

# Introduction to Natural Language Processing

a course taught as B4M36NLP at Open Informatics



by members of the Institute of Formal and Applied Linguistics



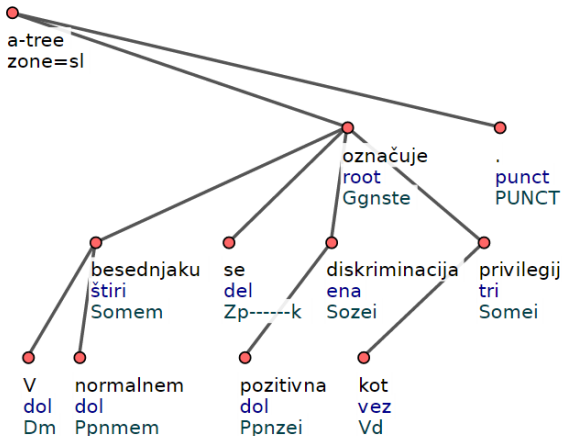
FACULTY  
OF MATHEMATICS  
AND PHYSICS  
Charles University

Today: **Week 6, lab**  
Today's topic: **Universal Dependencies**  
Today's teacher: **Daniel Zeman**

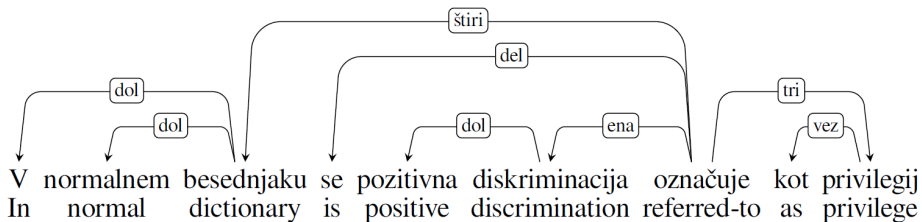
E-mail: [zeman@ufal.mff.cuni.cz](mailto:zeman@ufal.mff.cuni.cz)

WWW: <http://ufal.mff.cuni.cz/daniel-zeman>

# Dependency Treebanks

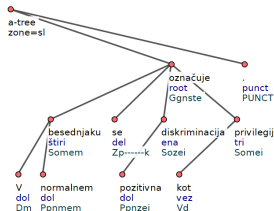


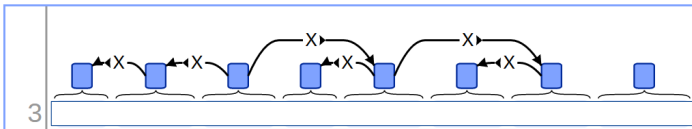
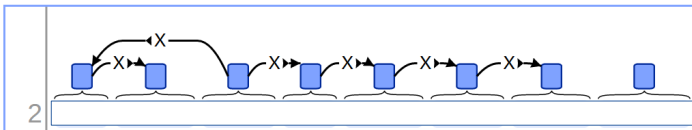
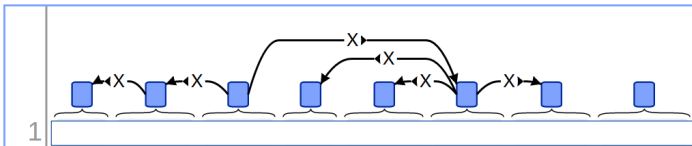
# Dependency Treebanks

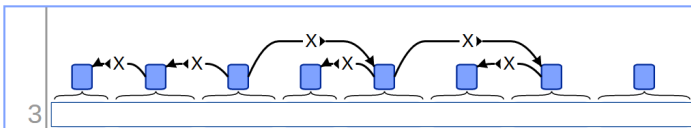
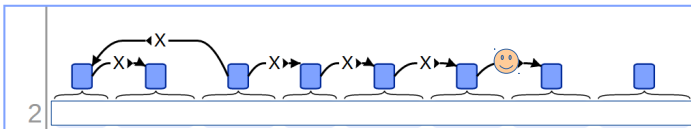
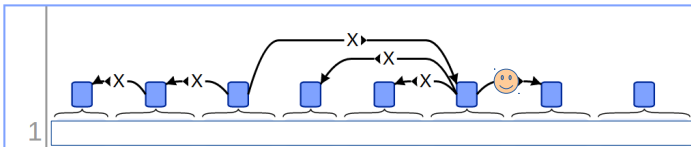


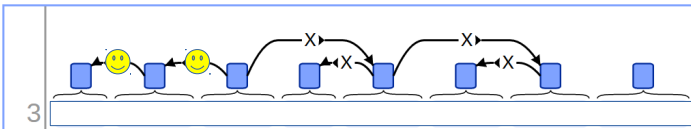
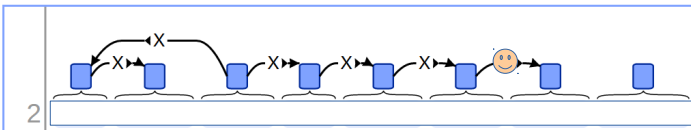
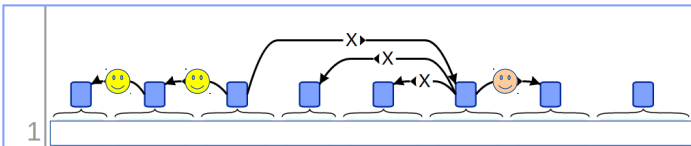
# Why?

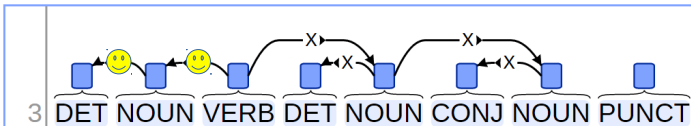
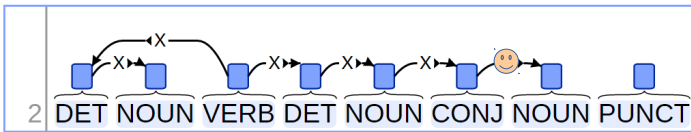
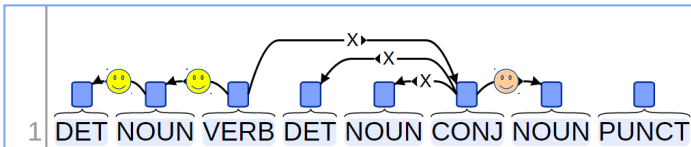
- Linguistic research
  - Corpus query
- Training tools (parsers) for NLP
  - Downstream applications



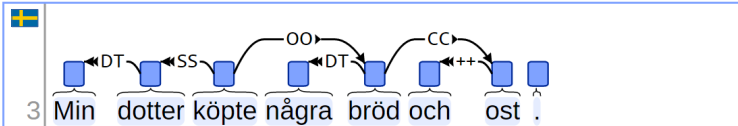
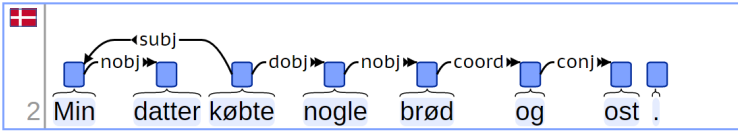
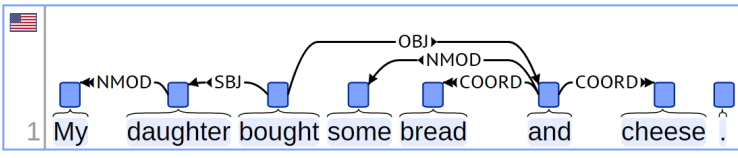












# Universal Dependencies

<http://universaldependencies.org/>



# Universal Dependencies

<http://universaldependencies.org/>

Stanford  
Dependencies

# Universal Dependencies

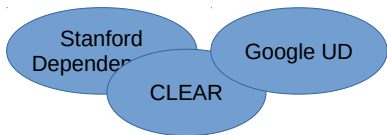
<http://universaldependencies.org/>

Stanford  
Depend

CLEAR

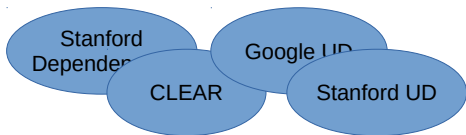
# Universal Dependencies

<http://universaldependencies.org/>



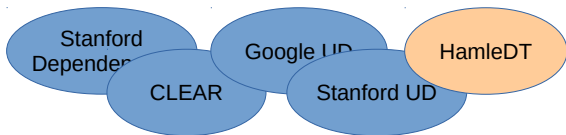
# Universal Dependencies

<http://universaldependencies.org/>



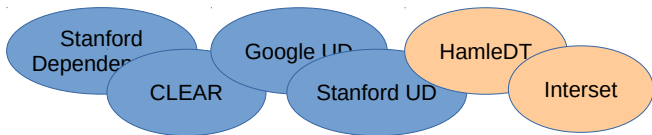
# Universal Dependencies

<http://universaldependencies.org/>



# Universal Dependencies

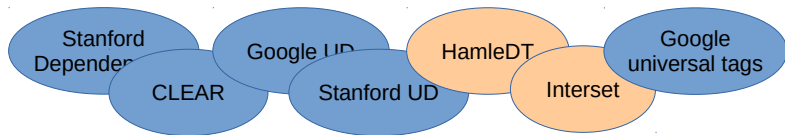
<http://universaldependencies.org/>





# Universal Dependencies

<http://universaldependencies.org/>



# Universal Dependencies

<http://universaldependencies.org/>

Universal Dependencies

# Universal Dependencies

<http://universaldependencies.org/>

- Milestones:

- 2014-04: EACL Göteborg, kick-off meeting
- 2014-10: UD guidelines version 1
- 2015-01: released 10 treebanks of **10** languages (UD 1.0)
- 2015-05: released 19 treebanks of **18** languages (UD 1.1)
- 2015-11: released 37 treebanks of **33** languages (UD 1.2)
- 2016-05: released 54 treebanks of **40** languages (UD 1.3)
- 2016-11: UD release 1.4, ~7 new languages
- 2016 fall: UD **guidelines version 2**



# Goals and Requirements

- Cross-linguistically consistent grammatical annotation



# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP



# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP
- Based on common usage and existing de facto standards

# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP
- Based on common usage and existing de facto standards
- Caveats:
  - Not a new linguistic theory –  
but linguistically informed and relevant

# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP
- Based on common usage and existing de facto standards
- Caveats:
  - Not a new linguistic theory –  
but linguistically informed and relevant
  - Not an ideal parsing representation –  
but useful for comparative evaluation



# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP
- Based on common usage and existing de facto standards
- Caveats:
  - Not a new linguistic theory –  
but linguistically informed and relevant
  - Not an ideal parsing representation –  
but useful for comparative evaluation
  - Not the ultimate annotation scheme –  
but a lightweight lingua franca

# Goals and Requirements

- Cross-linguistically consistent grammatical annotation
- Support multilingual research and development in NLP
- Based on common usage and existing de facto standards
- Caveats:
  - Not a new linguistic theory – but linguistically informed and relevant
  - Not an ideal parsing representation – but useful for comparative evaluation
  - Not the ultimate annotation scheme but a lightweight **lingua franca**

Not  
“Universal”  
in the strictly  
typological  
sense!

# 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

# 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
- Recoverability
  - Transparent mapping from input text to word segmentation

# Golden Rules

- Maximize parallelism
  - Don't annotate the same thing in different ways
  - Don't make different things look the same

# Golden Rules

- Maximize parallelism
  - Don't annotate the same thing in different ways
  - Don't make different things look the same
- But don't overdo it
  - Don't annotate things that are not there
  - Balance: is it still the same thing?
  - Allow **language-specific** extensions

# Morphology

Některé dívky si nicméně pochvalovaly zmrzlinu .





# Morphology

Některé	dívky	si	nicméně	pochvalovaly	zmrzlinu	.
některý	dívka	se	nicméně	pochvalovat	zmrzlina	.

- Lemma representing the semantic content of the word

# Morphology

Některé	dívky	si	nicméně	pochvalovaly	zmrzlinu	.
některý	dívka	se	nicméně	pochvalovat	zmrzlina	.
DET	NOUN	PRON	CONJ	VERB	NOUN	PUNCT

- Lemma representing the semantic content of the word
- Part-of-speech tag representing the abstract lexical category associated with the word

# Morphology

Některé	dívky	si	nicméně	pochvalovaly	zmrzlinu	.
některý	dívka	se	nicméně	pochvalovat	zmrzlina	.
DET	NOUN	PRON	CONJ	VERB	NOUN	PUNCT
PronType=Ind Gender=Fem Number=Plur Case=Nom	Gender=Fem Number=Plur Case=Nom	PronType=Prs Reflex=Yes Case=Dat		VerbForm=Part Tense=Past Voice=Act Aspect=Imp Gender=Fem Number=Plur	Gender=Fem Number=Sing Case=Acc	

- Lemma representing the semantic content of the word
- Part-of-speech tag representing the abstract lexical category associated with the word
- Features representing lexical and grammatical properties associated with the lemma or the particular word form

# Part-of-Speech Tags

Open	Closed	Other
ADJ	ADP	PUNCT
ADV	AUX	SYM
INTJ	CONJ	X
NOUN	DET	
PROPN	NUM	
VERB	PART	
	PRON	
	SCONJ	

- Taxonomy of 17 universal part-of-speech tags, based on the Google Universal Tagset (Petrov et al., 2012)
- All languages use the same inventory, but not all tags have to be used by all languages

# Features

Lexical	Inflectional / Nominal	Inflectional / Verbal
PronType	Gender	VerbForm
NumType	Animacy	Mood
Poss	Number	Tense
Reflex	Case	Aspect
	Definite	Voice
	Degree	Person
		Negative

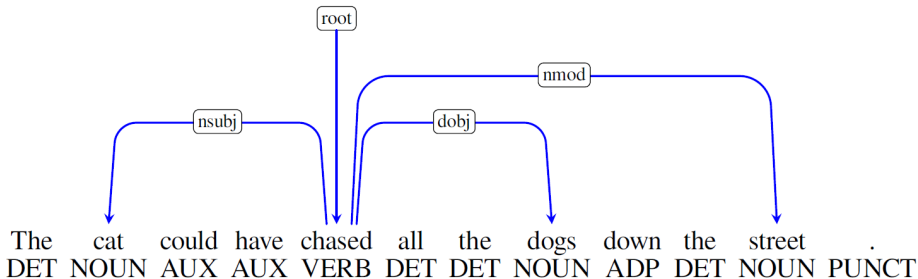
- Standardized inventory of morphological features, based on Interset (Zeman, 2008)
- Languages select relevant features and can add language-specific features or values with documentation

# Syntax

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

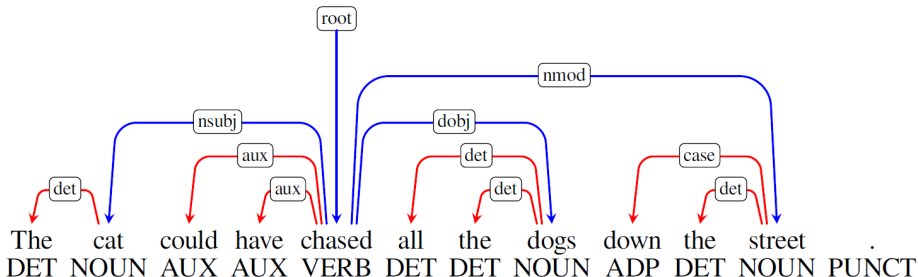


# Syntax



- Content words are related by dependency relations

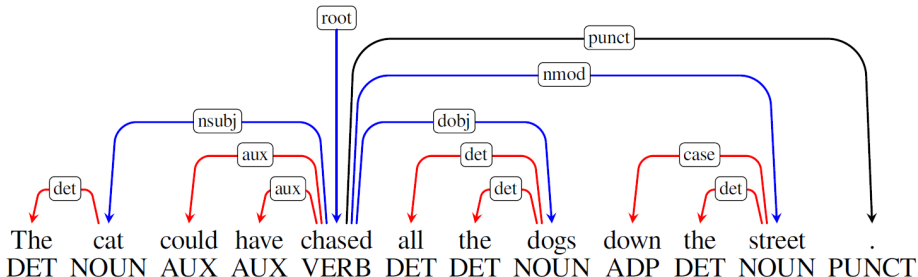
# Syntax



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

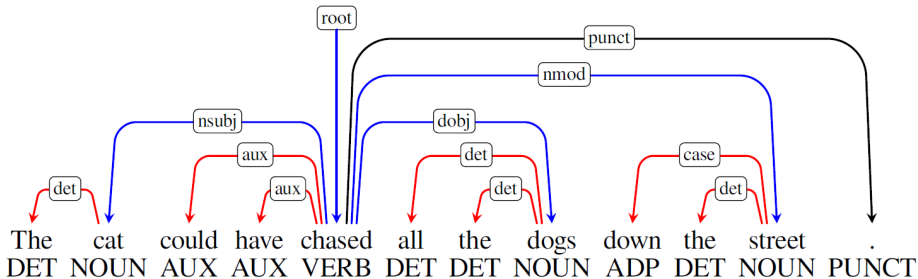


# Syntax



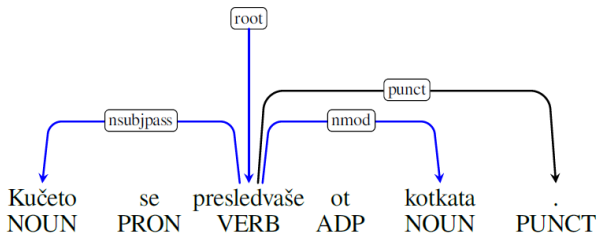
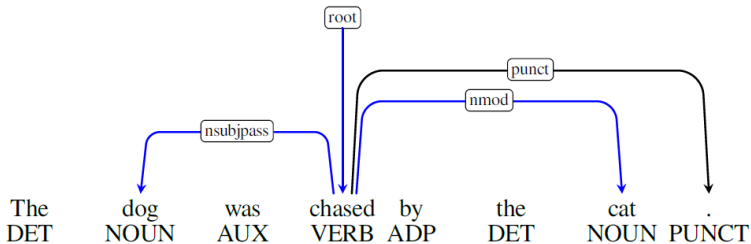
- Content words are related by dependency relations
- Function words attach to closest content word
- Punctuation attach to head of phrase or clause

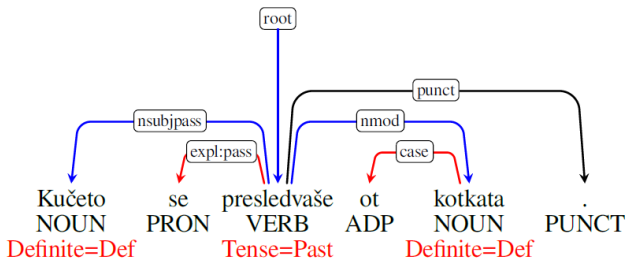
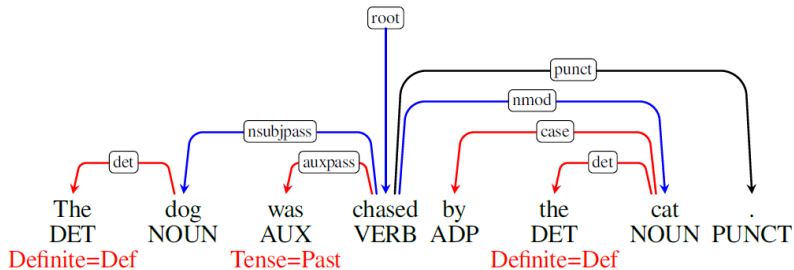
# Syntax

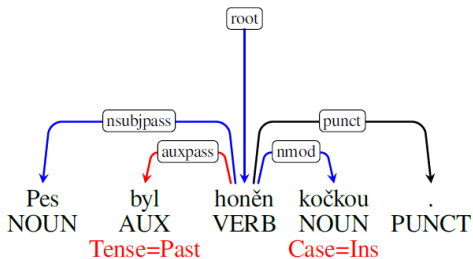
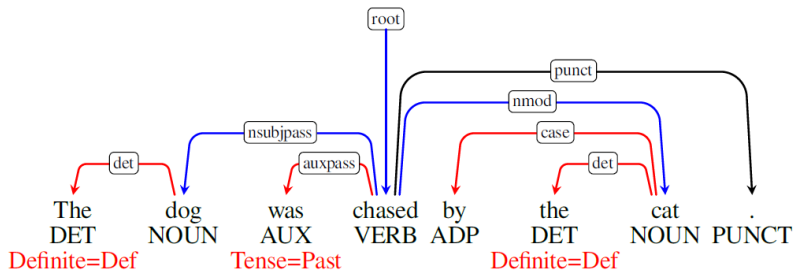


- Content words are related by dependency relations
- Function words attach to closest content word
- Punctuation attach to head of phrase

Not  
"dependency"  
in the strictly  
syntactic  
sense!







# Dependency Relations

- Taxonomy of 40 universal grammatical relations, broadly attested in language typology (de Marneffe et al., 2014)
  - Language-specific **subtypes** may be added

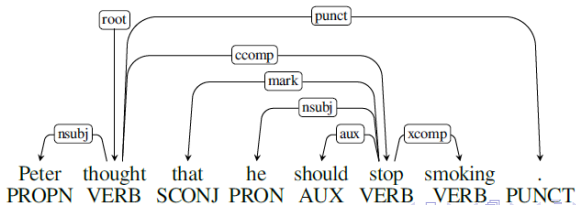
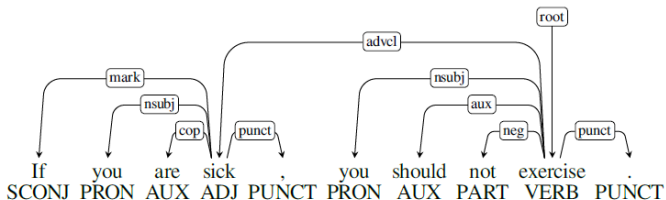
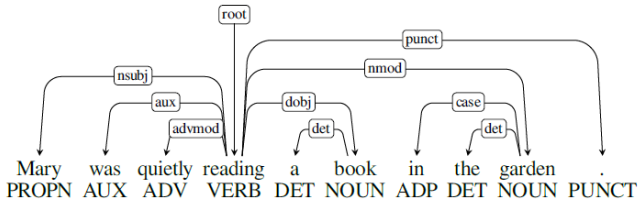
# Dependency Relations

- Taxonomy of 40 universal grammatical relations, broadly attested in language typology (de Marneffe et al., 2014)
  - Language-specific **subtypes** may be added
- Organizing principles
  - Three types of structures: nominals, clauses, modifiers
  - **Core** arguments vs. other dependents (**not** arguments vs. adjuncts)

# Dependents of Clausal Predicates

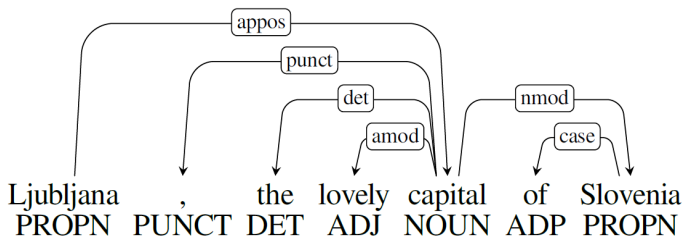
	Nominal	Clausal	Other
Core	nsubj nsubjpass dobj iobj	csubj csubjpass ccomp xcomp	
Non-Core	nmod vocative discourse expl	advcl	advmod neg aux auxpass cop mark punct



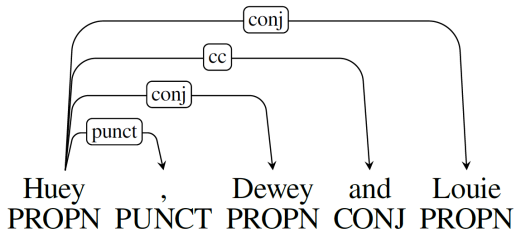
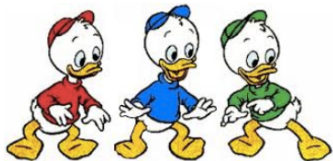


# Dependents of Nominals

Nominal	Clausal	Other
nmod appos nummod	acl	amod det neg case



# “Stanford-style” Coordination



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

# Multiword Expressions

Relation	Examples
mwe	<i>in spite of, as well as, ad hoc</i>
name	<i>Roger Bacon, New York</i>
compound	<i>phone book, four thousand, dress up</i>
goeswith	<i>notwith standing, with out</i>

- UD annotation does not permit “words with spaces”
  - Multiword expressions are analyzed using special relations
  - The **mwe**, **name** and **goeswith** relations are always head-initial
  - The **compound** relation reflects the internal structure

# Other Relations

Relation	Explanation
parataxis	Loosely linked clauses of same rank
list	Lists without syntactic structure
remnant	Orphans in ellipsis linked to parallel elements
reparandum	Disfluency linked to (speech) repair
foreign	Elements within opaque stretches of code switching
dep	Unspecified dependency
root	Syntactically independent element of clause/phrase

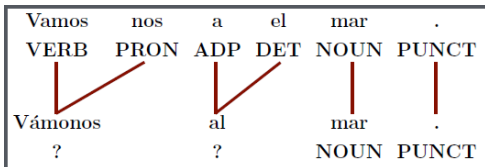
# Language-Specific Relations

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

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

# Word Segmentation

- Must be **reproducible** on new data
- Surface tokens vs. syntactic words
- Chinese, Vietnamese etc.: no clues, non-trivial algorithm
- Arabic, Tamil etc.: part of morphological analysis
- Spanish, German etc.: rather limited cases of contractions
- Others: only punctuation (low-level tokenization)



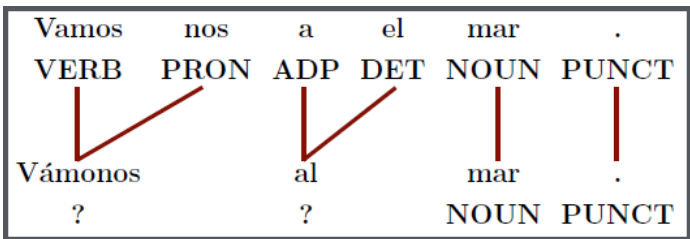
# Word Segmentation

- Fusions

- *al* = *a* + *el*
- *nañ* = *na* + *něj*

- Clitics

- *vámonos* = *vamos* + *nos*
- *изменяться* = *изменять* + *ся*
- *potrafilibyšmy*  
= *potrafilili* + *by* + *jestešmy*





# Where Are We Now?



# Where Are We Now?

- Two years of UD version 1
- 4 treebank releases (every 6 months)
- 54 (61) treebanks
- 40 (47) languages (over 50% world's population)
- Over 11M tokens; treebanks range from 1K to 1.5M
- Over 120 contributors
  - language group consistency SIGs
  - version 2 guidelines coming soon

# 47 Languages and Growing

	Ancient Greek-PROIEL	206K	LF	-	☰				
	Arabic	242K	LF	-	☰				
	Basque	121K	LF	📄	☰				
	Bulgarian	156K	LF	📄	☰	☑			
	Buryat	5K	L	-	☰	👤			
	Catalan	530K	LF	📄	☰	☑			
	Chinese	123K	F	📄	☰	☑			
	Coptic	4K	L	📄	☰	👤			
	Croatian	87K	LF	-	☰	☑			
	Czech	1,503K	LF	📄	☰	☑			
	Czech-CAC	493K	LF	📄	☰	☑			
	Czech-CLTT	35K	LF	📄	☰	☑			
	Danish	100K	LF	📄	☰	☑			
	Dutch	209K	LF	-	☰	☑			
	Dutch-LassySmall	98K	LF	-	☰	☑			
	English	254K	LF	📄	☰	👤			
	English-ESL	97K	L	📄	☰	👤			
	English-LinES	82K	L	📄	☰	☑			
	Estonian	234K	LF	-	☰	☑			
	Faroese	119K	F	-	☰	☑			
	Finnish	181K	LF	📄	☰	☑			
	Finnish-FTB	159K	LF	-	☰	☑			
	French	390K	LF	📄	☰	☑			
	Galician	138K	L	📄	☰	☑			
	German	293K	LF	-	☰	☑			
	Gothic	56K	LF	-	☰	☑			
	Greek	59K	LF	📄	☰	☑			
	Hebrew	115K	F	-	☰	☑			
	Hindi	351K	LF	-	☰	☑			
	Hungarian	42K	LF	📄	☰	👤			
	Indonesian	121K	-	-	☰	☑			
	Irish	23K	LF	-	☰	☑			
	Italian	252K	LF	📄	☰	☑			
	Japanese-KTC	267K	L	📄	☰	☑			
	Kazakh	4K	L	📄	☰	☑			
	Korean	-	-	-	☰	☑			
	Latin	47K	LF	-	☰	☑			
	Latin-ITT	291K	LF	-	☰	☑			
	Latin-PROIEL	165K	LF	-	☰	☑			
	Latvian	20K	LF	-	☰	☑			
	Norwegian	311K	LF	📄	☰	☑			
	Old Church Slavonic	57K	LF	-	☰	☑			
	Persian	151K	F	📄	☰	☑			
	Polish	83K	LF	-	☰	☑			
	Portuguese	209K	LF	-	☰	☑			
	Portuguese-BR	298K	F	-	☰	☑			
	Romanian	145K	LF	📄	☰	☑			
	Russian	99K	F	📄	☰	☑			
	Russian-SynTagRus	1,032K	LF	📄	☰	☑			
	Sanskrit	1K	LF	-	☰	☑			
	Slovenian	140K	LF	📄	☰	☑			
	Slovenian-SST	29K	LF	📄	☰	☑			
	Spanish	423K	LF	📄	☰	☑			
	Spanish-AnCora	547K	LF	📄	☰	☑			
	Swedish	96K	LF	📄	☰	☑			
	Swedish-LinES	79K	-	📄	☰	☑			
	Tamil	8K	LF	-	☰	☑			
	Turkish	56K	LF	📄	☰	☑			
	Ukrainian	-	-	-	☰	☑			
	Urdu	-	-	-	☰	☑			
	Uyghur	45K	F	-	☰	☑			
	Vietnamese	43K	L	-	☰	☑			



# Where Are We Going?

- UD guidelines **version 2** coming soon
- Consistency checking

# Common vocabulary is great ...

... because we finally  
understand each other ...





... almost

*Childs of you be  
vary acute!*














# Consistency Checking

- Automatic tests catch only a fraction
- Focus groups on
  - Romance, Germanic, **Slavic**, Uralic, Turkic languages



# Existing Slavic Treebanks

Language	Code	Treebank	Sent	Tok
Bulgarian 	[bg]	BulTreeBank	13,221	196K
Church Slavonic	[cu]	PROIEL	7,818	72K
Croatian 	[hr]	SETimes.HR	3,736	84K
Croatian 	[hr]	HOBS	4,626	117K
Czech 	[cs]	PDT	87,913	1504K
Czech 	[cs]	CAC	24,709	494K
Polish 	[pl]	IPI PAN	8,227	84K
Russian 	[ru]	SynTagRus	59,130	1033K
Russian 	[ru]	Google	5,030	99K
Slovak 	[sk]	SNK	63,238	994K
Slovenian 	[sl]	SSJ500K	27,829	500K
Slovenian 	[sl]	SST	3,188	29K





# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- Core arguments
- Reflexive pronouns (clitics)
- Auxiliary verbs and modal verbs
- Comparative constructions

# Pronouns and Determiners

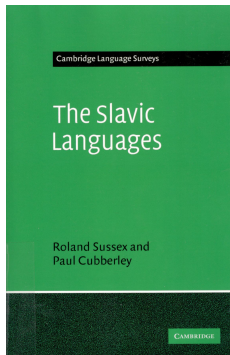
- English + Romance languages: **DET** = article or pronominal adjective (*this, which, every*)

# Pronouns and Determiners

- English + Romance languages: **DET** = article or pronominal adjective (*this, which, every*)
- We don't have this category! (Traditionally → PRON.)

# Pronouns and Determiners

- English + Romance languages: **DET** = article or pronominal adjective (*this, which, every*)
- We don't have this category! (Traditionally → PRON.)
- Some authors do recognize determiners in Slavic!



# Pronouns and Determiners

- English + Romance languages: **DET** = article or pronominal adjective (*this, which, every*)
- We don't have this category! (Traditionally → PRON.)
- We have the words (except for articles).

# Pronouns and Determiners

- English + Romance languages: **DET** = article or pronominal adjective (*this, which, every*)
- We don't have this category! (Traditionally → PRON.)
- We have the words (except for articles).
- Currently functional borderline (but ellipsis?)  
*This.DET car is expensive.*  
*This.PRON is expensive.*
- Less strict in UD v2.

# Pronouns Only

- Personal pronouns (including reflexives, but not possessives)
- Interrogative *who*, *what*
- Indefinite and negative derivatives
- Relative [cs] *jenž*
  - cs: *já, ty, on, my, vy, oni, se, kdo, co, někdo, něco, nikdo, nic*
  - sk: *ja, ty, on, my, vy, oni, sa, kto, čo, niekto, niečo, nikto, nič*
  - pl: *ja, ty, on, my, wy, oni, się, kto, co, ktoś, coś, nikt, nic*
  - ru: *я, ты, он, мы, вы, они, ся, кто, что, кто-нибудь, что-нибудь, никто, ничто*
  - sl: *jaz, ti, on, mi, vi, oni, se, kdo, kaj, nekdo, nekaj, nihče, nič*
  - hr: *ja, ti, on, mi, vi, oni, se, tko, što, neki, nešto, nitko, ništa*
  - bg: *аз, ти, ние, вие, се, кой, кое, някой, нещо, никой, нищо*
  - cu: *азъ, ты, мы, вы, и, са, кѣто, чѣто*



# Possessives: Determiners

- If they occur without a noun ... **ellipsis**

*Můj otec je starší. Tvůj má ale více zkušeností.  
My father is older. But yours is more experienced.*

- sl: *moj, tvoj, njegov, njen, najin, vajin, njun, naš, vaš, njihov, svoj*
- bg: *моѝ, твоѝ, негов, неин, наш, ваш, техен, своѝ*
- cs: *můj, tvůj, jeho, její, náš, váš, jejich, svůj*
- sk: *môj, tvoj, jeho, jej, náš, váš, ich, svoj*
- cu: *мои, твои, нашъ, вашъ, свои / ego, ея, ею, ихъ*





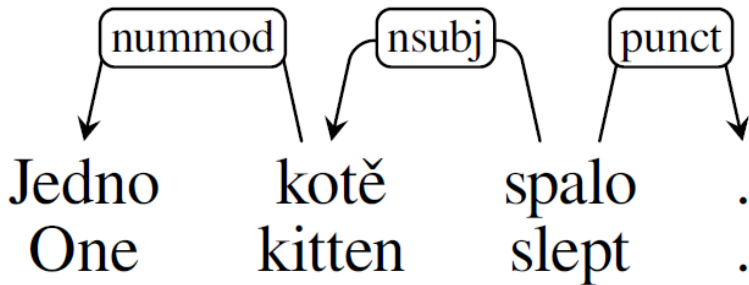
# Both Possible?

- Demonstratives
  - cs: *ten, to, tento, tenhle, tamten, ...*
  - sl: *ta, to, tisti, oni, takšen, ...*
- Adjectival interrogatives/relatives, indefinites, negatives
  - *jaký, který, čí, nějaký, některý, něčí, každý, žádný*
  - *všechen, všichni, všechno*
- Relative pronouns **cannot** be explained by **ellipsis!**
  - *Muž, kterého \*muže jsem vám představil.*
  - *The man, which \*man I introduced to you.*

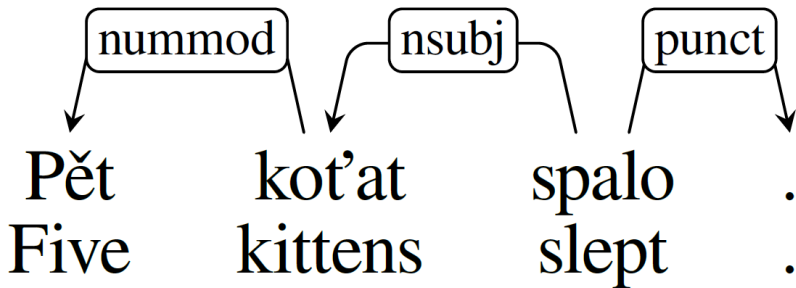
# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- Core arguments
- Reflexive pronouns (clitics)
- Auxiliary verbs and modal verbs
- Comparative constructions

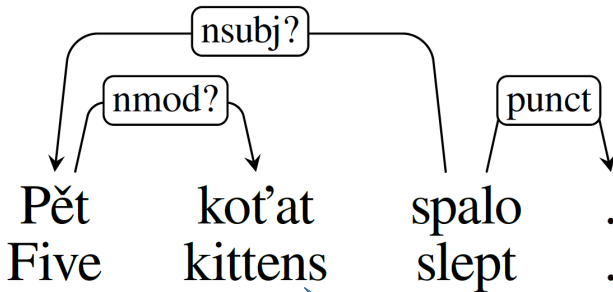
# Quantified Noun Phrase



# Quantified Noun Phrase

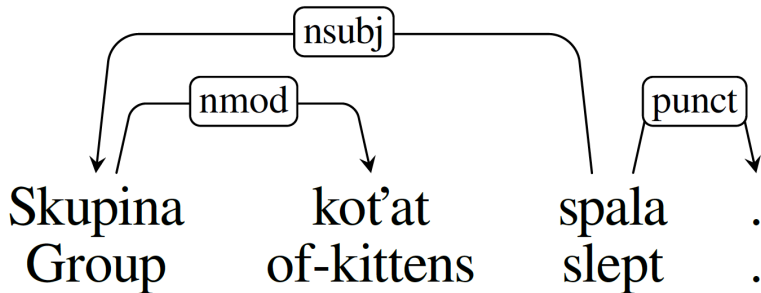


# Quantified Noun Phrase

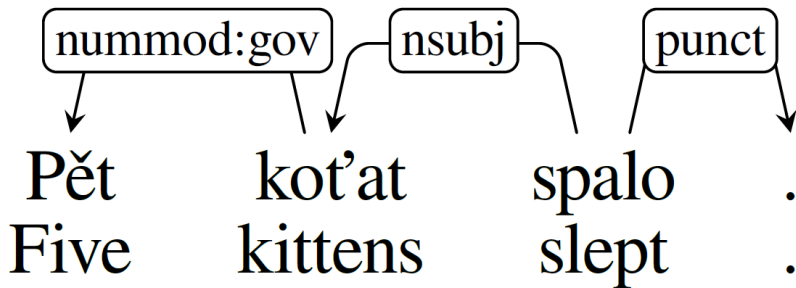


Genitive!

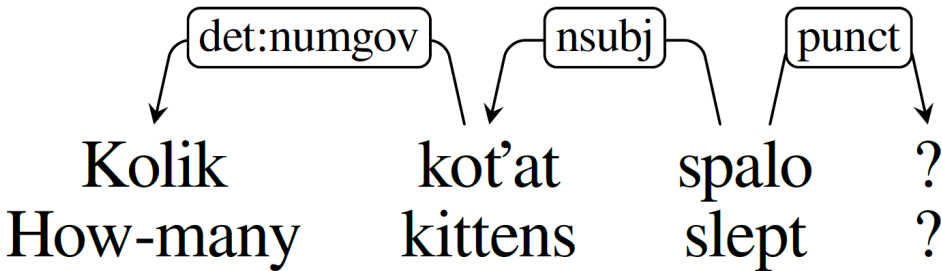
# Quantified Noun Phrase



# Quantified Noun Phrase



# Pronominal Quantifiers





# Language-Specific Labels

	<b>Numeric</b>	<b>Pronominal</b>
<b>Noun governs</b>	nummod	det:nummod
<b>Numeral governs</b>	nummod:gov	det:numgov

# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- Core arguments
- Reflexive pronouns (clitics)
- Auxiliary verbs and modal verbs
- Comparative constructions

# Verb Forms

- Conflicting terminologies in traditional grammars
- Participle ... verb or adjective?
- Converb ... verb or adverb?
  
- Tags and features apply to **individual words!**

# Verb Forms

- POS tags and features apply to individual words!
- *A ko **so se** leta 1942 **vračali**, ...*
  - past tense
- *... da ne **bi** v Atene **prišli** ...*
  - conditional mood
- *... v prihodnje ne **bodo vozili** zgolj les ...*
  - future tense

# Verb Forms

- POS tags and features apply to individual words!
- A ko *so se leta 1942 vračali*, ...

Present tense

- ... da ne *bi v Atene prišli* ...

Conditional mood

- ... v prihodnje ne *bodo vozili* zgolj les ...

- future tense

Future

# Verb Forms

- POS tags and features apply to individual words!

• A ko *so se* leta 1942 *vračali*, ...

Past???

Present  
tense

Participle

• ... da ne *bi* v Atene *prišli* ...

Conditional  
mood

Participle

• ... v prihodnje ne *bodo vozili* zgolj les ...

- future tense

Future

Participle

# Verb Forms

- *vračali, prišli, vozili*
- [cs] “active participle” / “past tense”
- [ru] “past tense” / “finite!”
  - Active participle is something else: *нарушивший*
- [bg] “participle + past (aorist) / imperfect” (two subtypes)
- [cu] “participle + resultative aspect” (lang-spec)
- “I-participle”
  - But that would be a language-specific verb form.

# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- **Core arguments**
- Reflexive pronouns (clitics)
- Auxiliary verbs and modal verbs
- Comparative constructions



# Core Arguments

- Easier cross-linguistically than argument-adjunct?
- **S**ubject of intransitive verb
- **A**gent of transitive verb
- **P**atient (direct object) of transitive verb
  
- Indirect object? Dative only?

# Core vs. Oblique Dependents

- **Core arguments:** what exactly is it?
- English:
  - *He gave **John** the book.* (iobj)
  - *He gave the book **to John**.* (nmod)
- Spanish:
  - *Dio el libro **a John**.* (iobj)
- Czech:
  - Every Obj is translated to dobj, regardless the case and the presence of preposition

# doj / iobj

- Not as easy as accusative vs. dative.
- Default: doj
- Heuristics for iobj
  - *Cením si vaší pomoci.* (Gen)  
I appreciate your help.
  - *Čelíme velkým problémům.* (Dat)  
We are facing big problems.
  - *Nedisponuje takovým rozpočtem.* (Ins)  
He does not have such budget.
  - *Učí mou dceru fyziku.* (2 × Acc)  
He teaches my daughter physics.

# All Slavic Treebanks Have Non-Accusative “Direct” Objects

- *podrobit se testu; odpovídají smlouvě; jednat s někým*
- *mówi o niej; używa wielkich słów*
- *от которых зависят; относятся к программам*
- *potrebuje informacij; slediti evropskim smernicam; ukvarjal se bom orožjem*
- *odriče se imuniteta; priključiti se naporima*
- *се характеризира с развитие; моля за внимание*

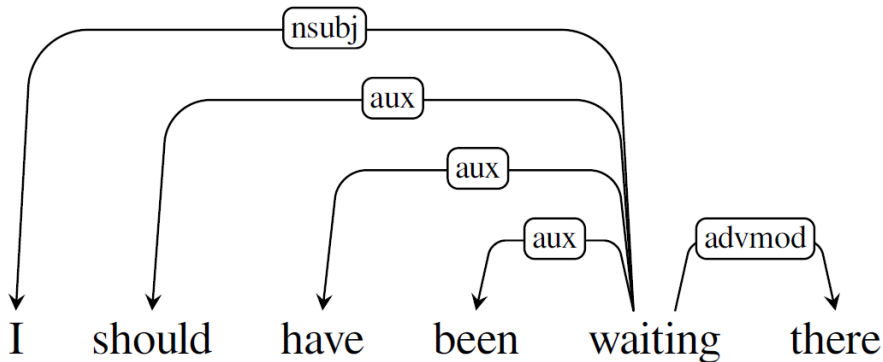
# Reflexive Pronouns

- Direct or indirect object (dobj, iobj):  
*Řízl se do prstu / Řízl ho do prstu.*
  - Including reciprocal usage:  
*Políbili se. / They kissed each other.*
- Inherently reflexive verbs: *smát se, bát se / laugh, fear*
  - **expl:pv** (pronominal verb; previously compound)
- Reflexive passive:  
*To se snadněji řekne než udělá. / That is easier said than done.*
  - **expl:pass** (previously auxpass: reflex)
- Impersonal construction (~ passive?):  
*Zde se mluví německy. / German is spoken here.*
  - **expl:impers**

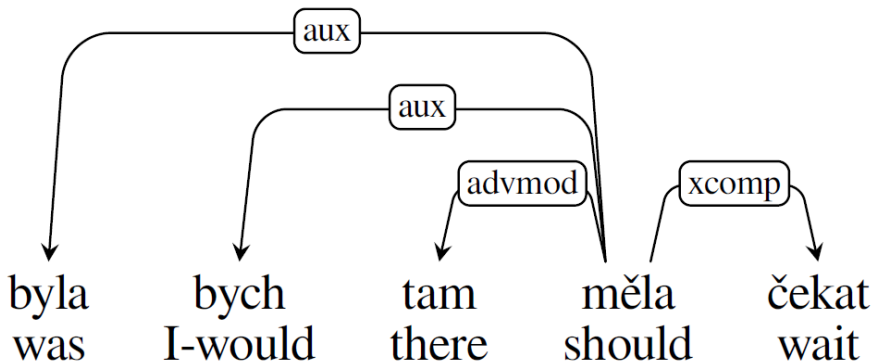
# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- Core arguments
- Reflexive pronouns (clitics)
- **Auxiliary verbs and modal verbs**
- Comparative constructions

# Modal Auxiliary in English

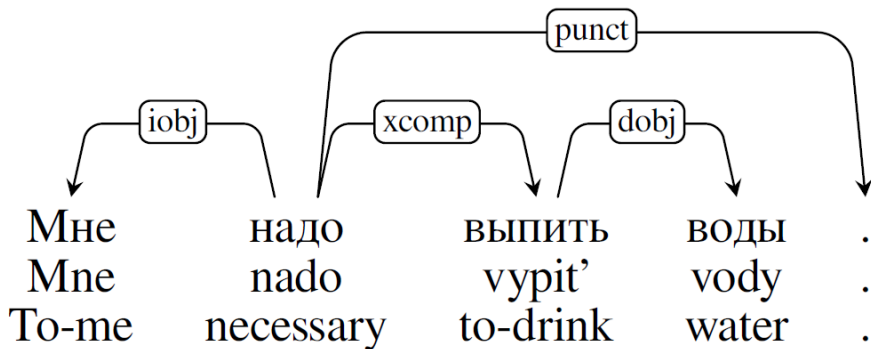


# Modal Verb in Czech

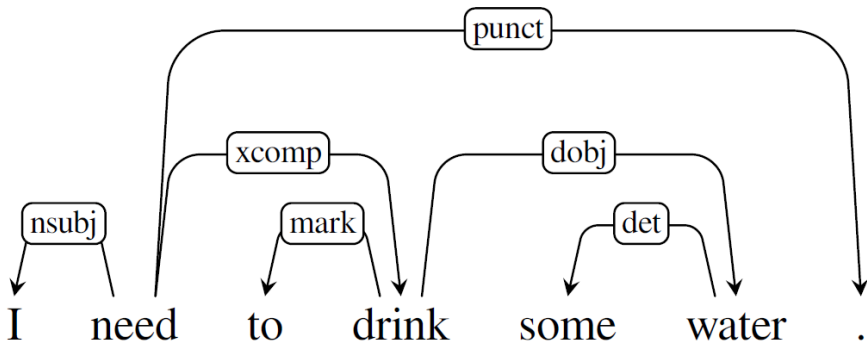




# Modal Adverb in Russian



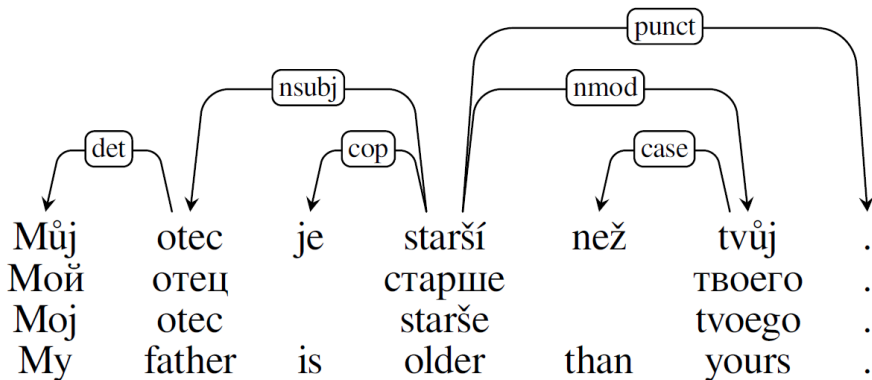
# Modal / Control Verb in English



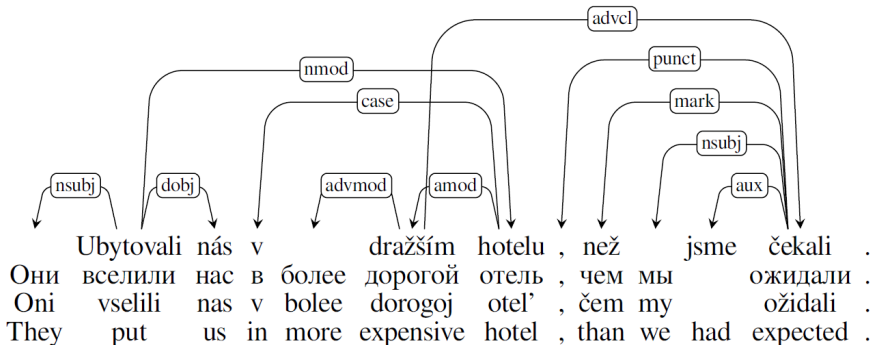
# Issues of Slavic Languages in UD

- Pronouns vs. determiners, numerals and quantifiers
- Attachment of cardinal numbers
- Verbs, participles, adjectives
- Core arguments
- Reflexive pronouns (clitics)
- Auxiliary verbs and modal verbs
- **Comparative constructions**

# Comparative Constructions



# Comparative Constructions



# Wrapping Up



# Wrapping Up

- UD has had a great start

# Wrapping Up

- UD has had a great start
- Still a long way to go.  
Consistency matters!



# Wrapping Up

- UD has had a great start
- Still a long way to go. Consistency matters!
- Get involved. It's fun!



Děkuji!  
Otázky?

Благодаря!  
Въпроси?

Đakujem!  
Otázky?

Благодаря!  
Въпроси?

Thank you!  
Questions?

Спасибо!  
Вопросы?

Dziękuję!  
Pytania?

Hvala!  
Vprašanja?

Hvala!  
Pitanja?



