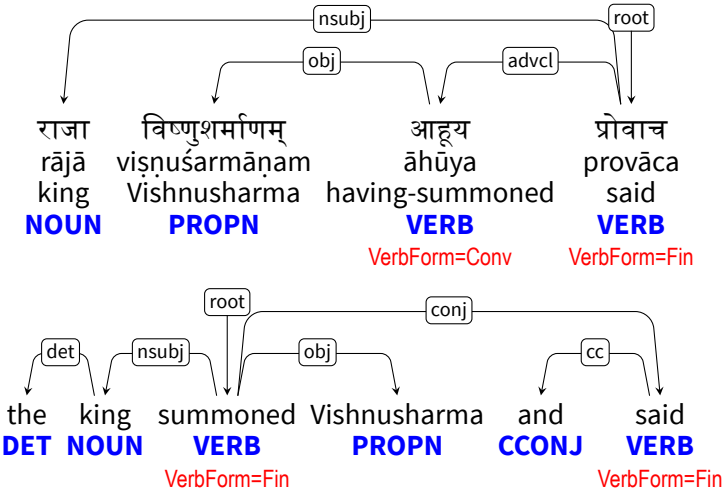
Same Meaning ≠ Same Construction!

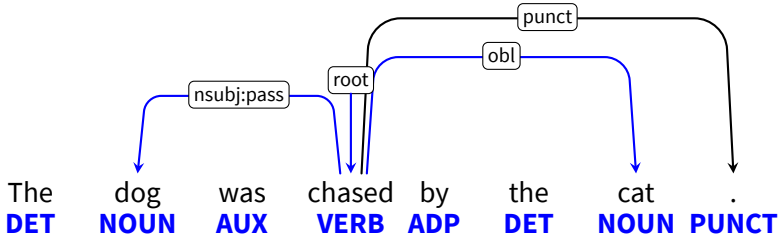


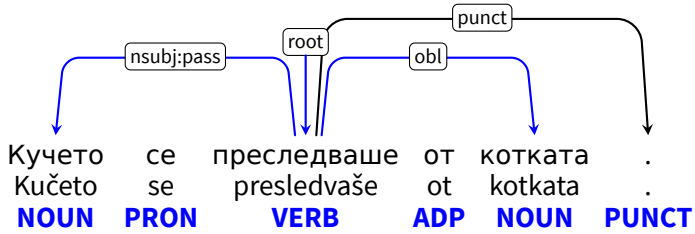
Language-specific Preferences

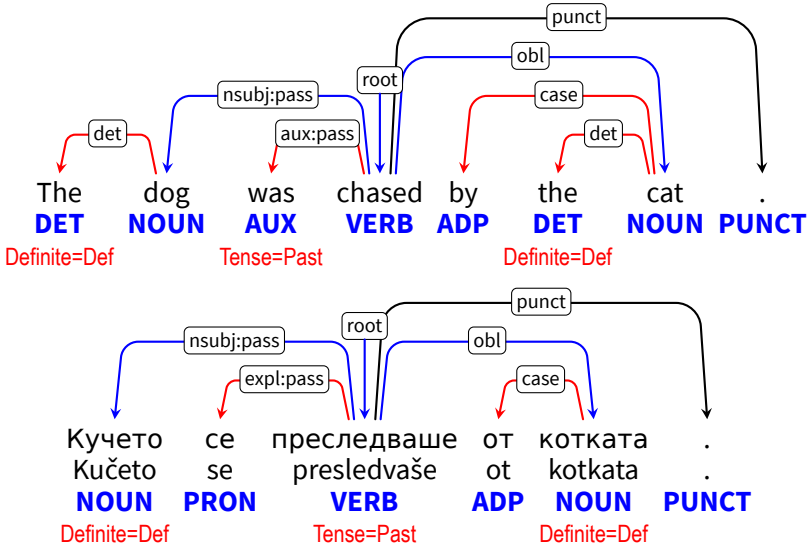


Language-specific Preferences




punct
obl
root
nsubj:pass
The dog was chased by the cat .
DET NOUN AUX VERB ADP DET NOUN PUNCT


punct
root
nsubj:pass
obl
Кучето се преследваше от котката .
Kučeto se presledvaše ot kotkata .
NOUN PRON VERB ADP NOUN PUNCT


punct
obl
root
nsubj:pass
case
det
aux:pass
det
The dog was chased by the cat .
DET NOUN AUX VERB ADP DET NOUN PUNCT
Definite=Def Tense=Past Definite=Def
punct
root
nsubj:pass
obl
expl:pass
case
Кучето се преследваше от котката .
Kučeto se presledvaše ot kotkata .
NOUN PRON VERB ADP NOUN PUNCT
Definite=Def Tense=Past Definite=Def

Dependents of Clauses (Verbal or Not)

	Nominal	Clausal	Modifier	Function
Core	**nsubj**	csubj		
Non-Core	**obl**	advcl	advmod	aux
	vocative		discourse	cop
	dislocated			mark
	expl			

Dependents of Verbs, Adjectives and Adverbs

	Nominal	Clausal	Modifier
Core	**obj**	ccomp	
	iobj	xcomp	
Non-Core	**obl**	advcl	advmod
	expl		

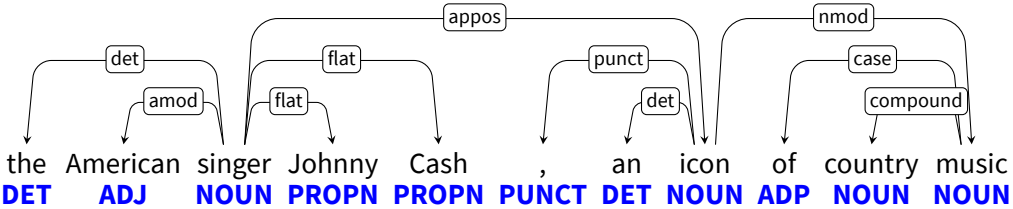
Dependents of Nominals

Nominal	Clausal	Modifier	Function
nmod	acl	amod	det

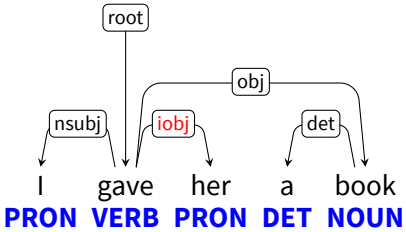
Noun Phrase

Dependents of Nominals

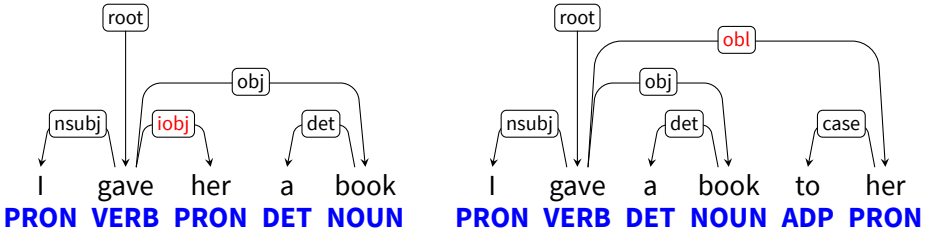
Nominal	**Clausal**	**Modifier**	**Function**
nmod	acl	amod	det
appos		nummod	case
compound			clf
flat			



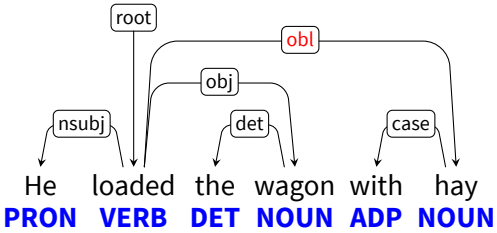
Information Packaging



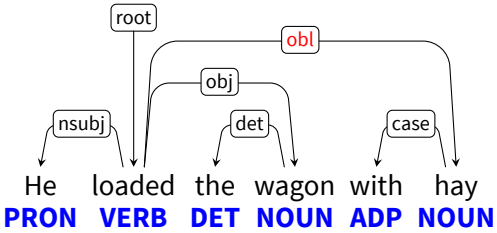
Information Packaging

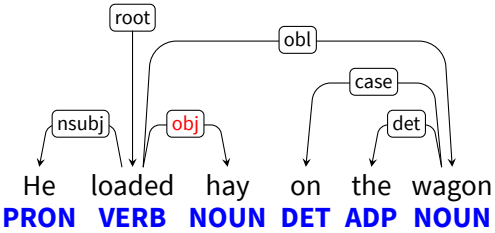


Information Packaging

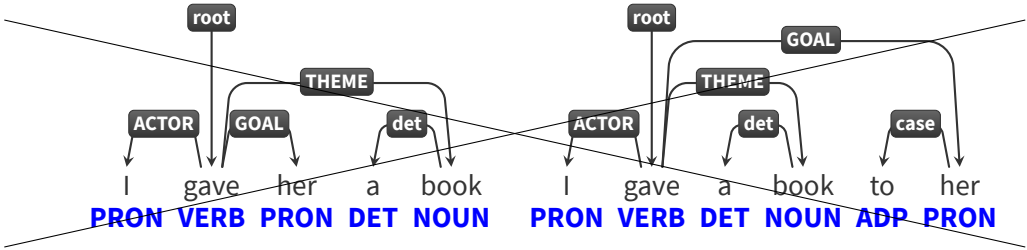


Information Packaging

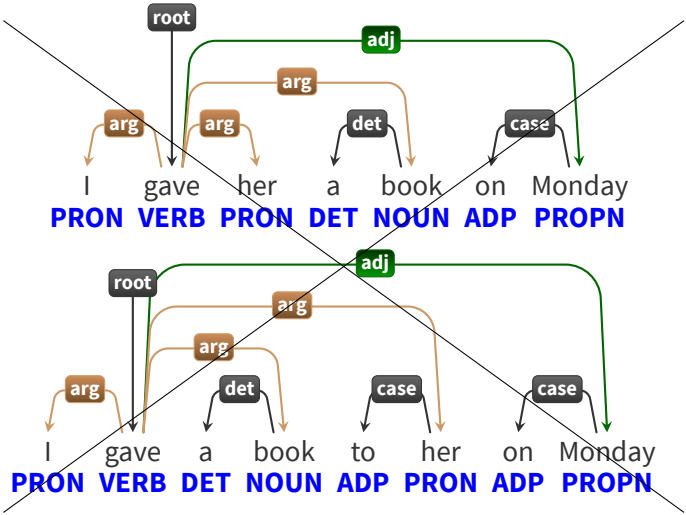




UD is NOT about Semantic Roles!



UD Avoids Argument-Adjunct Distinction!



So What Is Core and Why?

- UD v1 guidelines took core-oblique for granted
- English (simplified):
  - Bare noun phrase ⇒ core argument (`nsubj`, `obj`, `iobj`)
  - Prepositional phrase ⇒ oblique argument or adjunct (`obl`)

So What Is Core and Why?

- UD v1 guidelines took core-oblique for granted
- English (simplified):
  - Bare noun phrase ⇒ core argument (`nsubj`, `obj`, `iobj`)
  - Prepositional phrase ⇒ oblique argument or adjunct (`obl`)
- Other languages: not necessarily! (Spanish, Japanese)
  - But some people simply took the English rule…
  - Manning's law: non-linguists!

So What Is Core and Why?

- UD v1 guidelines took core-oblique for granted
- English (simplified):
  - Bare noun phrase ⇒ core argument (`nsubj`, `obj`, `iobj`)
  - Prepositional phrase ⇒ oblique argument or adjunct (`obl`)
- Other languages: not necessarily! (Spanish, Japanese)
  - But some people simply took the English rule…
  - Manning's law: non-linguists!
- Clash with traditional terminology
  - Grammars of German, Czech etc. define prepositional objects
  - But these are not necessarily core…
  - Yet some people took their national definition of object…

Language-specific Coding Strategy

- Idea:
  - Oblique arguments are marked similarly to adjuncts (prepositions, certain morphological cases…)
  - Core arguments are marked differently
    - ⇒ easy for annotators and non-linguists!
- Why are core arguments special?
  - They tend to be targeted by grammatical rules
    - Passivization
    - Control verbs
    - Reflexives
    - …

Language-specific Coding Strategy

- Core vs. oblique is not defined in traditional grammar
- How shall we define it?

Language-specific Coding Strategy

- Core vs. oblique is not defined in traditional grammar
- How shall we define it?

- Andrews, 2007 (In Shopen: Language Typology)
  - Identify primary transitive predicates
    - ★ We need semantic roles for this! (One-time only.)

Language-specific Coding Strategy

- Core vs. oblique is not defined in traditional grammar
- How shall we define it?

- Andrews, 2007 (In Shopen: Language Typology)
  - Identify primary transitive predicates
    - ★ We need semantic roles for this! (One-time only.)
  - Actor/agent = function A
  - Undergoer/patient = function P

Language-specific Coding Strategy

- Core vs. oblique is not defined in traditional grammar
- How shall we define it?

- Andrews, 2007 (In Shopen: Language Typology)
  - Identify primary transitive predicates
    - ★ We need semantic roles for this! (One-time only.)
  - Actor/agent = function A
  - Undergoer/patient = function P
  - Note the way they are coded
  - Note other grammatical rules that target them
  - Generalize to other predicates with same coding and rules

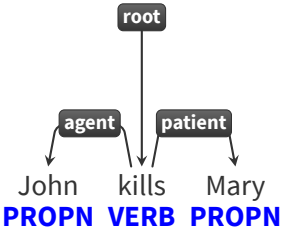
Language-specific Coding Strategy

- Core vs. oblique is not defined in traditional grammar
- How shall we define it?

- Andrews, 2007 (In Shopen: Language Typology)
  - Identify primary transitive predicates
    - We need semantic roles for this! (One-time only.)
  - Actor/agent = function A
  - Undergoer/patient = function P
  - Note the way they are coded
  - Note other grammatical rules that target them
  - Generalize to other predicates with same coding and rules

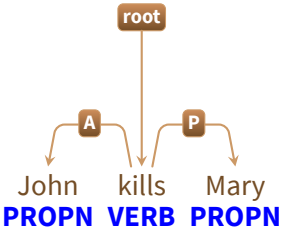
  - Then define:
    - function A ⇒ **nsubj**
    - function P ⇒ **obj**

Transitive Predicates in English



(primary transitive)

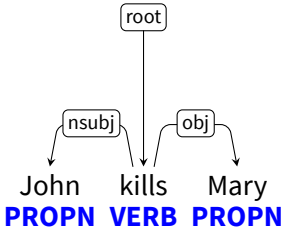
Transitive Predicates in English



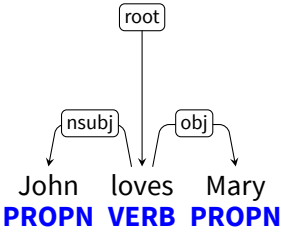
(primary transitive)

(generalized transitive)

Transitive Predicates in English

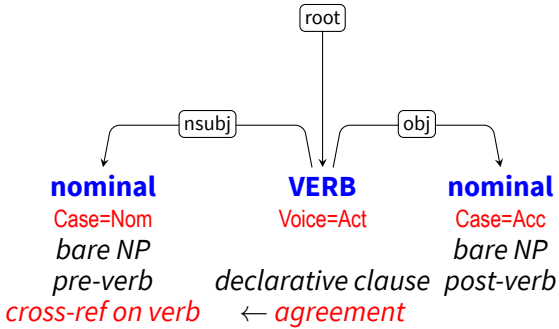


(primary transitive)

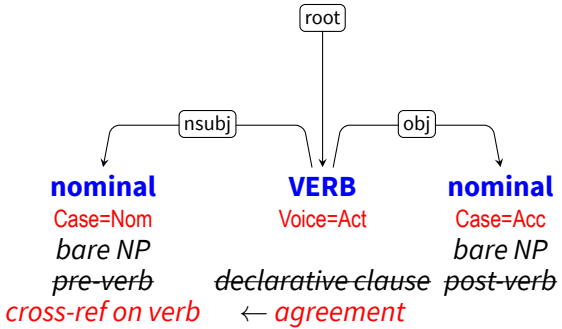


(generalized transitive)

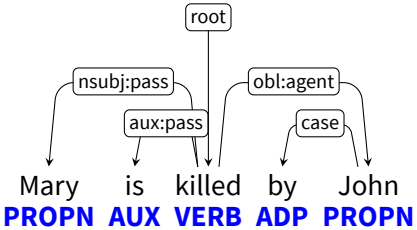
English Transitive Clauses

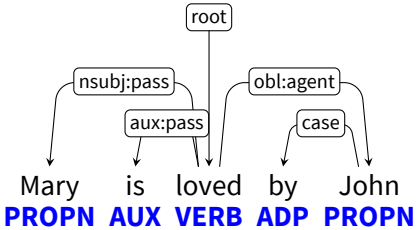


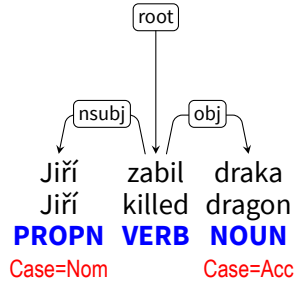
Czech Transitive Clauses

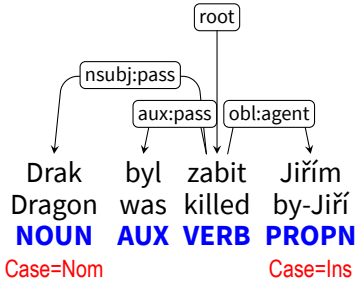


Passivization in English and Czech

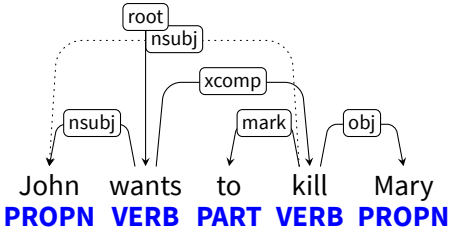


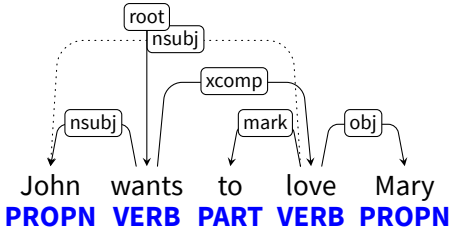


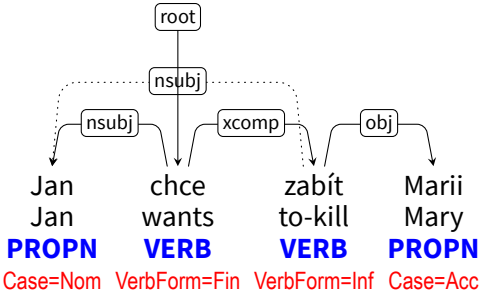


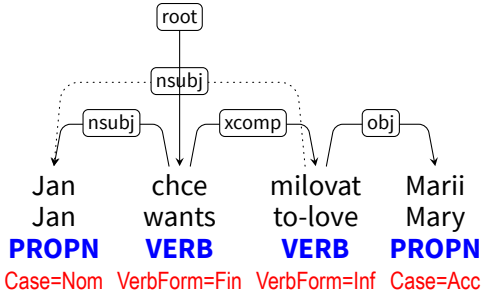


Subject Control in English and Czech

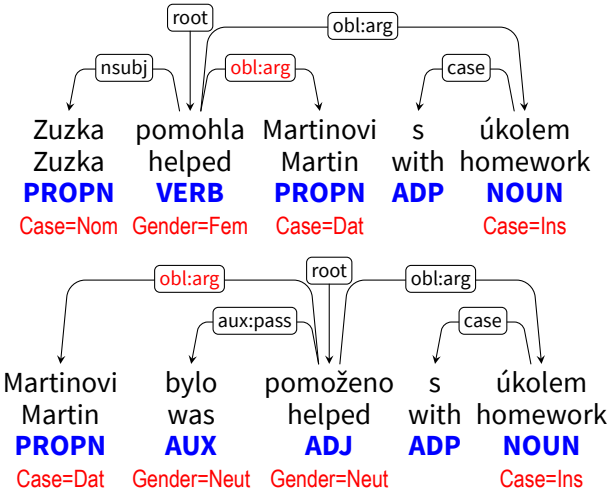






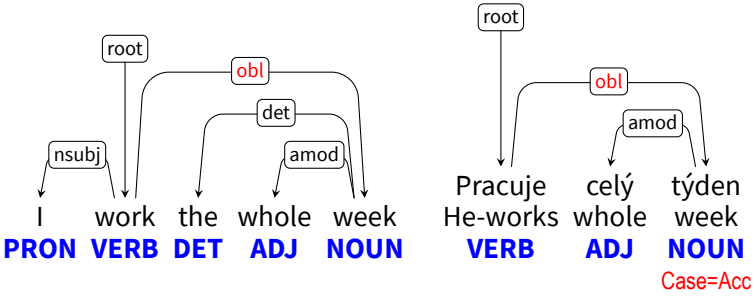


Czech: Bare Nominal in Dative/Genitive/Instrumental is Oblique



Some Problems

- Some temporal adjuncts are bare noun phrases
  - *I work the whole week.*
  - *I work every Friday.*



- At least it cannot passivize:
  - **The whole week is worked by me.*
  - **Every Friday is worked by me.*
- But…

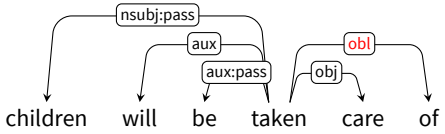
Some Problems

- Some transitive verbs cannot passivize
  - *John has a new car.*
    - **A new car is had by John.*
  - *Friday does not suit me.*
    - **I am not suited by Friday.*

Some Problems

- Some transitive verbs cannot passivize
  - *John has a new car.*
    - **A new car is had by John.*
  - *Friday does not suit me.*
    - **I am not suited by Friday.*

- Some prepositional verbs can passivize
  - *You can rely on Ben.*
    - *Ben can be relied on.*
  - *They will take care of your children.*
    - *Your children will be taken care of.*



Bare Temporal Adjuncts: Any Other Criteria?

- *I work the whole week.*
- *I work every Friday.*

- English has a fixed word order; adjuncts are less fixed than objects (not true in Czech):
  - *I work every Friday in Paris.*
  - *I work in Paris every Friday.*
  - *I spend every Friday in Paris.*
  - **I spend in Paris every Friday.*

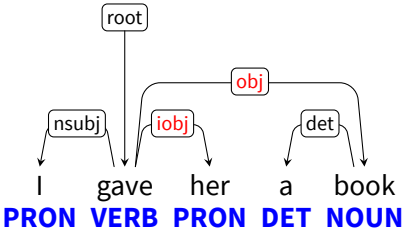
- Unlike objects, adjuncts cannot be replaced by pronouns (both English and Czech):
  - *Where do you spend this Friday? I spend it in Paris.*
  - *Where do you work this Friday? *I work it in Paris.*

Intransitive Predicates

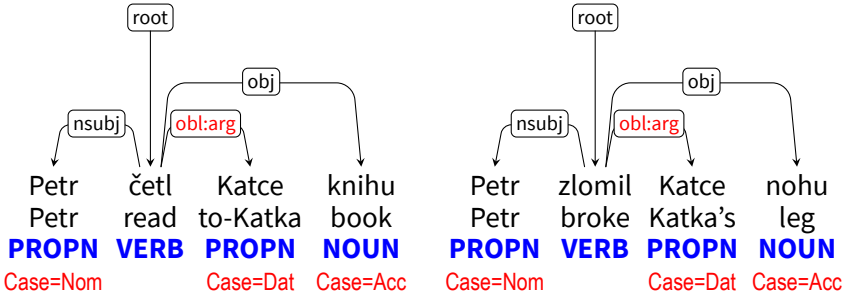
- Just one core argument
  - We already “know” how to find out if there are two
- ⇒ function S
  - Regardless of semantic role:
    - *John runs.*
    - *John sleeps.*
    - *John falls.*
- Then define:
  - function S ⇒ **`nsubj`**

Ditransitive Predicates

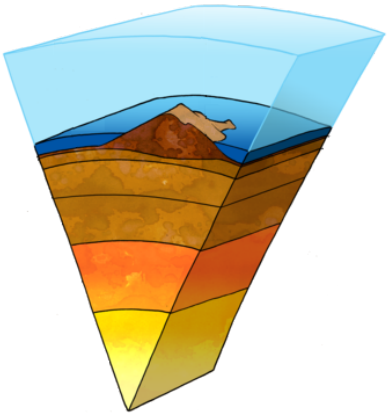
- Three core arguments, two of them objects
- To distinguish them, one object is *indirect*: **`iobj`**
- In English, `iobj` is closer to the verb (the recipient)



Oblique Dative: Recipient vs. Beneficiary



Multiple Layers of Dependencies



Form

- Surface syntax
- Deep syntax
- Semantics

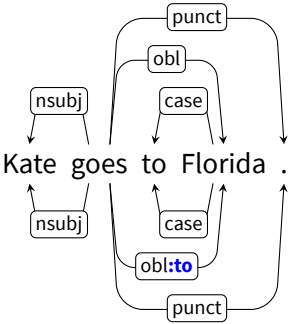
Meaning

Enhanced Universal Dependencies

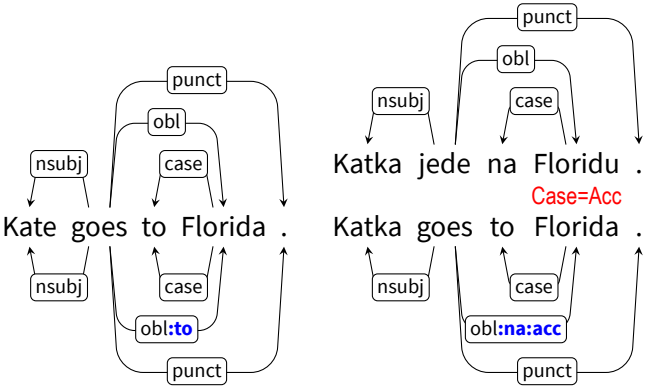
Currently available (at least partially) for 17 languages:

 Arabic,  Bulgarian,  Czech,  Dutch,  English,  Estonian,  Finnish,  French,  Italian,  Latvian,  Lithuanian,  Polish,  Russian,  Slovak,  Swedish,  Tamil,  Ukrainian

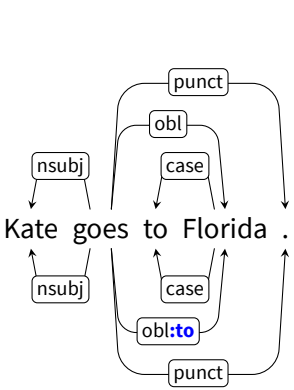
Enhanced UD: Case Information in Dependency Label

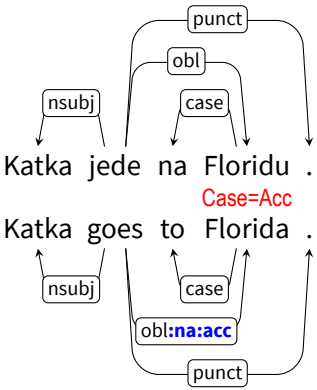


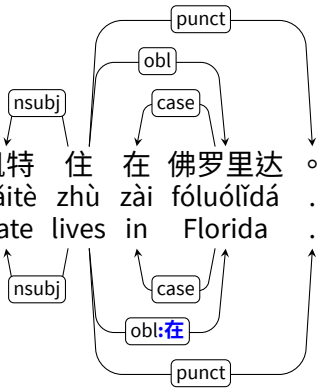
Enhanced UD: Case Information in Dependency Label



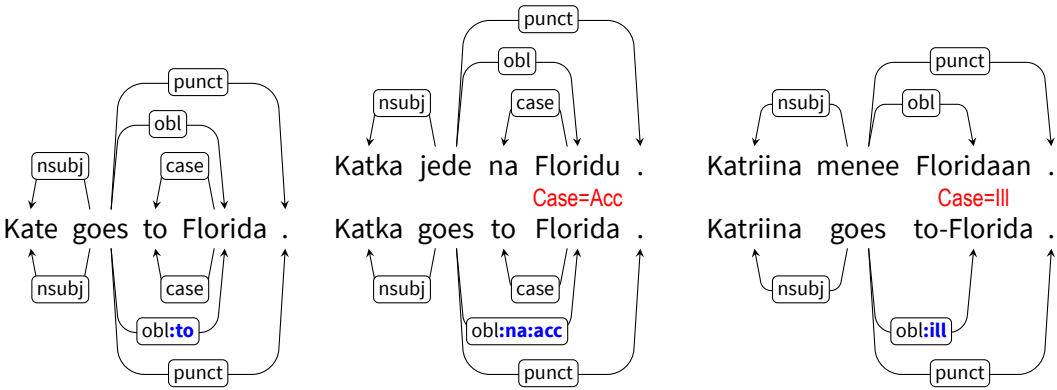
Enhanced UD: Case Information in Dependency Label



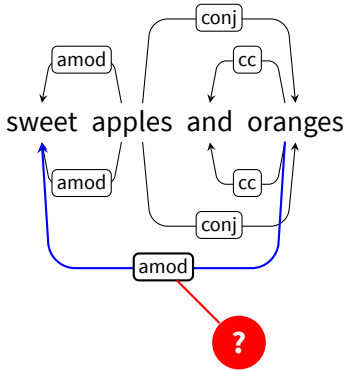




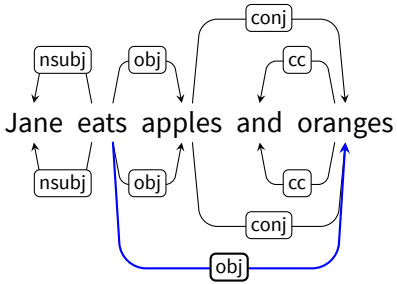
Enhanced UD: Case Information in Dependency Label



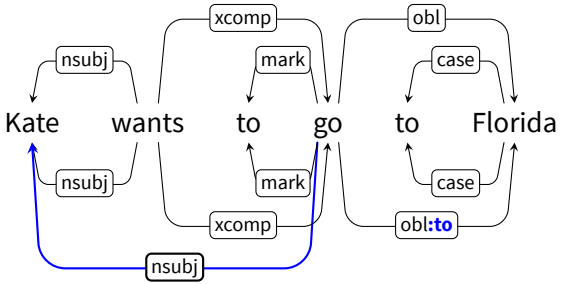
Enhanced UD: Shared Dependent of Coordination



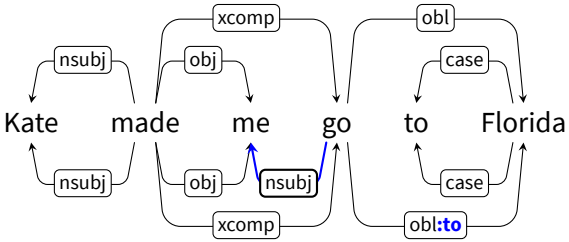
Enhanced UD: Parent of Coordination



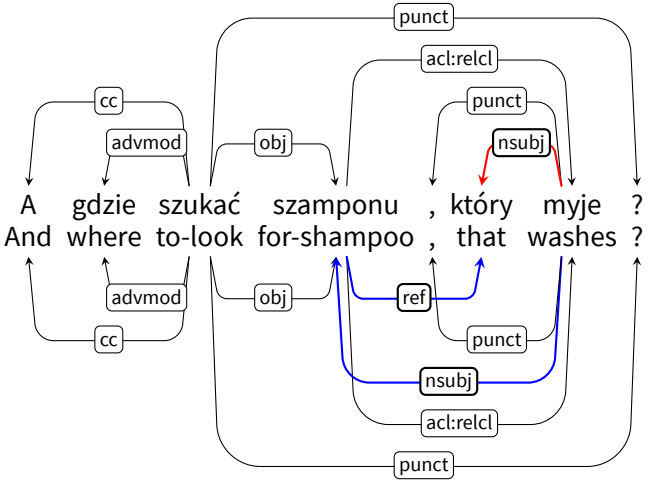
Enhanced UD: External Subject of Controlled Predicate



Enhanced UD: External Subject in Object-Control Construction

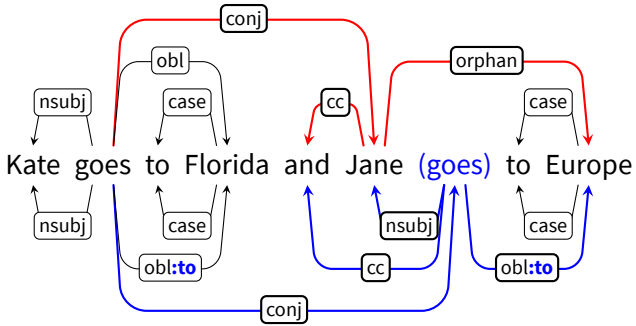


Enhanced UD: Relative Clauses

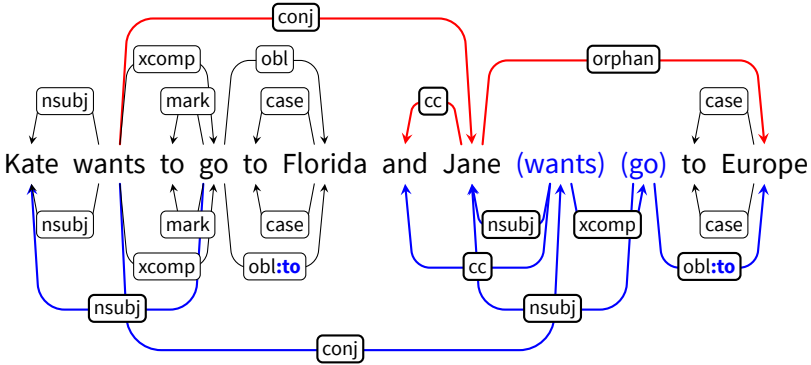


“And where to look for shampoo that works?”

Enhanced UD: Gapping and Stripping



Enhanced UD: Gapping + Control Verb Construction



Enhanced UD: Five (Six) Enhancements

- Null nodes for gapping (20 treebanks in UD 2.6)
- Dependency propagation in coordination
  - Common parent of coordination (24 treebanks)
  - Shared dependents of coordination (22 treebanks)
- External subjects of controlled predicates (18 treebanks)
- Cyclic dependencies to/from relative clauses (16 treebanks)
- Case-enhanced dependency labels (19 treebanks)

- All 5 types: 11 treebanks, 6 languages
- At least 1 type: 26 treebanks, 16 languages
- Only basic UD: 137 treebanks

IWPT 2020 Shared Task in Parsing EUD

- `https://universaldependencies.org/iwpt20/`
- Co-organized with Djamé Seddah and Gosse Bouma

- End-to-end from raw text to EUD
- Data in 17 languages (UD 2.5 + French)

- Baseline: UDPipe + Stanford Enhancer … 61.1%
- Winner: TurkuNLP (compress graph into tree) … **84.5%**

Towards “Deep” Universal Dependencies

joint work with
Kira Droganova

Our “Enhanced Plus”

- Enhanced UD help us identify more predicate-argument relations
- But some patterns are still not handled…

- Adverbial **infinitives**
  - *They will meet to discuss a contract.*
  - But not `ccomp` infinitives: *He recommended to replace the tyres.*

Our “Enhanced Plus”

- Enhanced UD help us identify more predicate-argument relations
- But some patterns are still not handled…

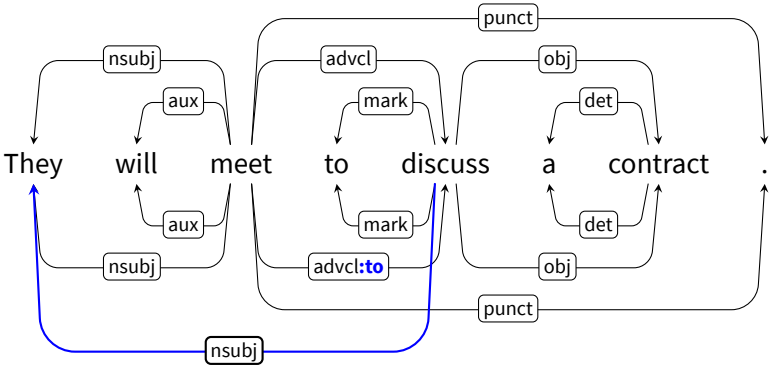
- Adverbial **infinitives**
  - *They will meet to discuss a contract.*
  - But not `ccomp` infinitives: *He recommended to replace the tyres.*
- Adverbial **converbs (gerunds)**
  - *Terrorists detonated a bomb killing five people.*

Our “Enhanced Plus”

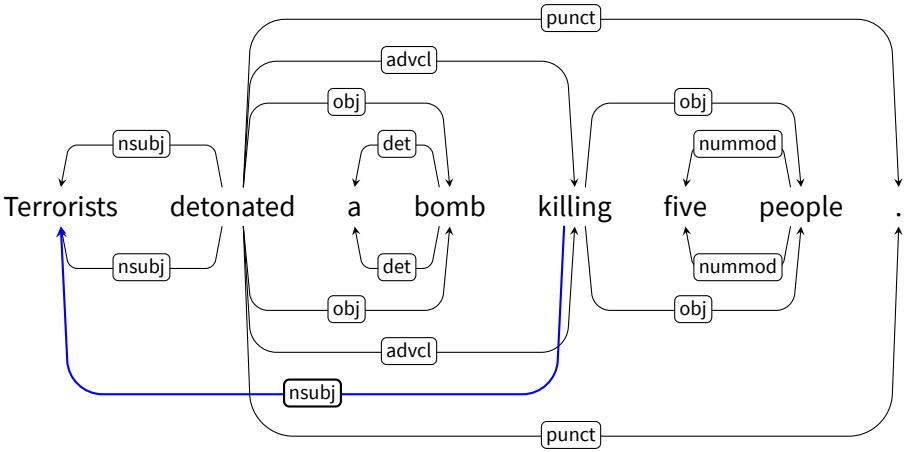
- Enhanced UD help us identify more predicate-argument relations
- But some patterns are still not handled…

- Adverbial **infinitives**
  - *They will meet to discuss a contract.*
  - But not `ccomp` infinitives: *He recommended to replace the tyres.*
- Adverbial **converbs (gerunds)**
  - *Terrorists detonated a bomb killing five people.*
- Attributive **participles**
  - *The shares reflected on your statement.*

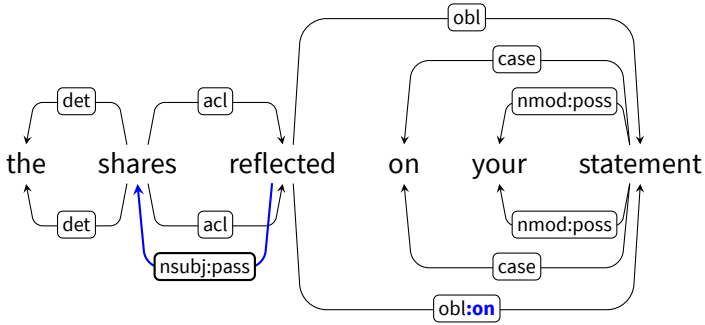
External Subject of Adverbial Infinitive



External Subject of Adverbial Gerund

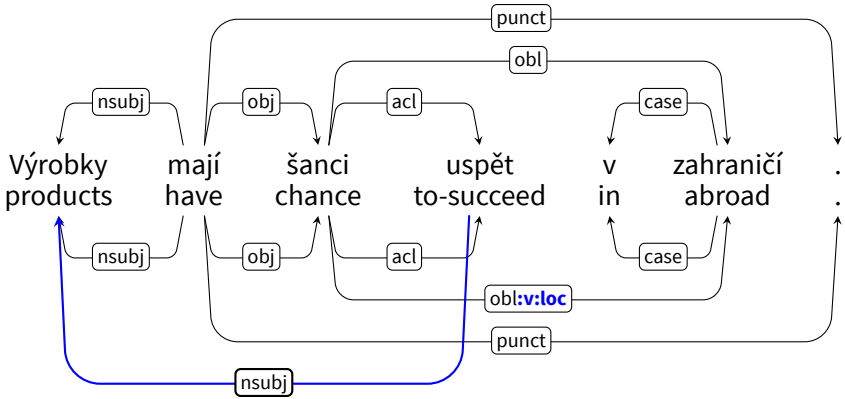


External Subject of Participle



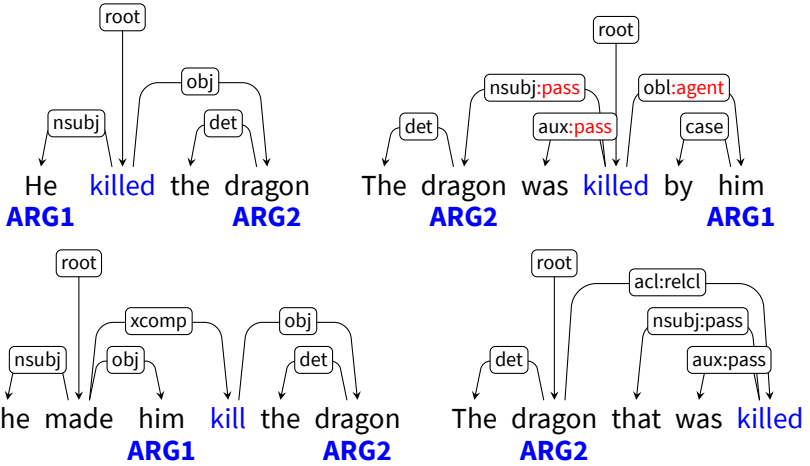
Possible Addition: Control by Light-Verb Constructions

We do not do this yet:



“The products have a chance to succeed abroad.”

Deep UD: Normalization of Syntactic Alternations



Numbered Arguments

- Degree of salience of arguments derived from surface syntax:
  - Subject of active clause ⇒ **ARG1**
  - Direct object of active clause ⇒ **ARG2**
  - Indirect object of active clause ⇒ **ARG3**
  - Subject of passive clause ⇒ **ARG2**
  - etc…
- Numbers can be mapped to semantic roles if we have a valency dictionary
- Like in PropBank… SOMEWHAT!

Numbered Arguments

- Degree of salience of arguments derived from surface syntax:
  - Subject of active clause ⇒ **ARG1**
  - Direct object of active clause ⇒ **ARG2**
  - Indirect object of active clause ⇒ **ARG3**
  - Subject of passive clause ⇒ **ARG2**
  - etc…
- Numbers can be mapped to semantic roles if we have a valency dictionary
- Like in PropBank… SOMEWHAT!

	Deep UD	PropBank
John broke the window.	*John*.ARG1 *window*.ARG2	*John*.ARG0 *window*.ARG1
The window was broken by John.	*John*.ARG1 *window*.ARG2	*John*.ARG0 *window*.ARG1
The window broke.	*window*.ARG1	*window*.ARG1

Predicate Identfiers

- They could be sense/frame identifiers
- But now we just take lemmas

- Exception:
  - Germanic phrasal verbs: *come_up*
  - Inherently reflexive verbs: `[cs]` *smát_se* “laugh”
  - Other compound verbs (incl. light & serial verb constructions)

Summary

- Deep UD is closer to meaning ⇒ easier to discover **who did what to whom**

- Deep UD 2.6 (`http://hdl.handle.net/11234/1-3242`)
- 140 treebanks of 82 languages
- Enhanced graphs in all treebanks
- (Enhanced Plus: infinitives, gerunds, participles)
- Normalized active-passive

- Can be regenerated after each UD release (every 6 months)

Thanks!
Ačiū!