The State of Universal Dependencies

Marie-Catherine de Marneffe, Joakim Nivre, **Daniel Zeman**
UCLouvain / Ohio State University
RISE Research Institutes of Sweden / Uppsala University
Charles University

What is Universal Dependencies?

“Universal Dependencies (UD) is a project developing cross-linguistically consistent treebank annotation for many languages, with the goal of facilitating multilingual parser development, cross-lingual learning, and parsing research from a language typology perspective.”

(https://universaldependencies.org/introduction.html)

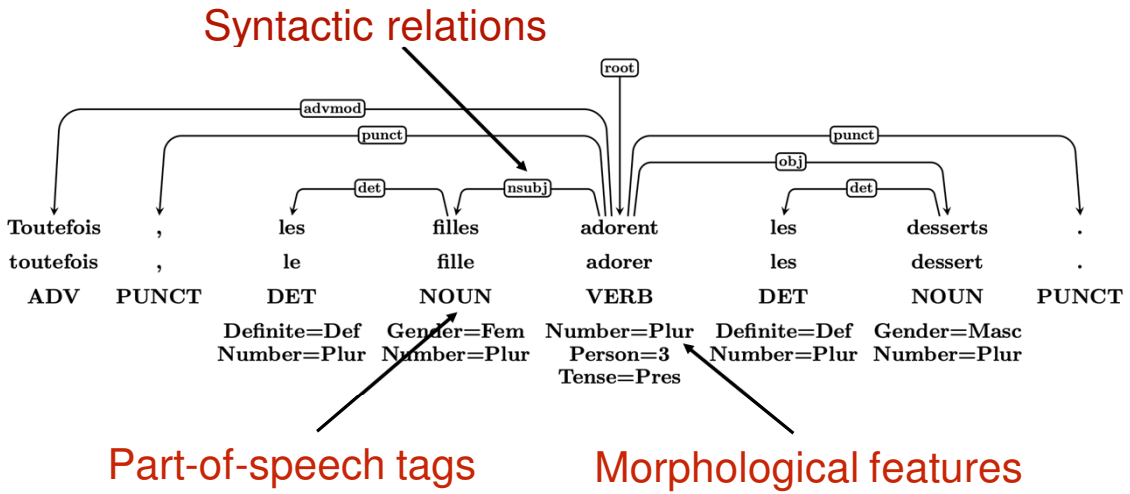
A project – but also an annotation scheme, a data repository and a community.

The UD Annotation Scheme

Basic design principles:

- Words as basic units
- Binary syntactic relations

Built on de facto standards



J. Nivre, M.-C. de Marneffe, F. Ginter, Y. Goldberg, J. Hajič, C. D. Manning, R. McDonald, S. Petrov, S. Pyysalo, N. Silveira, R. Tsarfaty, and D. Zeman (2016) Universal Dependencies v1: A Multilingual Treebank Collection. In *Proceedings of LREC*, 1659–1666.

J. Nivre, M.-C. de Marneffe, F. Ginter, J. Hajič, C. D. Manning, S. Pyysalo, S. Schuster, F. Tyers, and D. Zeman (2020) Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection. In *Proceedings of LREC*, 4034–4043.

M.-C. de Marneffe, C. D. Manning, J.Nivre, and D. Zeman (2021) Universal Dependencies. *Computational Linguistics*, 47(2):255–308, 2021.

CoNLL-U Format

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC
1	Es	es	PRON	_	_	2	nsubj	_	_
2	unterscheidet	unterscheiden	VERB	_	_	0	root	_	_
3	sich	sich	PRON	_	_	2	expl:pv	_	_
4-5	vom	_	_	_	_	_	_	_	_
4	von	von	ADP	_	_	7	case	_	_
5	dem	der	DET	_	_	7	det	_	_
6	westlichen	westlich	ADJ	_	_	7	amod	_	SpaceAfter=No
7	Teil	Teil	NOUN	_	_	2	obl	_	_
8	des	der	DET	_	_	9	det	_	_
9	Landes	Land	NOUN	_	_	7	nmod	_	SpaceAfter=No
10	.	.	PUNCT	_	_	2	punct	_	_

- Revised and extended version of CoNLL-X format
- Two-level segmentation and enhanced dependencies

Word Segmentation

Let's go to the sea.

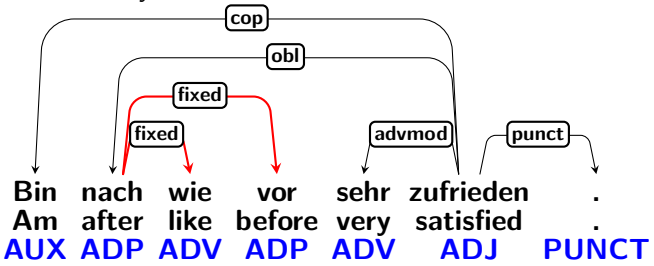
Vámonos	**al**	**mar**	**.**	**Vamos**	**nos**	**a**	**el**	**mar**	**.**
VERB?	**X**	**NOUN**	**PUNCT**	**VERB**	**PRON**	**ADP**	**DET**	**NOUN**	**PUNCT**

- **Syntactic word** vs. orthographic word
- **Multi-word tokens**
- Two-level scheme:
  - Tokenization (low level, punctuation, concatenative)
  - Word segmentation (higher level, not necessarily concatenative)

Fixed Expressions

One syntactic word spans several orthographic words?

I am still very satisfied.



Part-of-Speech Tags

http://universaldependencies.org/u/pos/index.html

Open		Closed		Other	
NOUN	common noun	**PRON**	pronoun	**PUNCT**	punctuation
PROPN	proper noun	**DET**	determiner	**SYM**	symbol
VERB	verb	**AUX**	auxiliary	**X**	unknown
ADJ	adjective	**NUM**	numeral		
ADV	adverb	**ADP**	adposition		
INTJ	interjection	**SCONJ**	subordinator		
		CCONJ	coordinator		
		PART	particle		

- Taxonomy of 17 universal POS tags
- All languages use the same inventory
  - Not all tags have to be used by all languages
  - Need extensions? Use features!

Features

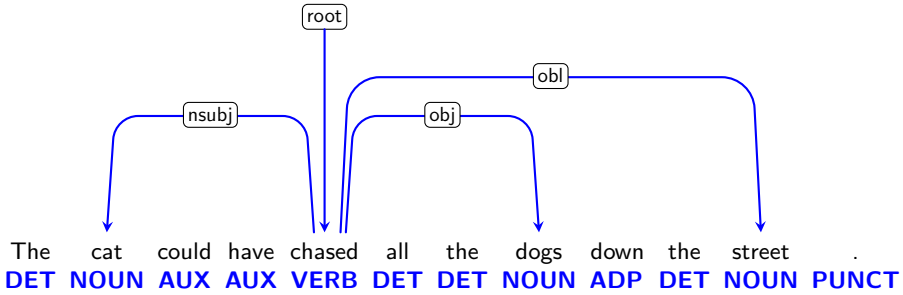
Lexical	Inflectional ("Nominal")	Inflectional ("Verbal, Pronominal")
PronType	Gender	VerbForm
NumType	Animacy	Mood
Poss	NounClass	Tense
Reflect	Number	Aspect
Foreign	Case	Voice
	Definite	Evident
	Degree	Polarity
Abbr		Person
Typo		Polite
		Clusivity

- 24 features, each with a number of possible *values*
- Languages select relevant features
- May add language-specific features or values

Syntax

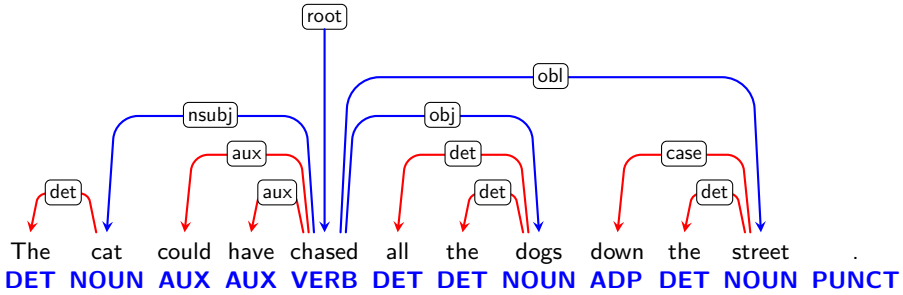
The	cat	could	have	chased	all	the	dogs	down	the	street	.
DET	**NOUN**	**AUX**	**AUX**	**VERB**	**DET**	**DET**	**NOUN**	**ADP**	**DET**	**NOUN**	**PUNCT**

Syntax



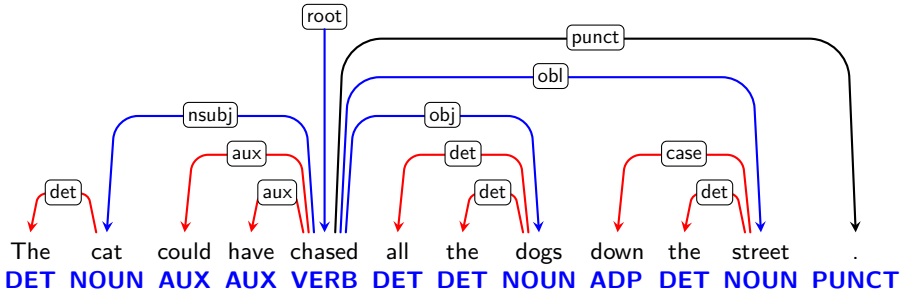
- Content words are related by dependency relations

Syntax



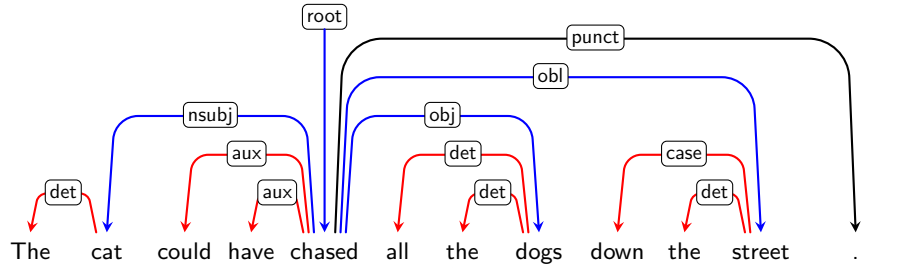
- Content words are related by dependency relations
- Function words attach to closest content words

Syntax

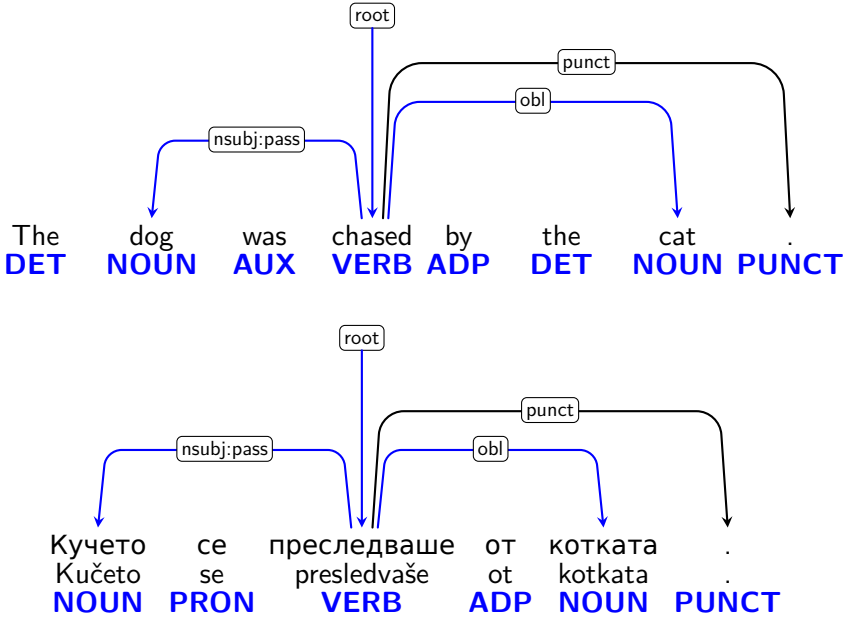


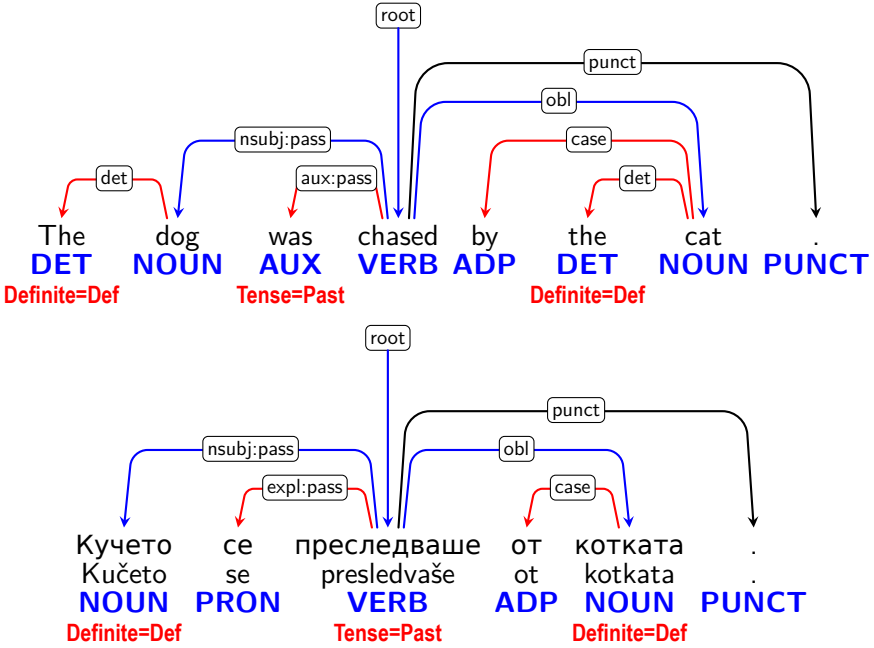
- Content words are related by dependency relations
- Function words attach to closest content words
- Punctuation attach to head of phrase or clause

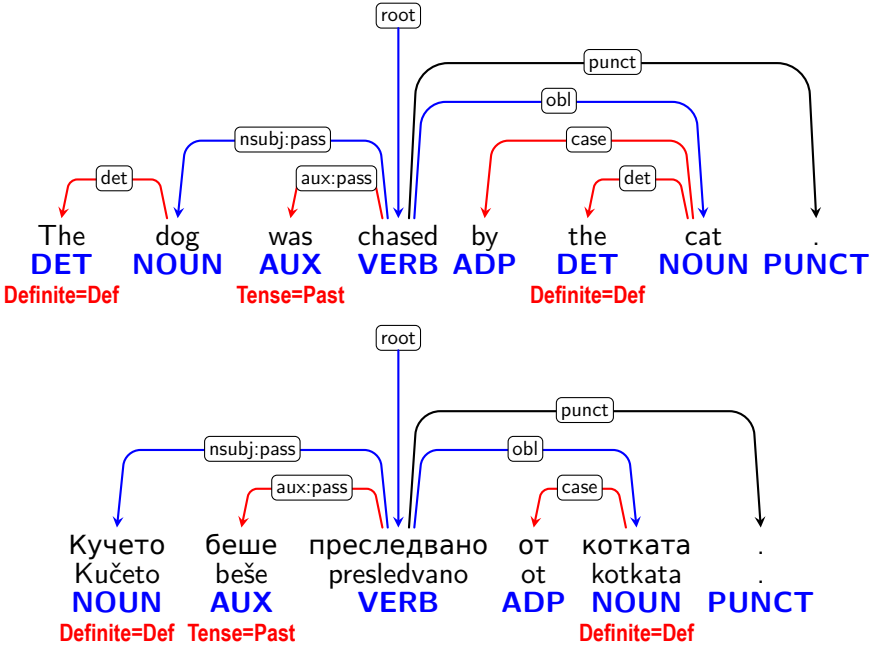
Syntax

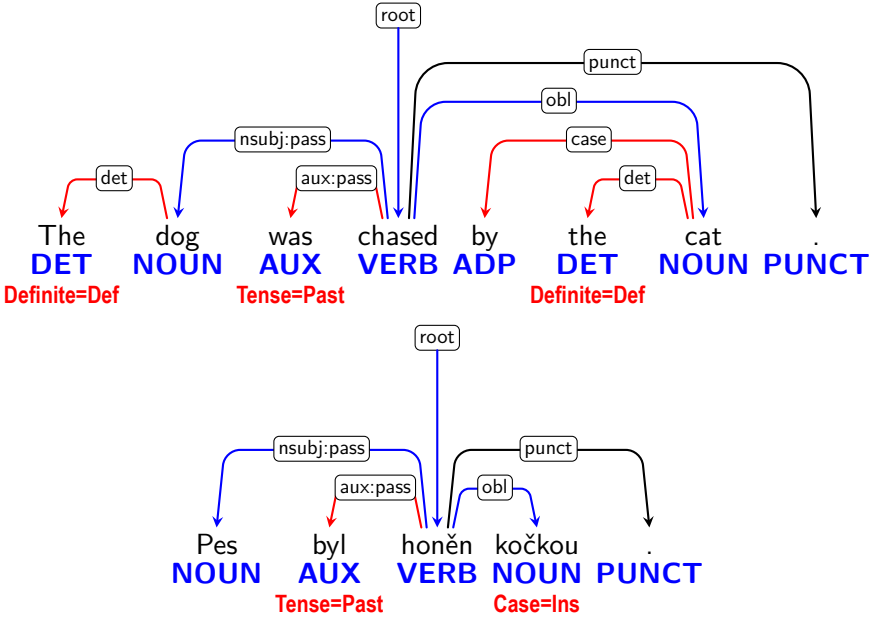


Not "dependency" in the strictly syntactic sense!


root
punct
obl
nsubj:pass
The dog was chased by the cat .
DET NOUN AUX VERB ADP DET NOUN PUNCT
root
punct
nsubj:pass
obl
Кучето се преследваше от котката .
Kučeto se presledvaše ot kotkata .
NOUN PRON VERB ADP NOUN PUNCT


root
punct
obl
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
root
punct
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


root
punct
obl
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
root
punct
nsubj:pass
obl
aux:pass
case
Кучето беше преследвано от котката .
Kučeto beše presledvano ot kotkata .
NOUN AUX VERB ADP NOUN PUNCT
Definite=Def
Tense=Past
Definite=Def


root
punct
obl
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
root
nsubj:pass
punct
aux:pass
obl
Pes byl honěn kočkou .
NOUN AUX VERB NOUN PUNCT
Tense=Past
Case=Ins

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		

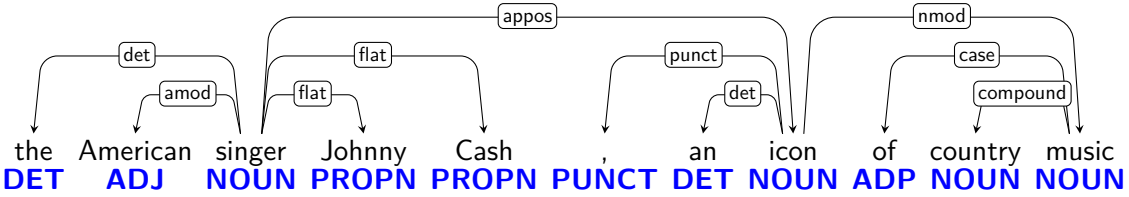
Dependents of Nominals

Nominal	Clausal	Modifier	Function
nmod	acl	amod	det
appos		nummod	case

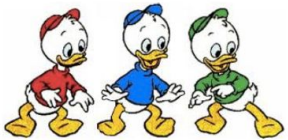
Noun Phrase

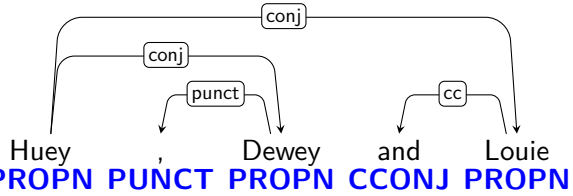
Dependents of Nominals

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



Coordination





- Coordinate structures are headed by the first conjunct
  - Subsequent conjuncts depend on it via the conj relation
  - Conjunctions depend on the next conjunct via the cc relation
  - Punctuation marks depend on the next conjunct via the punct relation

Multiword Expressions

Relation	Examples
fixed	*as well, by and large, according to, more than*
flat	*president Havel, New York, four thousand*
compound	*phone book, dress up*
goeswith	*notwith standing, with out*

- UD annotation almost does not permit "words with spaces"
  - Multiword expressions are analyzed using special relations
  - The fixed, flat and goeswith relations are always head-initial
  - The compound relation reflects the internal structure
- Words with spaces allowed in exceptional cases:
  - Vietnamese (spaces delimit syllables, not words)
  - Numbers ("1 000 000")
  - Possibly other approved cases, e.g. multi-word abbreviations

Other Relations

Other Relations

Relation	Explanation
parataxis	Loosely linked clauses of same rank
list	Lists without syntactic structure
orphan	Orphans in ellipsis linked together
reparandum	Disfluency linked to (speech) repair
dep	Unspecified dependency
root	The single syntactically independent element of the sentence

Language-specific Relation Subtypes

- Language-specific relations are subtypes of universal relations added to capture important phenomena
- Subtyping permits us to "back off" to universal relations

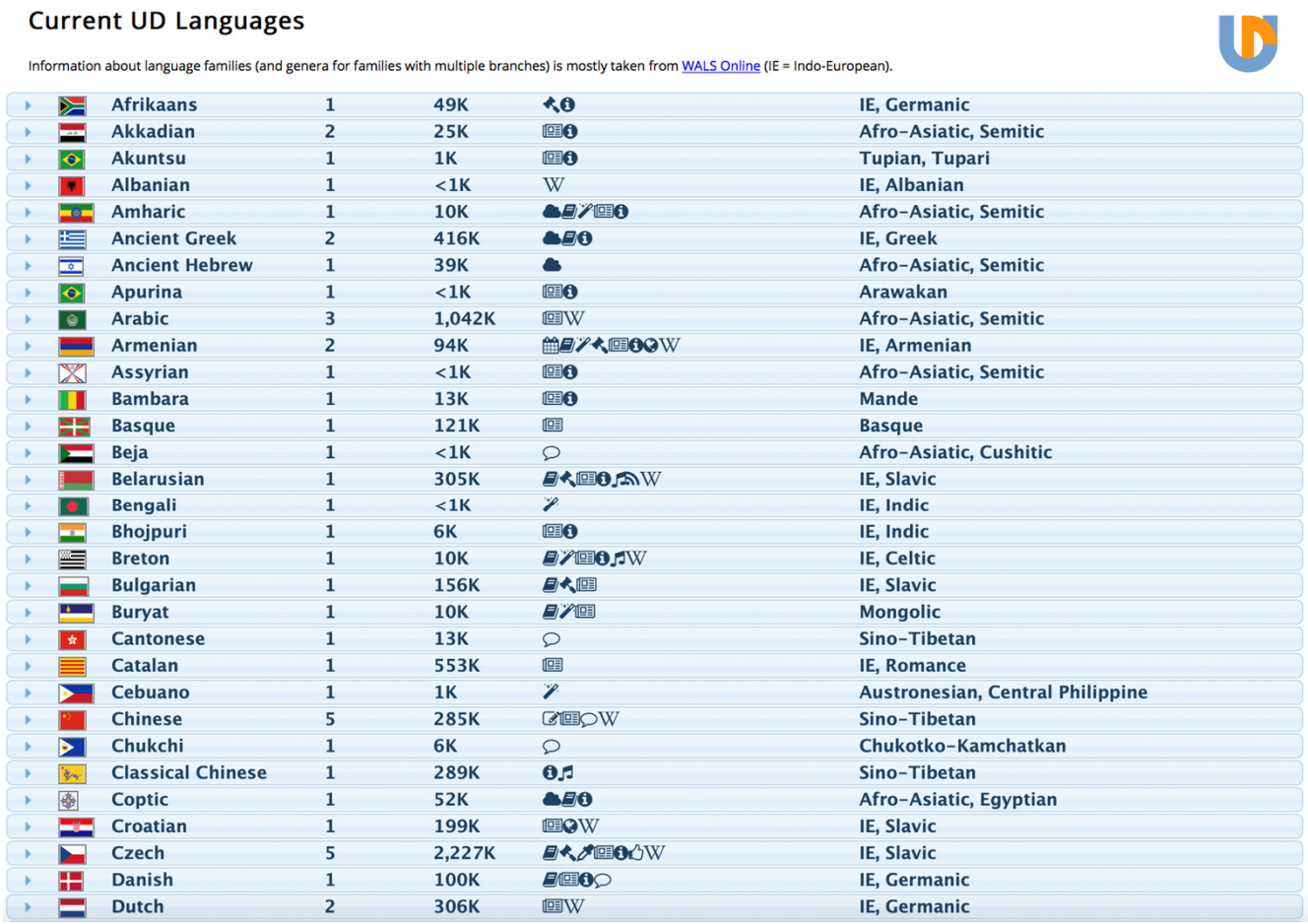
Language-specific Relation Subtypes

Relation	Explanation
acl:relcl	Relative clause (the boy who lived)
compound:prt	Verb particle (dress up)
nmod:poss	Possessive nominal (Mary 's book)
obl:agent	Agent in passive (saved by the bell)
cc:preconj	Preconjunction (both ... and)
det:predet	Predeterminer (all those ...)

The UD Repository

UD v2.11:

- 25 language families
- 138 languages
- 243 treebanks
- 535 contributors
- 1.8 million sentences
- 31 million words

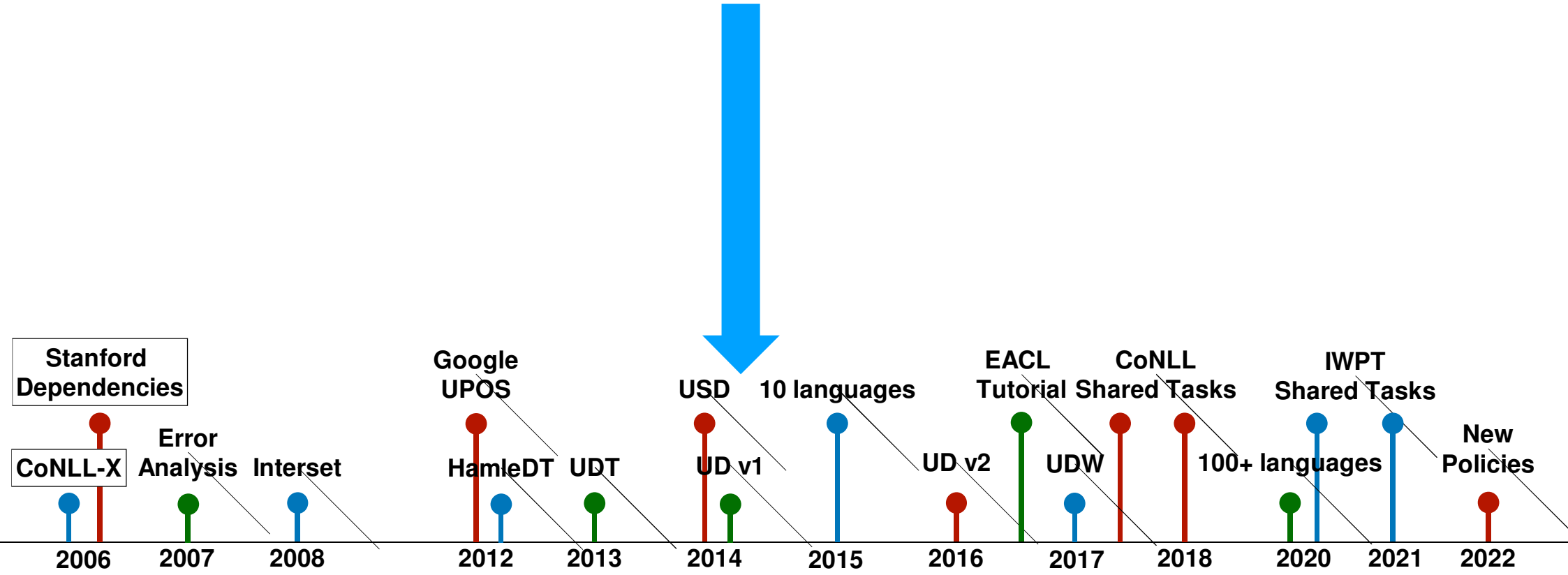


Current UD Languages

Information about language families (and genera for families with multiple branches) is mostly taken from WALS Online (IE = Indo-European).

Afrikaans	1	49K		IE, Germanic
Akkadian	2	25K		Afro-Asiatic, Semitic
Akuntsu	1	1K		Tupian, Tupari
Albanian	1	<1K	W	IE, Albanian
Amharic	1	10K		Afro-Asiatic, Semitic
Ancient Greek	2	416K		IE, Greek
Ancient Hebrew	1	39K		Afro-Asiatic, Semitic
Apurina	1	<1K		Arawakan
Arabic	3	1,042K	W	Afro-Asiatic, Semitic
Armenian	2	94K	W	IE, Armenian
Assyrian	1	<1K		Afro-Asiatic, Semitic
Bambara	1	13K		Mande
Basque	1	121K		Basque
Beja	1	<1K		Afro-Asiatic, Cushitic
Belarusian	1	305K	W	IE, Slavic
Bengali	1	<1K		IE, Indic
Bhojpuri	1	6K		IE, Indic
Breton	1	10K	W	IE, Celtic
Bulgarian	1	156K		IE, Slavic
Buryat	1	10K		Mongolic
Cantonese	1	13K		Sino-Tibetan
Catalan	1	553K		IE, Romance
Cebuano	1	1K		Austronesian, Central Philippine
Chinese	5	285K	W	Sino-Tibetan
Chukchi	1	6K		Chukotko-Kamchatkan
Classical Chinese	1	289K		Sino-Tibetan
Coptic	1	52K		Afro-Asiatic, Egyptian
Croatian	1	199K	W	IE, Slavic
Czech	5	2,227K	W	IE, Slavic
Danish	1	100K		IE, Germanic
Dutch	2	306K	W	IE, Germanic

A Brief History of UD



So where are we now?

Achievements:

- Amazing growth in terms of contributors and treebanks
- Data sets widely used for research in NLP and linguistics

Challenges:

- Annotation scheme and documentation can be improved
- Data sets are uneven in terms of quality and quantity
- Relevance to the research community cannot be taken for granted

Annotation Scheme

Old policy:

- A version change requires a complete revision of all guidelines
- Guidelines cannot be changed between major versions
- Documentation can be updated only with clarifications

Considerations:

- The old policy favors stability at the expense of improvement
- Implementing a major version change has become increasingly difficult
- We need to be able to make incremental changes to the guidelines

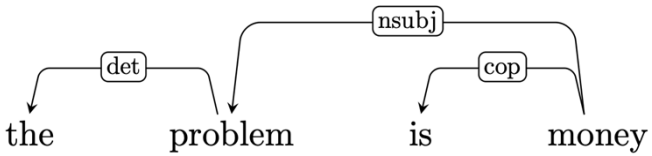
Changes

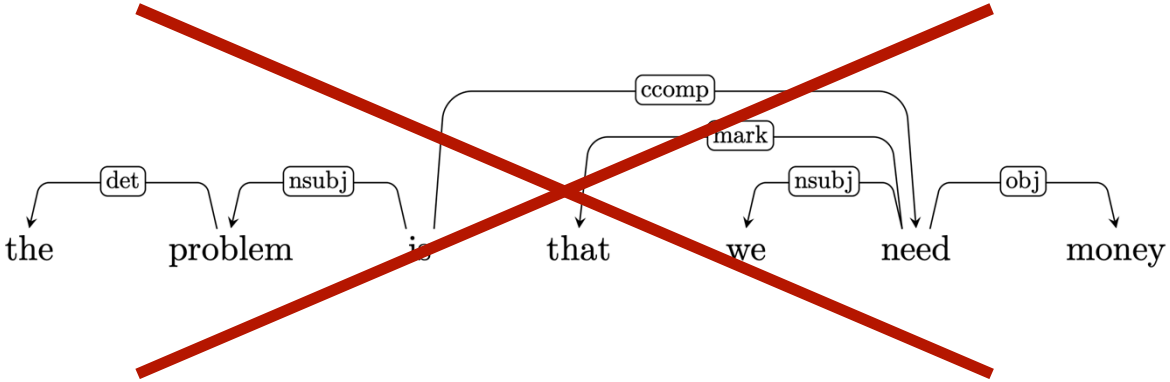
Change	Explanation
CLARIFICATION	Resolution of an ambiguity in the guidelines (without contradicting any previous version).
AMENDMENT	Change that negates a policy articulated in an earlier version of the guidelines.
VALIDATION	Important change (clarification or amendment) that will be checked by the UD validator.

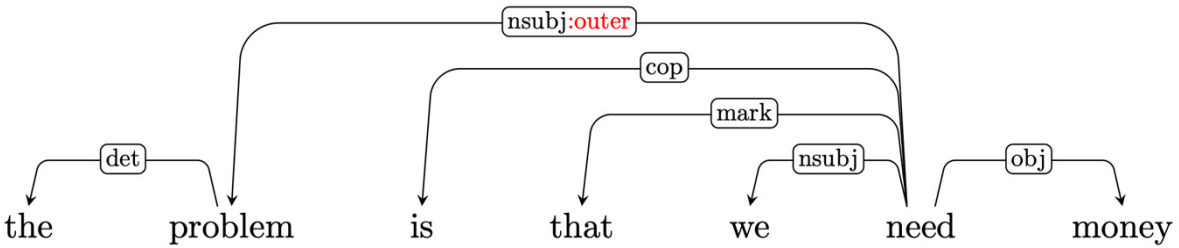
Guidelines remain at UDv2 as long as the inventories of universal part-of-speech tags and syntactic relations remain the same.

https://universaldependencies.org/changes.html

Recent Changes





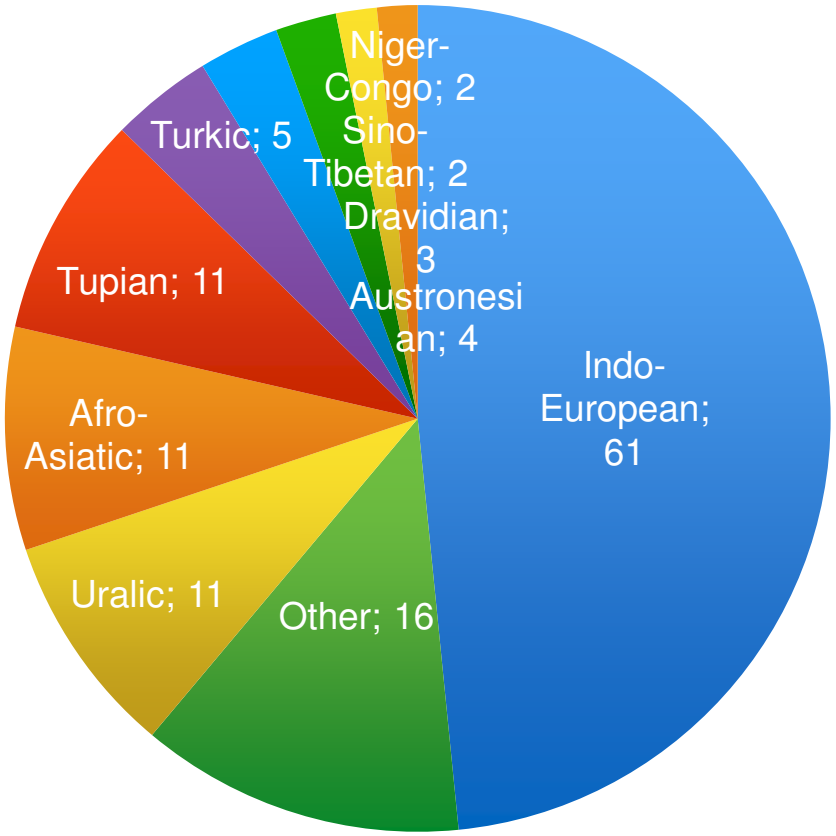


Nonverbal predication with clausal predicates

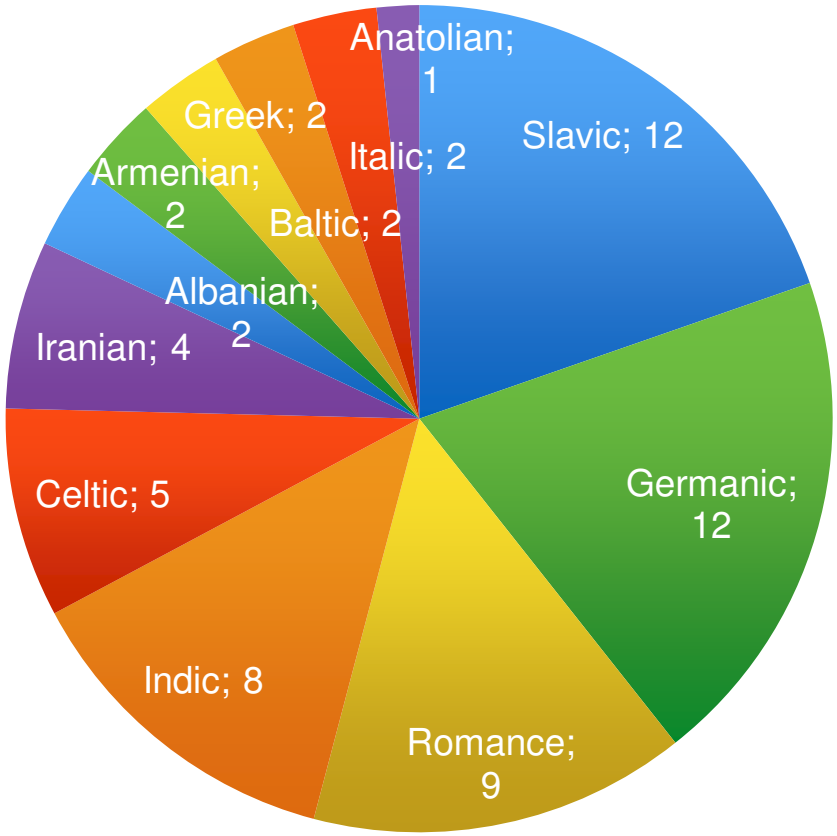
Treebank Status

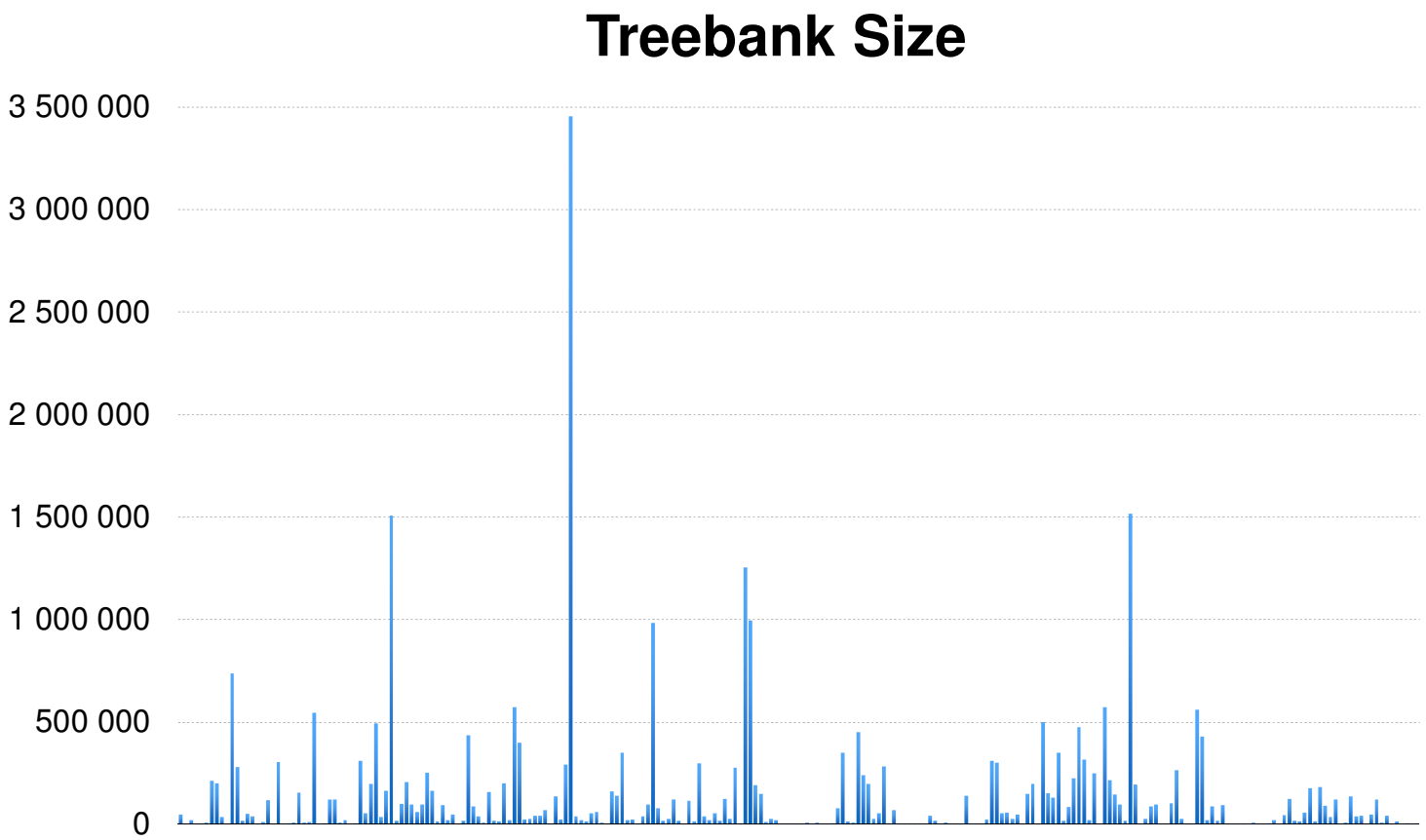
Status	Explanation
SAPLING	New treebank not yet released.
VALID	1. Passes all validation tests. 2. Contains at least 20 sentences and 100 words.
LEGACY	1. Fails at least one validation test. 2. Oldest validation error < 3 years old.
NEGLECTED	1. Fails at least one validation test. 2. Oldest validation error > 3 and < 4 years old.
INVALID	1. Fails at least one validation test. 2. Oldest validation error > 4 years old.

Language Family



Indo-European




Treebank Size
3 500 000
3 000 000
2 500 000
2 000 000
1 500 000
1 000 000
500 000
0

Extensions

- Enhanced Universal Dependencies
- Deep Universal Dependencies
- Universal Proposition Banks
- CorefUD
- Named Entities
- Surface Syntactic UD (SUD)
- PARSEME

Universal PropBank

PARSEME

Thanks to all UD contributors!

Daniel Zeman, Joakim Nivre, Mitchell Abrams, Elia Ackermann, Noëmi Aepli, Hamid Aghaei, Željko Agić, Amir Ahmadi, Lars Ahrenberg, Chika Kennedy Ajede, Salih Furkan Akkurt, Gabrielė Aleksandravičiūtė, Ika Alfina, Avner Algom, Chiara Alzetta, Erik Andersen, Lene Antonsen, Katya Aplonova, Angelina Aquino, Carolina Aragon, Glyd Aranes, Maria Jesus Aranzabe, Bilge Nas Arıcan, Þórunn Arnardóttir, Gashaw Arutie, Jessica Naraiswari Arwidarasti, Masayuki Asahara, Katla Ásgeirsdóttir, Deniz Baran Aslan, Cengiz Asmazoğlu, Luma Ateyah, Furkan Atmaca, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Elena Badmaeva, Keerthana Balasubramani, Miguel Ballesteros, Esha Banerjee, Sebastian Bank, Verginica Barbu Mititelu, Starkaður Barkarson, Rodolfo Basile, Victoria Basmov, Colin Batchelor, John Bauer, Seyyit Talha Bedir, Juan Belieni, Kepa Bengoetxea, Yifat Ben Moshe, Gözde Berk, Yevgeni Berzak, Irshad Ahmad Bhat, Riyaz Ahmad Bhat, Erica Biagetti, Eckhard Bick, Agnė Bielinskienė, Kristín Bjarnadóttir, Rogier Blokland, Victoria Bobicev, Loïc Boizou, Emanuel Borges Völker, Carl Börstell, Cristina Bosco, Gosse Bouma, Sam Bowman, Adriane Boyd, Anouck Braggaar, Kristina Brokaitė, Aljoscha Burchardt, Marie Candito, Bernard Caron, Gauthier Caron, Lauren Cassidy, Maria Clara Castro, Tatiana Cavalcanti, Gülşen Cebiroğlu Eryiğit, Flavio Massimiliano Cecchini, Giuseppe G. A. Celano, Slavomír Čéplö, Neslihan Cesur, Savas Cetin, Özlem Çetinoğlu, Fabricio Chalub, Liyanage Chamila, Shweta Chauhan, Ethan Chi, Taishi Chika, Yongseok Cho, Jinho Choi, Jayeol Chun, Juyeon Chung, Alessandra T. Cignarella, Silvie Cinková, Aurélie Collomb, Çağrı Çöltekin, Miriam Connor, Daniela Corbetta, Marine Courtin, Mihaela Cristescu, Philemon Daniel, Elizabeth Davidson, Leonel Figueiredo de Alencar, Mathieu Dehouck, Martina de Laurentiis, Marie-Catherine de Marneffe, Valeria de Paiva, Mehmet Oguz Derin, Elvis de Souza, Arantza Diaz de Ilarraza, Carly Dickerson, Arawinda Dinakaramani, Elisa Di Nuovo, Bamba Dione, Peter Dirix, Kaja Dobrovoljc, Timothy Dozat, Kira Droganova, Puneet Dwivedi, Christian Ebert, Hanne Eckhoff, Sandra Eiche, Marhaba Eli, Ali Elkahky, Binyam Ephrem, Olga Erina, Tomaž Erjavec, Aline Etienne, Wograine Evelyn, Sidney Facundes, Richárd Farkas, Federica Favero, Jannatul Ferdaousi, Marília Fernanda, Hector Fernandez Alcalde, Jennifer Foster, Cláudia Freitas, Kazunori Fujita, Katarína Gajdošová, Daniel Galbraith, Federica Gamba, Marcos Garcia, Moa Gärdenfors, Sebastian Garza, Fabrício Ferraz Gerardi, Kim Gerdes, Filip Ginter, Gustavo Godoy, 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, Hinrik Hafsteinsson, Jan Hajič, Jan Hajič jr., Mika Hämäläinen, Linh Hà Mỹ, Na-Rae Han, Muhammad Yudistira Hanifmuti, Takahiro Harada, Sam Hardwick, Kim Harris, Dag Haug, Johannes Heinecke, Oliver Hellwig, Felix Hennig, Barbora Hladká, Jaroslava Hlaváčová, Florinel Hociung, Petter Hohle, Marivel Huerta Mendez, Jena Hwang, Takumi Ikeda, Anton Karl Ingason, Radu Ion, Elena Irimia, Ọlájídé Ishola, Artan Islamaj, Kaoru Ito, Siratun Jannat, Tomáš Jelínek, Apoorva Jha, Katharine Jiang, Anders Johannsen, Hildur Jónsdóttir, Fredrik Jørgensen, Markus Juutinen, Hüner Kaşıkara, Andre Kaasen, Nadezhda Kabaeva, Sylvain Kahane, Hirosh Kanayama, Jenna Kanerva, Neslihan Kara, Ritván Karahóǧa, Boris Katz, Tolga Kayadelen, Sarveswaran Kengatharaiyer, Jessica Kenney, Václava Kettnerová, Jesse Kirchner, Elena Klementieva, Elena Klyachko, Arne Köhn, Abdullatif Köksal, Kamil Kopacewicz, Timo Korkiakangas, Mehmet Köse, Alexey Koshevoy, Natalia Kotsyba, Jolanta Kovalevskaitė, Simon Krek, Parameswari Krishnamurthy, Sandra Kübler, Adrian Kuqi, Oğuzhan Kuyrukçu, Aslı Kuzgun, Sookyoung Kwak, Veronika Laippala, Lucia Lam, Lorenzo Lambertino, 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, Yixuan Li, Yuan Li, KyungTae Lim, Bruna Lima Padovani, Krister Lindén, Nikola Ljubešić, Olga Loginova, Stefano Lusito, Andry Luthfi, Mikko Luukko, Olga Lyashevskaya, Teresa Lynn, Vivien Macketanz, Menel Mahamdi, Jean Maillard, Ilya Makarchuk, Aibek Makazhanov, Michael Mandl, Christopher Manning, Ruli Manurung, Büşra Marşan, Cătălina Mărănduc, David Mareček, Katrin Marheinecke, Stella Markantonatou, Héctor Martínez Alonso, Lorena Martín Rodríguez, André Martins, Jan Mašek, Hiroshi Matsuda, Yuji Matsumoto, Alessandro Mazzei, Ryan McDonald, Sarah McGuinness, Gustavo Mendonça, Tatiana Merzhevich, Niko Miekka, Karina Mischenkova, Margarita Misirpashayeva, Anna Missilä, Cătălin Mititelu, Maria Mitrofan, Yusuke Miyao, AmirHossein Mojiri Foroushani, Judit Molnár, Amirsaeid Moloodi, Simonetta Montemagni, Amir More, Laura Moreno Romero, Giovanni Moretti, Keiko Sophie Mori, Shinsuke Mori, Tomohiko Morioka, Shigeki Moro, Bjartur Mortensen, Bohdan Moskalevskyi, Kadri Muischnek, Robert Munro, Yugo Murawaki, Kaili Müürisep, Pinkey Nainwani, Mariam Nakhlé, Juan Ignacio Navarro Horñiacek, Anna Nedoluzhko, Gunta Nešpore-Bērzkalne, Manuela Nevaci, Lương Nguyễn Thị, Huyền Nguyễn Thị Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Alireza Nourian, Hanna Nurmi, Stina Ojala, Atul Kr. Ojha, Hulda Óladóttir, Adédayọ̀ Olúòkun, Mai Omura, Emeka Onwuegbuzia, Noam Ordan, Petya Osenova, Robert Östling, Lilja Øvrelid, Şaziye Betül Özateş, Betül Özateş, Merve Özçelik, Arzucan Özgür, Balkız Öztürk Başaran, Teresa Paccosi, Alessio Palmero Aprosio, Anastasia Panova, Hyunji Hayley Park, Niko Partanen, Elena Pascual, Marco Passarotti, Agnieszka Patejuk, Guilherme Paulino-Passos, Giulia Pedonese, Angelika Peljak-Łapińska, Siyao Peng, Cenel-Augusto Perez, Natalia Perkova, Guy Perrier, Slav Petrov, Daria Petrova, Andrea Peverelli, Jason Phelan, Jussi Piitulainen, Rodrigo Pintucci, Tommi A Pirinen, Emily Pitler, Magdalena Plamada, Barbara Plank, Thierry Poibeau, Larisa Ponomareva, Martin Popel, Lauma Pretkalniņa, Sophie Prévost, Prokopis Prokopidis, Adam Przepiórkowski, Robert Pugh, Tiina Puolakainen, Sampo Pyysalo, Peng Qi, Andriela Rääbis, Alexandre Rademaker, Mizanur Rahoman, Taraka Rama, Loganathan Ramasamy, Carlos Ramisch, Fam Rashel, Mohammad Sadegh Rasooli, Vinit Ravishankar, Livy Real, Petru Rebeja, Siva Reddy, Mathilde Regnault, Georg Rehm, Ivan Riabov, Michael Rießler, Erika Rimkutė, Larissa Rinaldi, Laura Rituma, Putri Rizqiyah, Luisa Rocha, Eiríkur Rögnvaldsson, Ivan Roksandic, Mykhailo Romanenko, Rudolf Rosa, Valentin Roșca, Davide Rovati, Ben Rozonoyer, Olga Rudina, Jack Rueter, Kristján Rúnarsson, Shoval Sadde, Pegah Safari, Benoît Sagot, Aleksi Sahala, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Stephanie Samson, Manuela Sanguinetti, Ezgi Sanıyar, Dage Särg, Marta Sartor, Mitsuya Sasaki, Baiba Saulīte, Yanin Sawanakunanon, Shefali Saxena, Kevin Scannell, Salvatore Scarlata, Nathan Schneider, Sebastian Schuster, Lane Schwartz, Djamé Seddah, Wolfgang Seeker, Mojgan Seraji, Syeda Shahzadi, Mo Shen, Atsuko Shimada, Hiroyuki Shirasu, Yana Shishkina, Muh Shohibussirri, Maria Shvedova, Janine Siewert, Einar Freyr Sigurðsson, João Ricardo Silva, Aline Silveira, Natalia Silveira, Maria Simi, Radu Simionescu, Katalin Simkó, Mária Šimková, Haukur Barri Símonarson, Kiril Simov, Dmitri Sitchinava, Maria Skachedubova, Aaron Smith, Isabela Soares-Bastos, Barbara Sonnenhauser, Shafi Sourov, Carolyn Spadine, Rachele Sprugnoli, Vivian Stamou, Steinþór Steingrímsson, Antonio Stella, Abishek Stephen, Milan Straka, Emmett Strickland, Jana Strnadová, Alane Suhr, Yogi Lesmana Sulestio, Umut Sulubacak, Shingo Suzuki, Daniel Swanson, Zsolt Szántó, Chihiro Taguchi, Dima Taji, Yuta Takahashi, Fabio Tamburini, Mary Ann C. Tan, Takaaki Tanaka, Dipta Tanaya, Mirko Tavoni, Samson Tella, Isabelle Tellier, Marinella Testori, Guillaume Thomas, Sara Tonelli, Liisi Torga, Marsida Toska, Trond Trosterud, Anna Trukhina, Reut Tsarfaty, Utku Türk, Francis Tyers, Sveinbjörn Þórðarson, Vilhjálmur Þorsteinsson, Sumire Uematsu, Roman Untilov, Zdeňka Urešová, Larraitz Uria, Hans Uszkoreit, Andrius Utka, Elena Vagnoni, Sowmya Vajjala, Rob van der Goot, Martine Vanhove, Daniel van Niekerk, Gertjan van Noord, Viktor Varga, Uliana Vedenina, Giulia Venturi, Eric Villemonte de la Clergerie, Veronika Vincze, Natalia Vlasova, Aya Wakasa, Joel C. Wallenberg, Lars Wallin, Abigail Walsh, Jing Xian Wang, Jonathan North Washington, Maximilan Wendt, Paul Widmer, Shira Wigderson, Sri Hartati Wijono, Vanessa Berwanger Wille, Seyi Williams, Mats Wirén, Christian Wittern, Tsegay Woldemariam, Tak-sum Wong, Alina Wróblewska, Mary Yako, Kayo Yamashita, Naoki Yamazaki, Chunxiao Yan, Koichi Yasuoka, Marat M. Yavrumyan, Arife Betül Yenice, Olcay Taner Yıldız, Zhuoran Yu, Arlisa Yuliawati, Zdeněk Žabokrtský, Shorouq Zahra, Amir Zeldes, He Zhou, Hanzhi Zhu, Anna Zhuravleva, Rayan Ziane