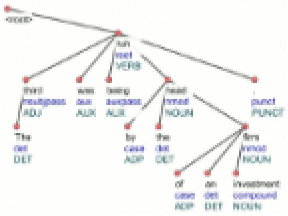
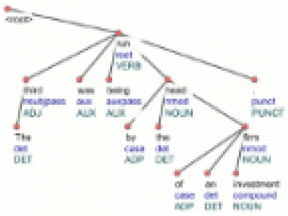


Universal Dependencies
UDPipe
Udapi

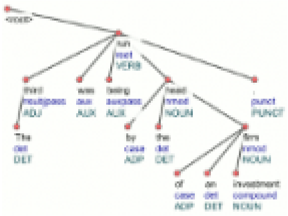
Martin Popel, Charles University, ÚFAL

TextLink training school, Prague, February 9, 2017



Universal Dependencies
UDPipe
Udapi

None of these three support discourse!



Universal Dependencies
UDPipe
Udapi

None of these three support discourse!

But

- Syntax and morphology are helpful when analyzing discourse.
- Universal Dependencies (UD) is a great multi-lingual resource
- and a source of inspiration wrt annotation guidelines, success.
- Udapi plans to support discourse, coreference, alignment,…

Universal Dependencies

Joakim Nivre, **Dan Zeman,** Filip Ginter, Sampo Pyysalo, Chris Manning, Marie-Catherine de Marneffe, Natalia Silveira, Slav Petrov, Ryan McDonald, Tim Dozat, Jan Hajič, Jinho Choi, Reut Tsarfaty, Yoav Goldberg, Simonetta Montemagni, Alessandro Lenci, Maria Simi, Cristina Bosco, Veronika Vincze, Richárd Farkas, Teresa Lynn, Jennifer Foster, Prokopis Prokopidis, Jenna Kanerva, Juha Kuokkala, Veronika Laippala, Krister Lindén, Anna Missilä, Hanna Nurmi, Jussi Piitulainen, Aaron Smith, Željko Agić, Nikola Ljubešić, Maria Jesus Aranzabe, Aitziber Atutxa, Iakes Goenaga, Koldo Gojenola, Anders Trærup Johannsen, Hèctor Martínez, Barbara Plank, Petya Osenova, Kiril Simov, Mojgan Seraji, Wolfgang Seeker, Fran Tyers, Aibek Makazhanov, Jon Washington, Çağrı Çöltekin, Arne Skjærholt, Lilja Øvrelid, Miguel Ballesteros, Elena Pascual, Giuseppe Celano, Marco Passarotti, Martin Popel, Christophe Onambélé, Dag Haug, Nizar Habash, Riyaz Ahmad, Verginica Mititelu, Catalina Mărănduc, Kaja Dobrovoljc, Tomaž Erjavec, Simon Krek, Yusuke Miyao, Shinsuke Mori, Takaaki Tanaka, Hiroshi Kanayama, Masayuki Asahara, Sumire Uematsu, Rob Voigt, …

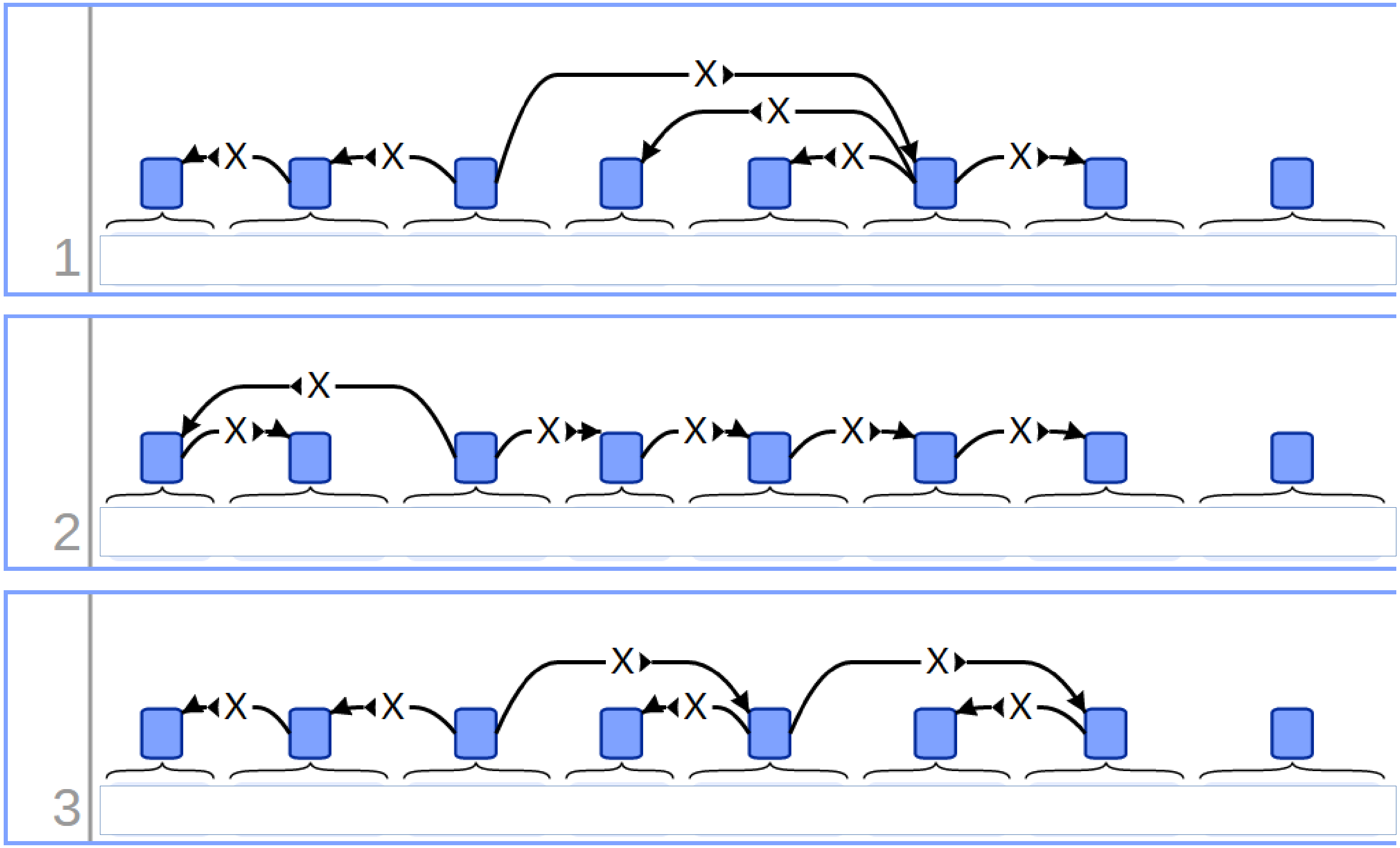
Introduction slides stolen from Joakim Nivre

Universal Dependencies

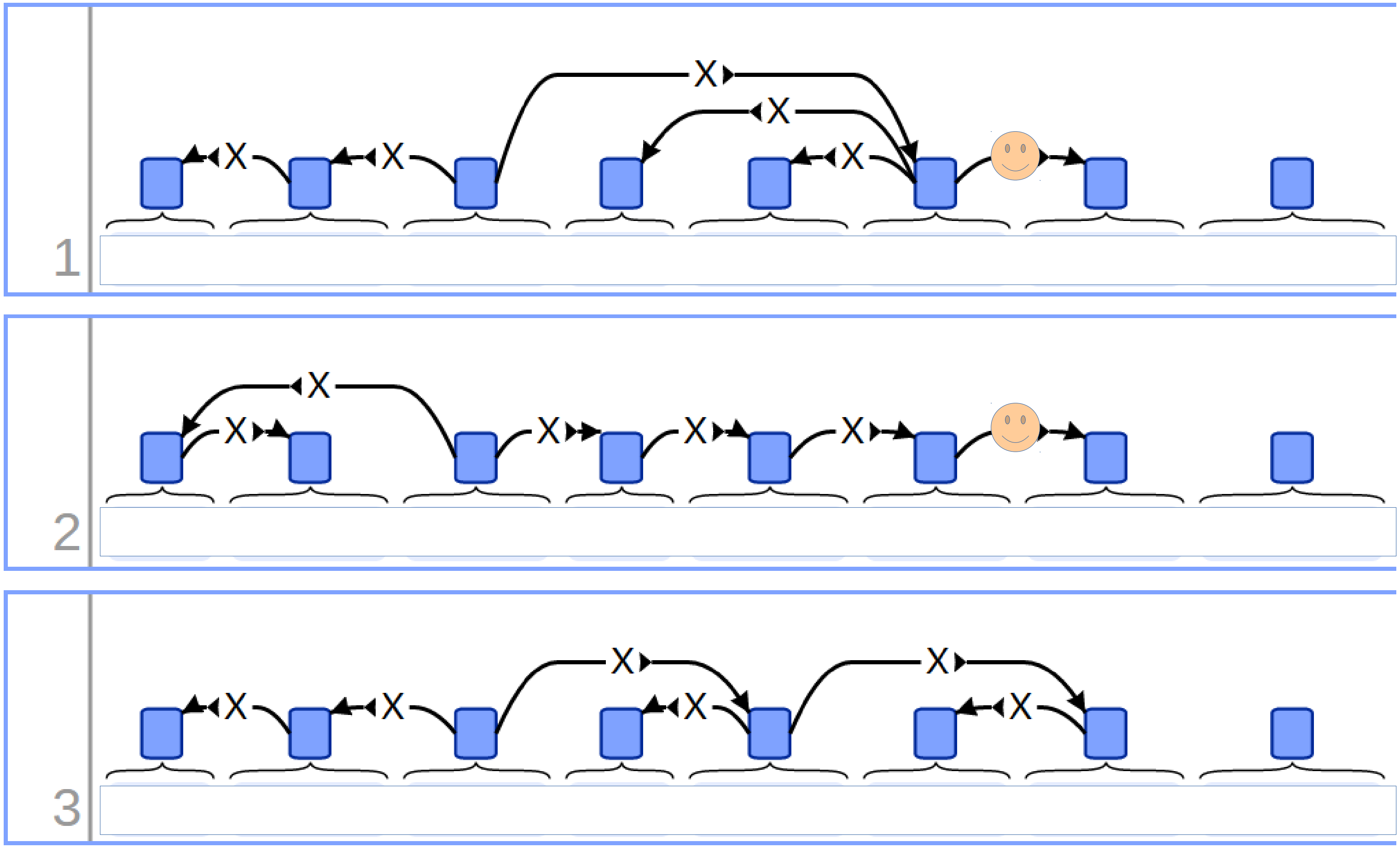
Joakim Nivre, Dan Zeman, Filip Ginter, Sampo Pyysalo, Chris Manning, Marie-Catherine de Marneffe, Natalia Silveira, Slav Petrov, Ryan McDonald, Tim Dozat, Jan Hajič, Jinho Choi, Reut Tsarfaty, Yoav Goldberg, Simonetta Montemagni, Alessandro Lenci, Maria Simi, Cristina Bosco, Veronika Vincze, Richárd Farkas, Teresa Lynn, Jennifer Foster, Prokopis Prokopidis, Jenna Kanerva, Juha Kuokkala, Veronika Laippala, Krister Lindén, Anna Missilä, Hanna Nurmi, Jussi Piitulainen, Aaron Smith, Željko Agić, Nikola Ljubešić, Maria Jesus Aranzabe, Aitziber Atutxa, Iakes Goenaga, Koldo Gojenola, Anders Trærup Johannsen, Hèctor Martínez, Barbara Plank, Petya Osenova, Kiril Simov, Mojgan Seraji, Wolfgang Seeker, Fran Tyers, Aibek Makazhanov, Jon Washington, Çağrı Çöltekin, Arne Skjærholt, Lilja Øvrelid, Miguel Ballesteros, Elena Pascual, Giuseppe Celano, Marco Passarotti, **Martin Popel**, Christophe Onambélé, Dag Haug, Nizar Habash, Riyaz Ahmad, Verginica Mititelu, Catalina Mărănduc, Kaja Dobrovoljc, Tomaž Erjavec, Simon Krek, Yusuke Miyao, Shinsuke Mori, Takaaki Tanaka, Hiroshi Kanayama, Masayuki Asahara, Sumire Uematsu, Rob Voigt, …

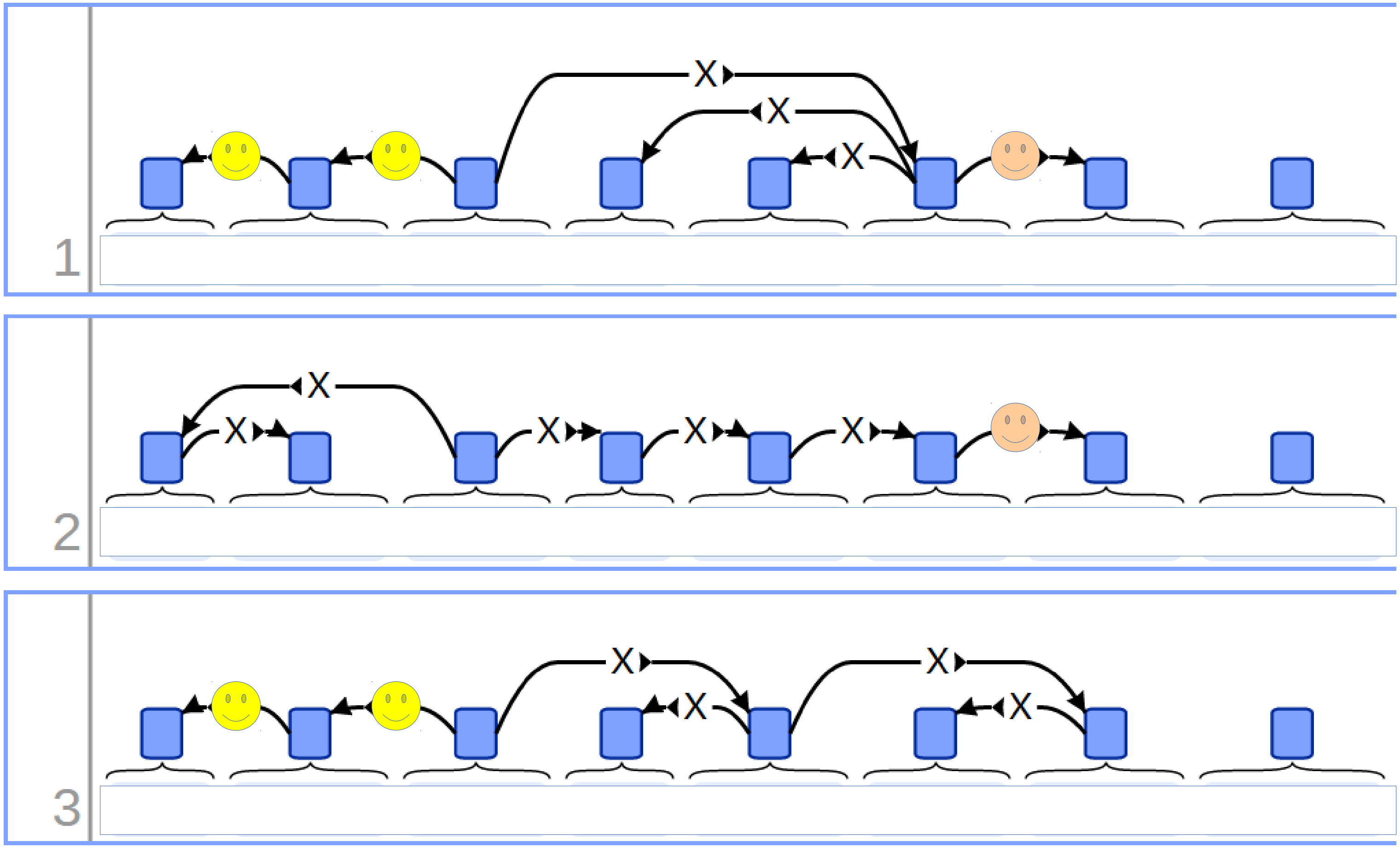
Introduction slides stolen ~~from Joakim Nivre~~

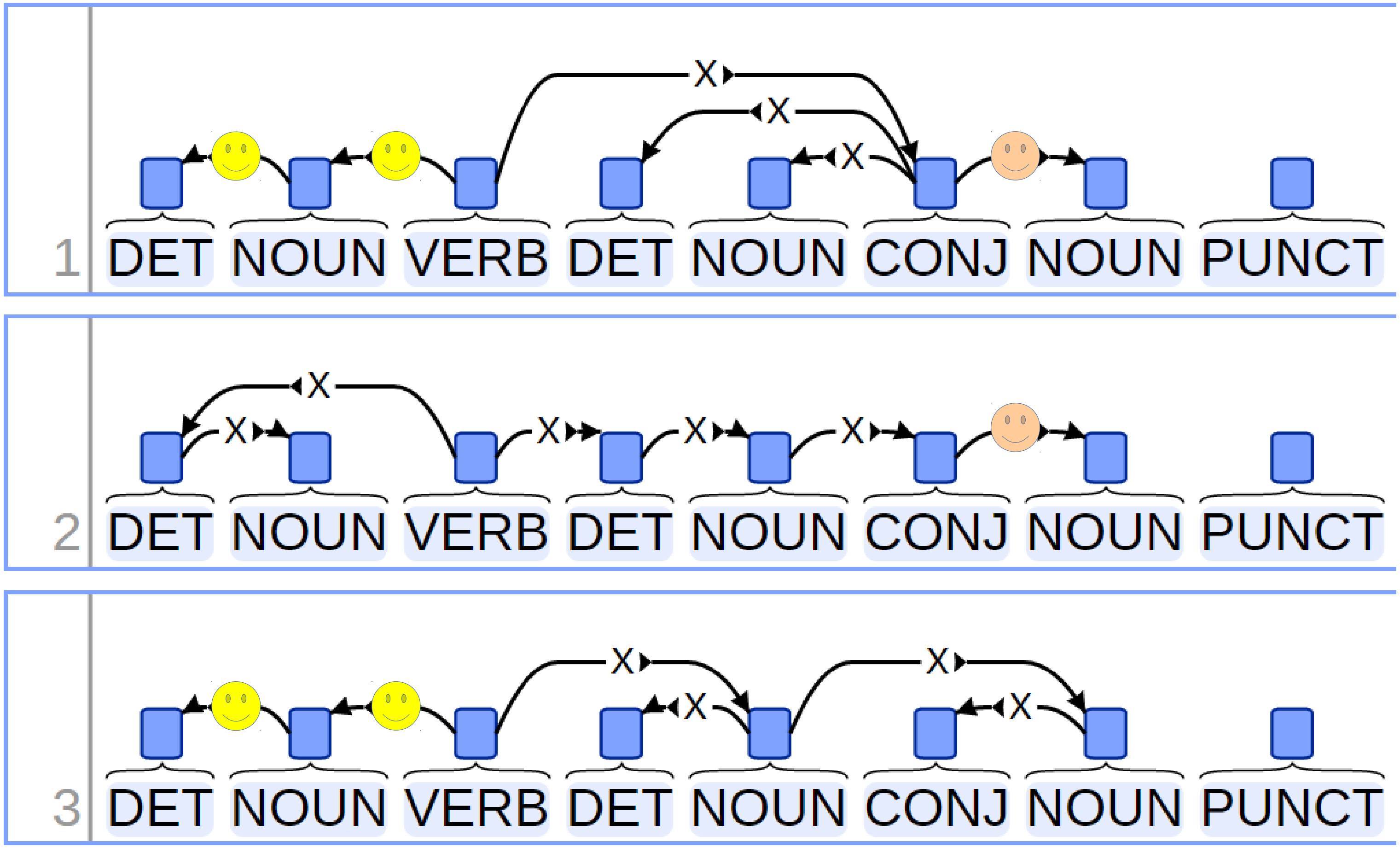
from Dan Zeman

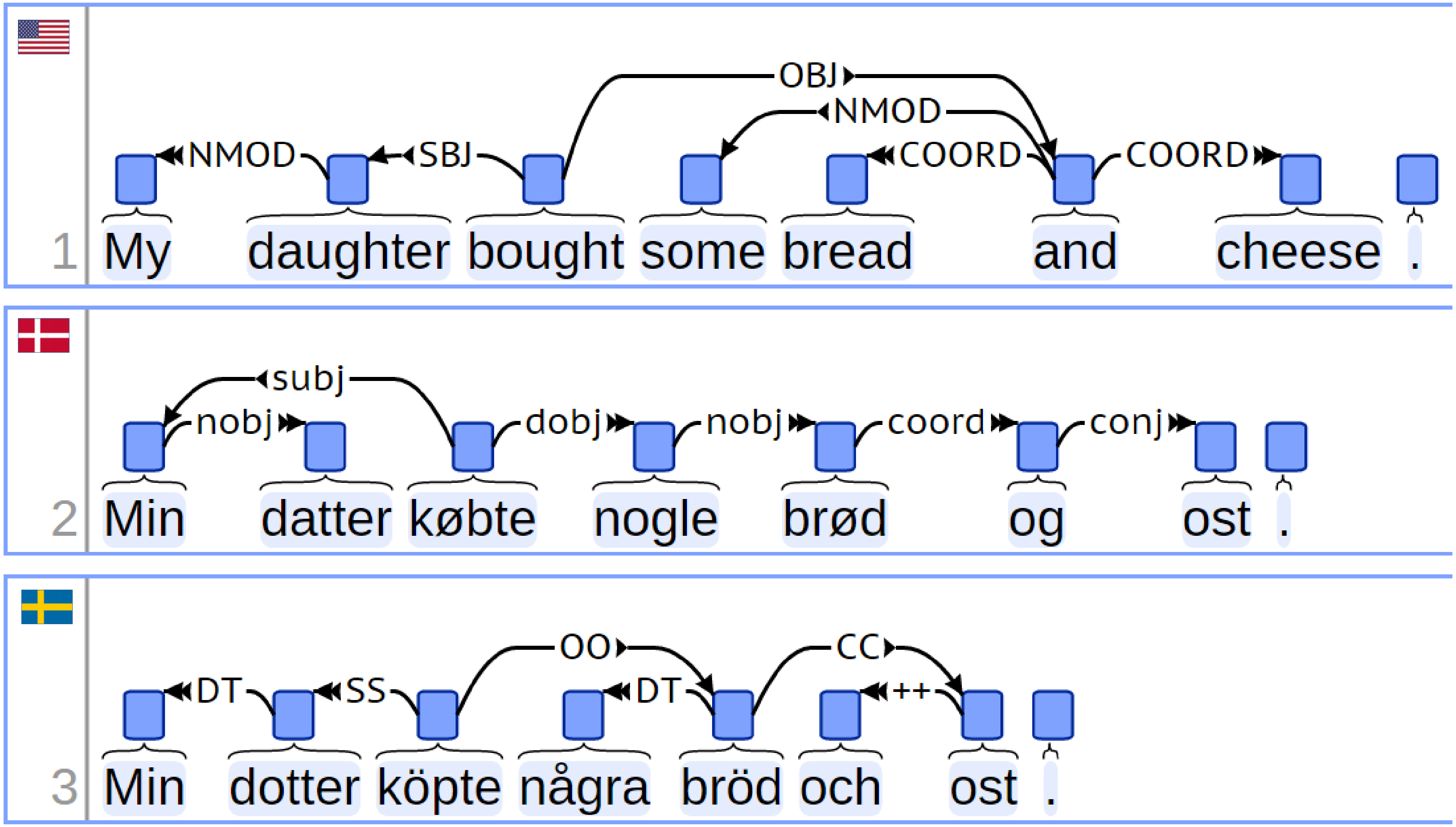


X
X
X
X
X
X
X
1
X
X
X
X
X
X
2
X
X
X
X
X
X
3


1
X
X
X
X
X
X
2
X
X
X
X
X
3
X
X
X
X
X
X
X


1
X
X
X
2
X
X
X
X
X
3
X
X
X
X


1
DET NOUN VERB DET NOUN CONJ NOUN PUNCT
X
X
X
2
DET NOUN VERB DET NOUN CONJ NOUN PUNCT
X
X
X
X
X
3
DET NOUN VERB DET NOUN CONJ NOUN PUNCT
X
X
X
X

1 My daughter bought some bread and cheese .
NMOD
SBJ
OBJ
NMOD
COORD
COORD
2 Min datter købte nogle brød og ost .
subj
nobj
dobj
nobj
coord
conj
3 Min dotter köpte några bröd och ost .
DT
SS
OO
DT
CC
++

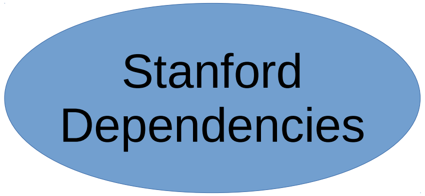
Universal Dependencies

http://universaldependencies.org

https://github.com/UniversalDependencies/

Universal Dependencies

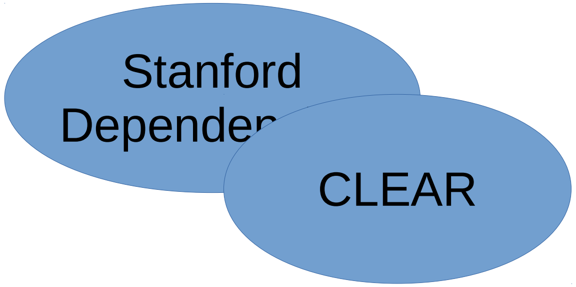
http://universaldependencies.org



Stanford Dependencies

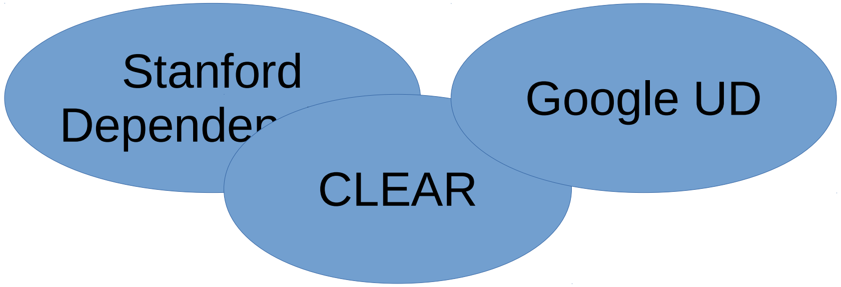
Universal Dependencies

http://universaldependencies.org



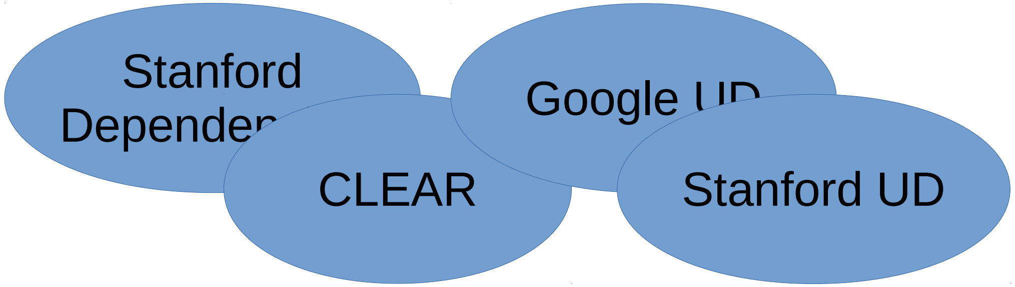
Universal Dependencies

http://universaldependencies.org



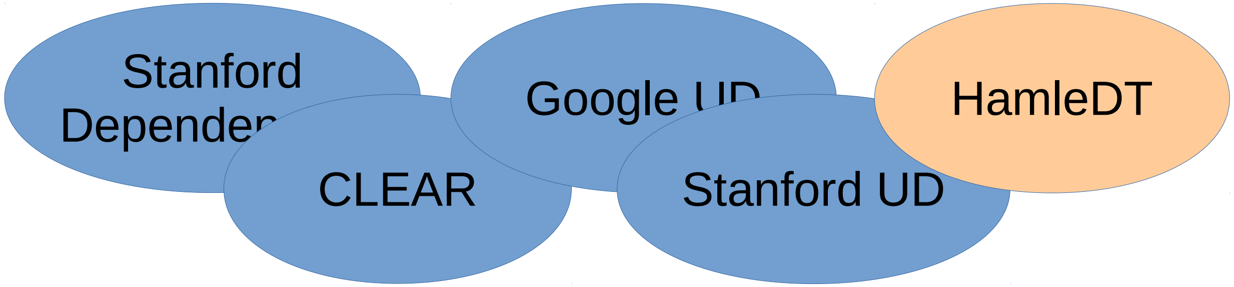
Universal Dependencies

http://universaldependencies.org



Universal Dependencies

http://universaldependencies.org



Stanford Dependen

Google UD

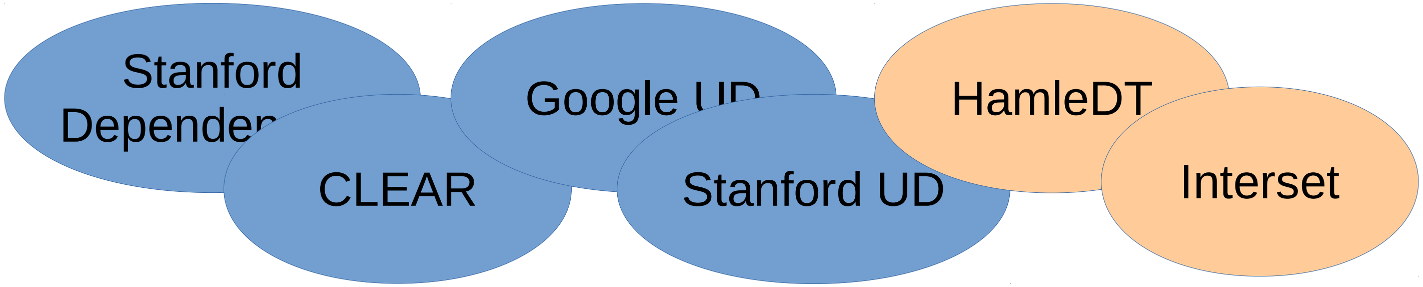
HamleDT

CLEAR

Stanford UD

Universal Dependencies

http://universaldependencies.org



Stanford Dependen

Google UD

CLEAR

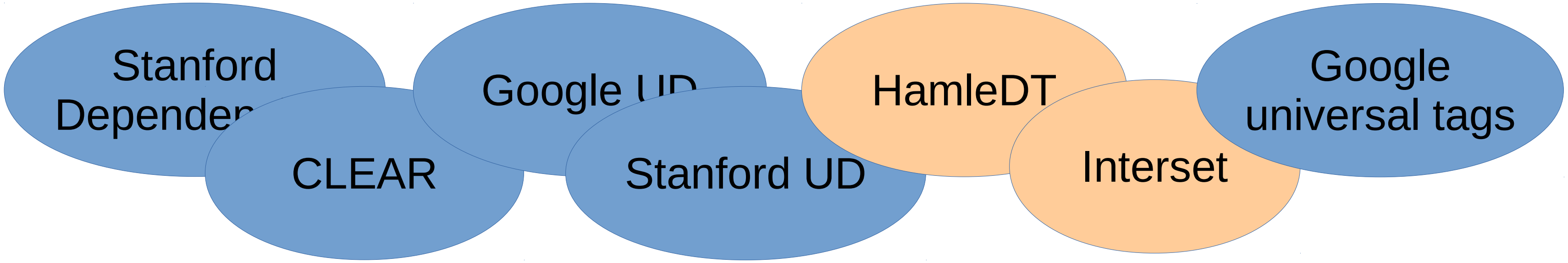
Stanford UD

HamleDT

Interset

Universal Dependencies

http://universaldependencies.org



Stanford Dependen

CLEAR

Google UD

Stanford UD

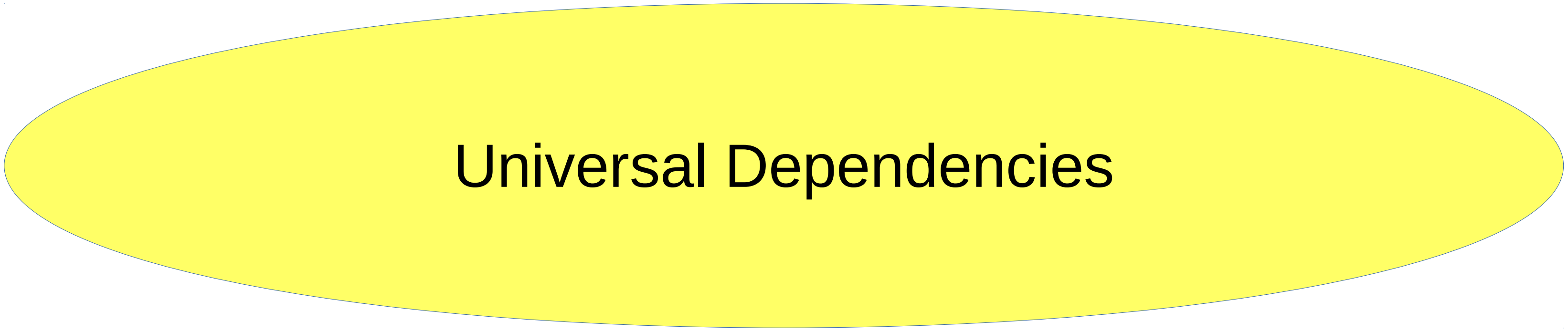
HamleDT

Interset

Google universal tags

Universal Dependencies

http://universaldependencies.org



Universal Dependencies

http://universaldependencies.org

9 treebanks with spoken
9 with social/blog/reviews

- 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 18 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: released 64 treebanks of 47 languages (UD 1.4), total 12M tokens
  - 2017-03: UD 2.0 planned (goal: 80 treebanks of 56 languages)

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
  - Not an ideal parsing representation –
    but useful for comparative evaluation
  - Not the ultimate annotation scheme –
    but a lightweight lingua franca

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
  - Languages select from a universal pool of categories
  - 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	CCONJ	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	CCONJ	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	CCONJ	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

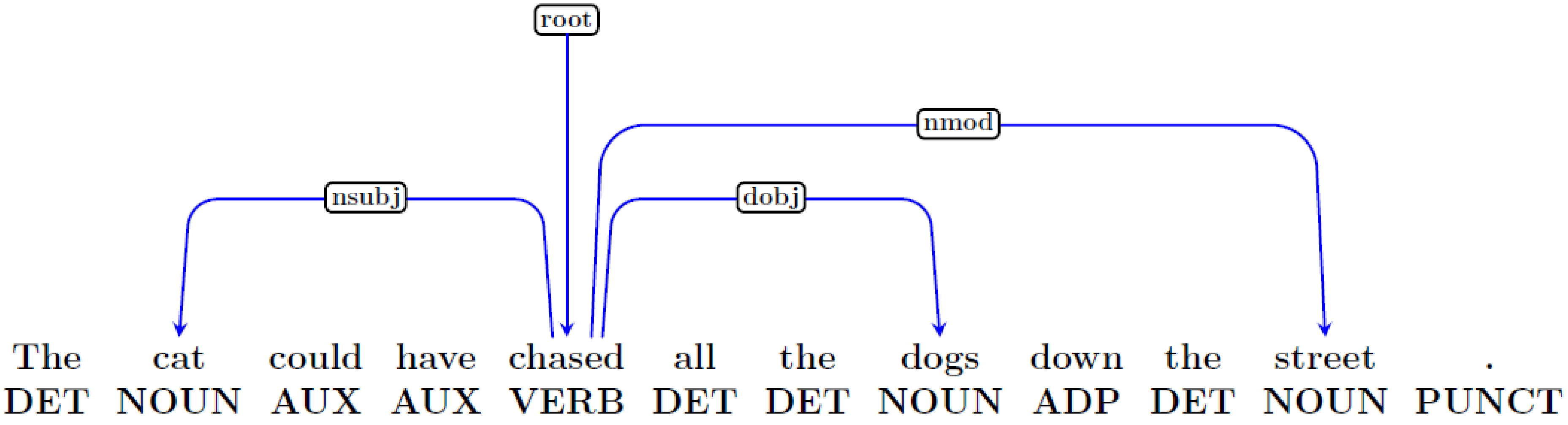
Lexical	Inflectional / Nominal	Inflectional / Verbal
PronType	Gender	VerbForm
NumType	Animacy	Mood
Poss	Number	Tense
Reflex	Case	Aspect
Foreign	Definite	Voice
Abbr	Degree	Person
		Polarity
		Polite, Evident

- Standardized inventory of 21 morphological features, based on Interset (Zeman, 2008)
- Languages select relevant features and can add language-specific features or values with documentation

Syntax

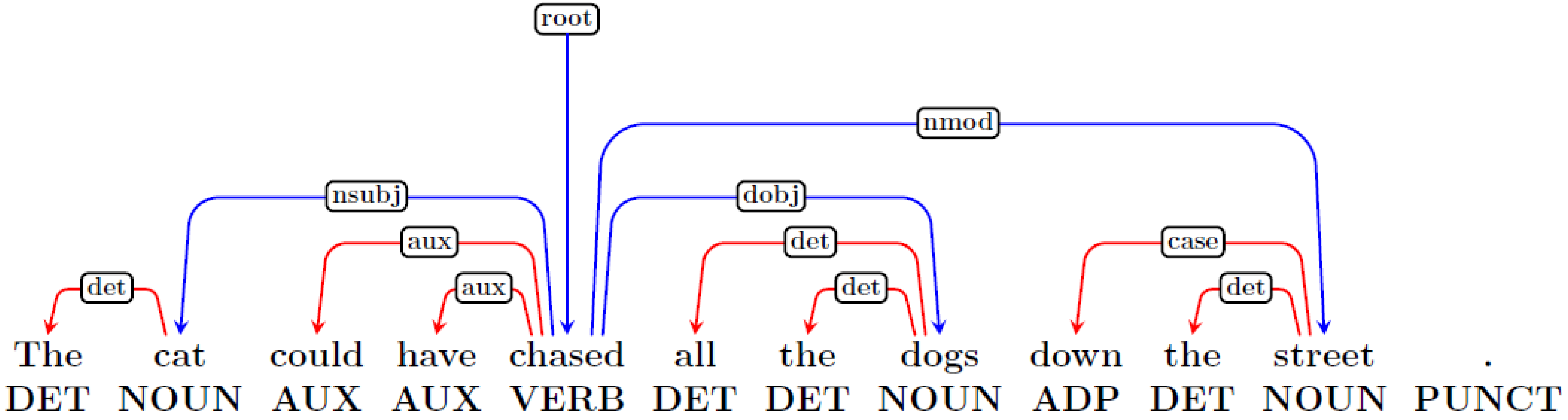
The	cat	could	have	chased	all	the	dogs	down	the	street	.
DET	NOUN	AUX	AUX	VERB	DET	DET	NOUN	ADP	DET	NOUN	PUNCT

Syntax



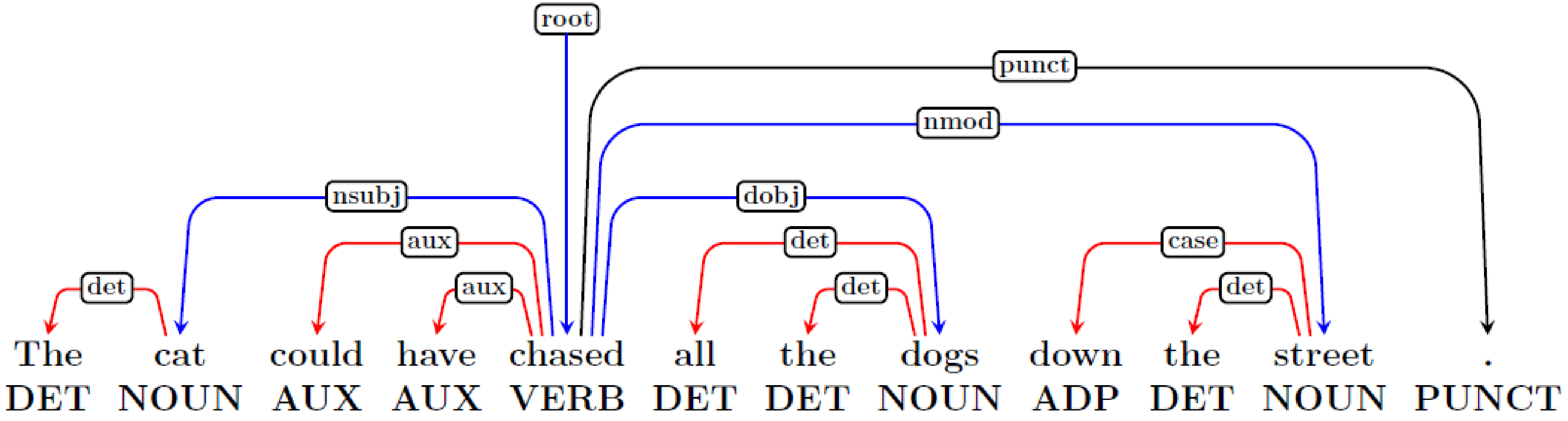
- Content words are related by dependency relations

Syntax

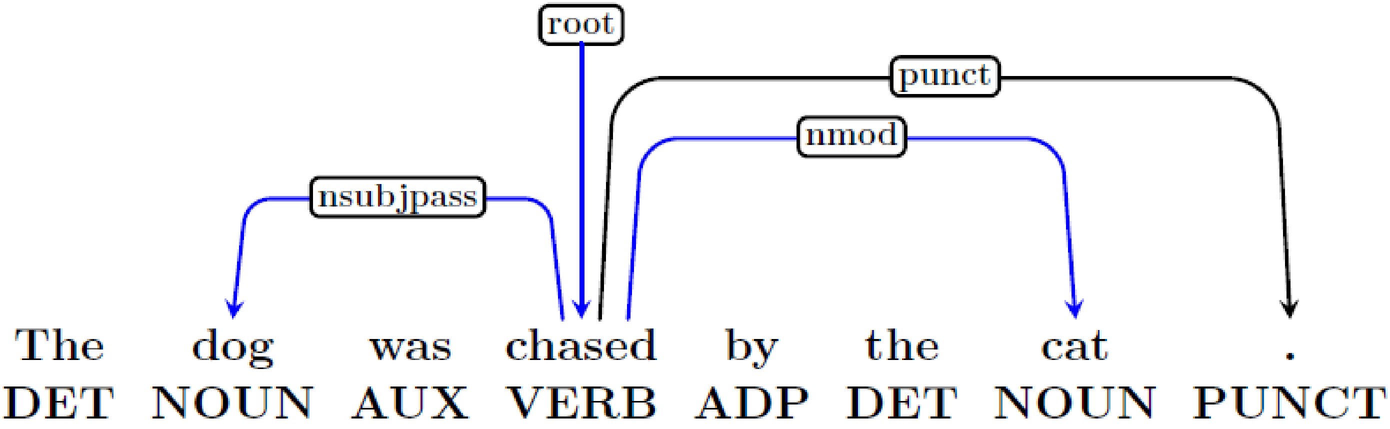


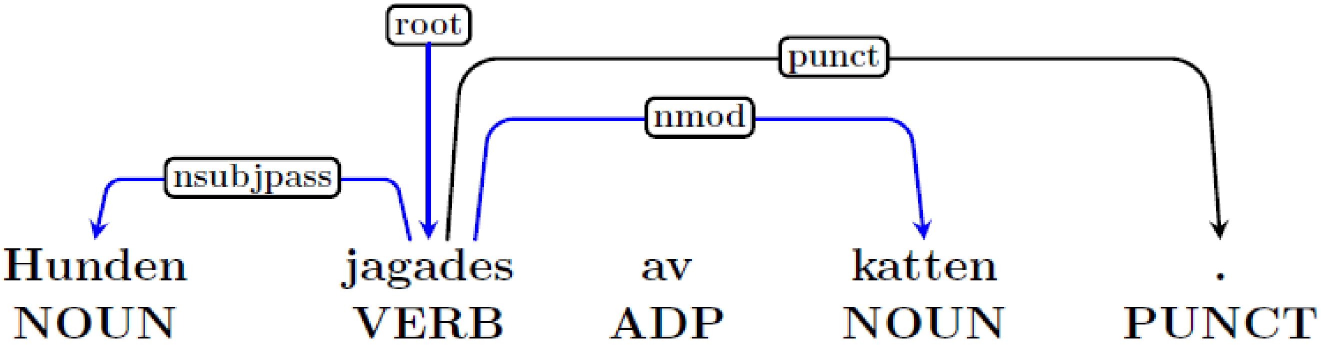
- Content words are related by dependency relations
- Function words attach to closest content word

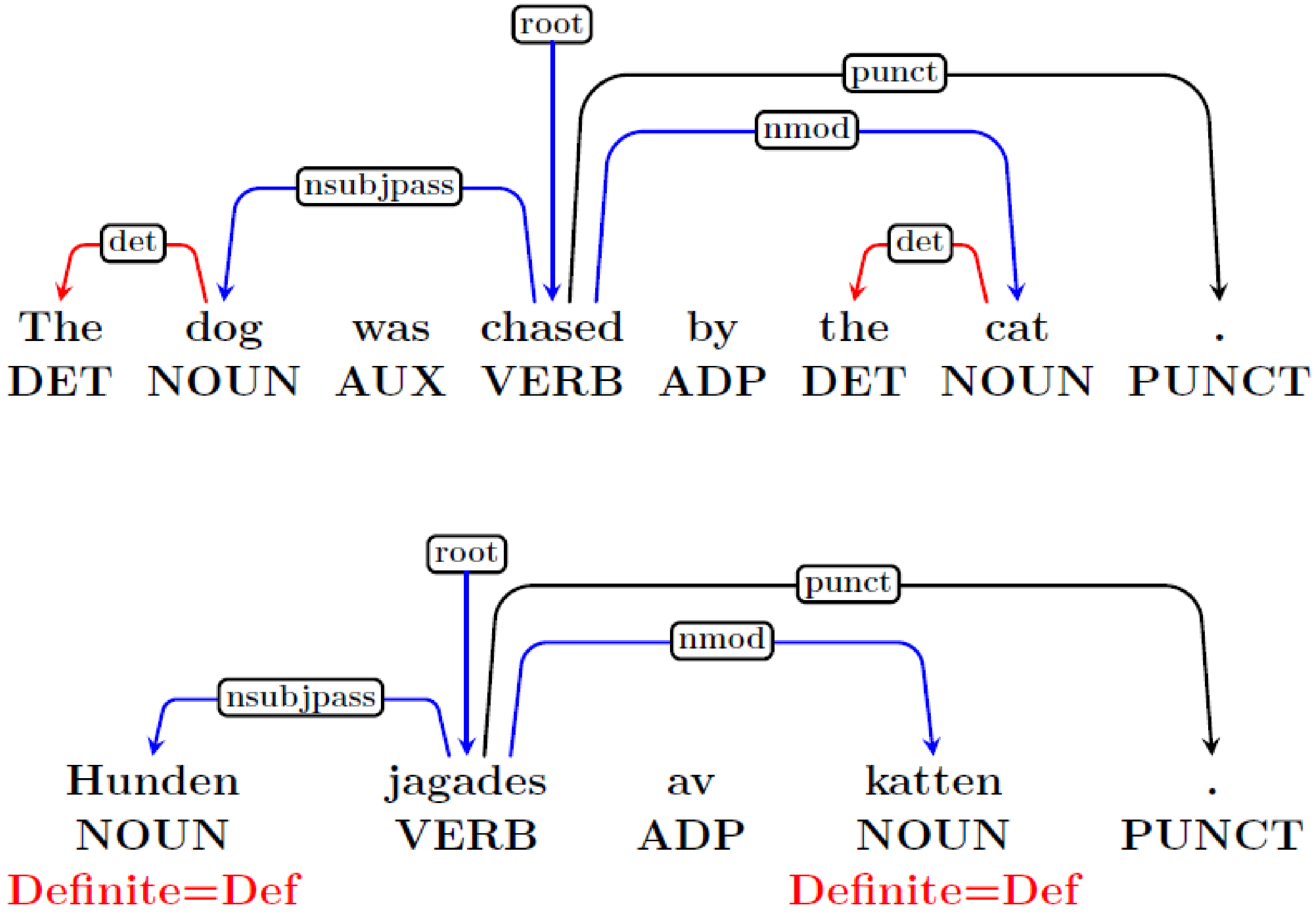
Syntax

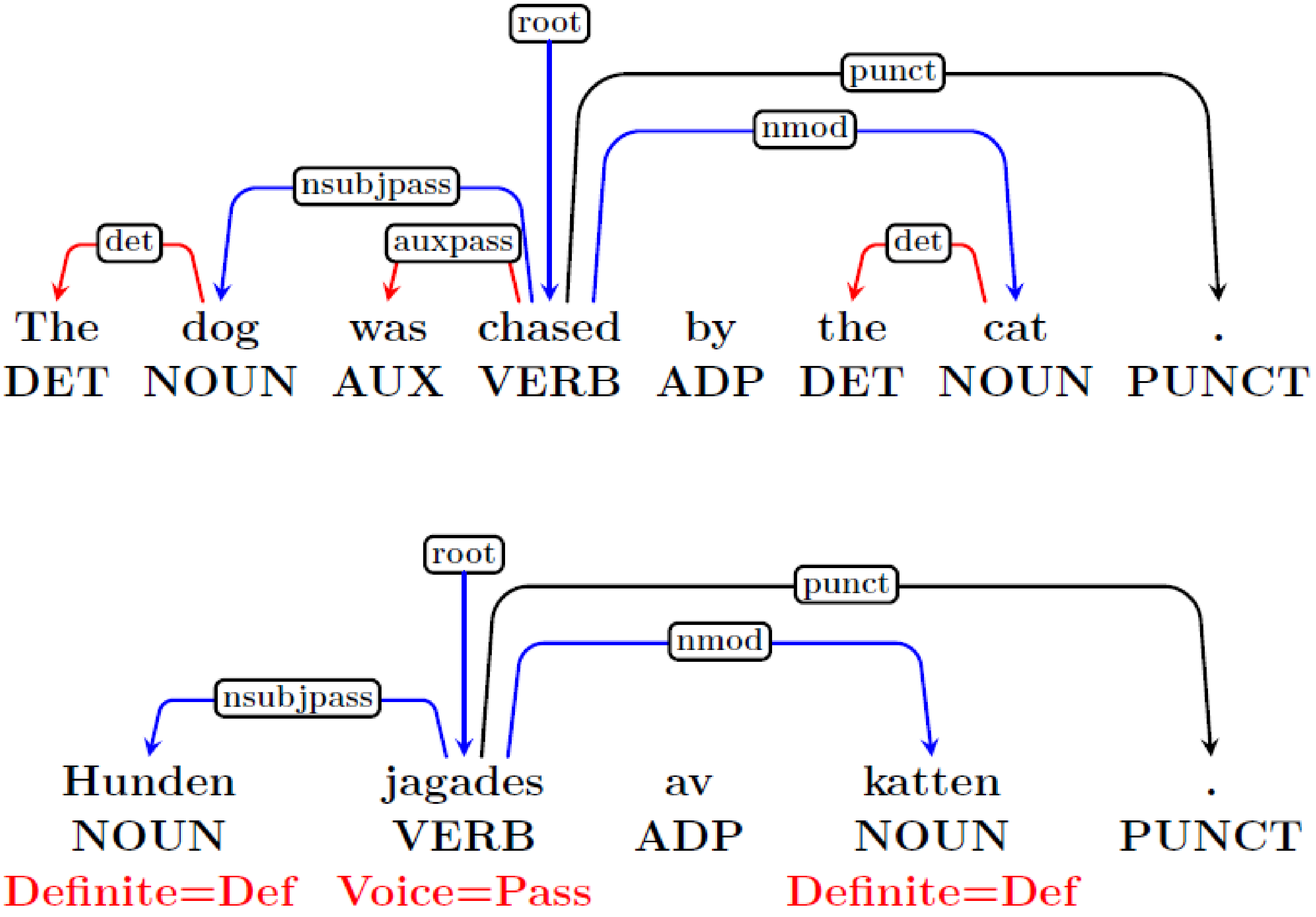


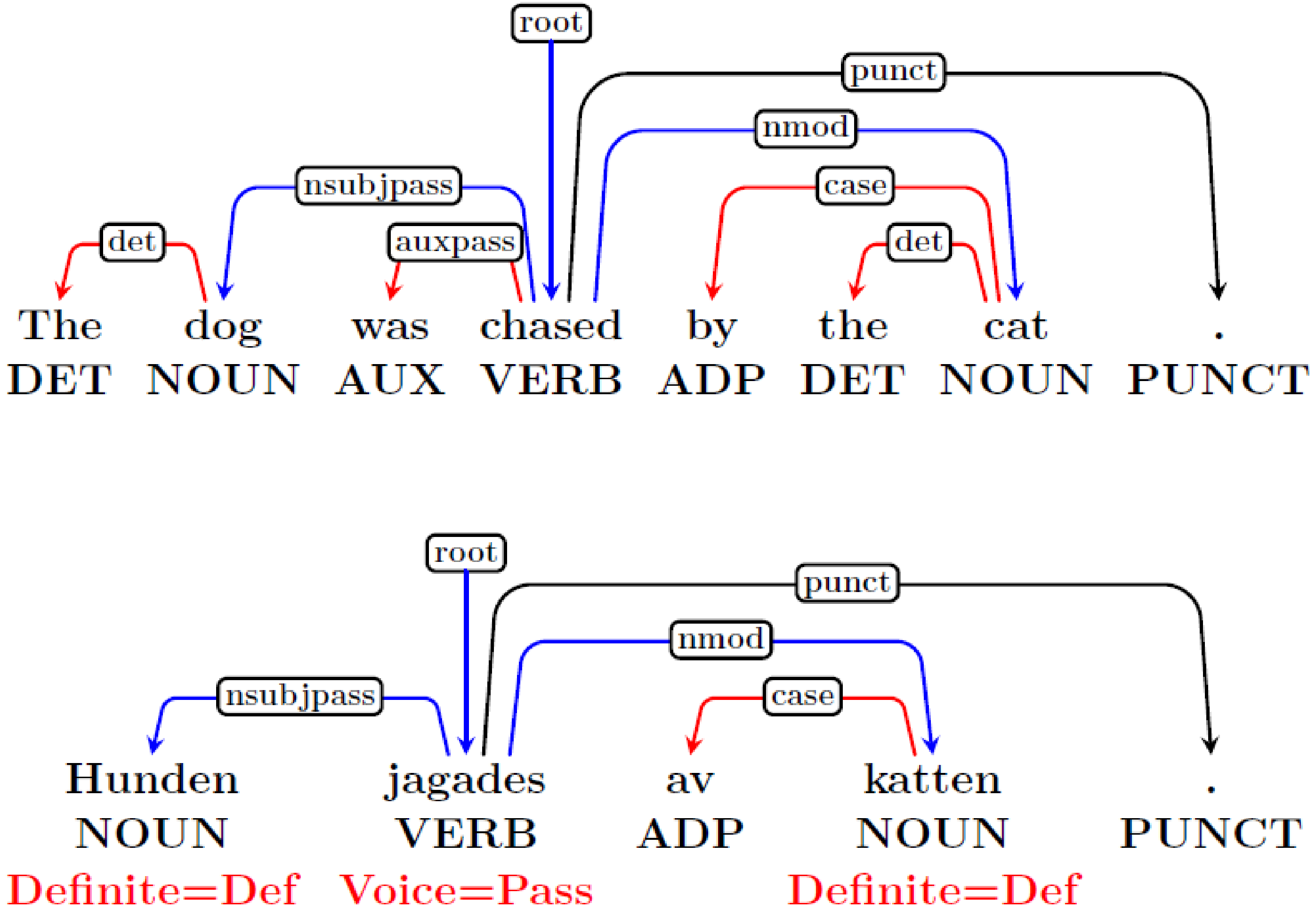
- Content words are related by dependency relations
- Function words attach to closest content word
- Punctuation attach to head of phrase or clause


root
punct
nmod
nsubjpass
The
DET
dog
NOUN
was
AUX
chased
VERB
by
ADP
the
DET
cat
NOUN
.
PUNCT


root
punct
nmod
nsubjpass
Hunden
NOUN
jagades
VERB
av
ADP
katten
NOUN
.
PUNCT


root
punct
nmod
nsubjpass
det
det
The dog was chased by the cat .
DET NOUN AUX VERB ADP DET NOUN PUNCT
root
punct
nmod
nsubjpass
Hunden jagades av katten .
NOUN VERB ADP NOUN PUNCT
Definite=Def
Definite=Def


root
punct
nmod
nsubjpass
auxpass
det
det
The dog was chased by the cat .
DET NOUN AUX VERB ADP DET NOUN PUNCT
root
punct
nmod
nsubjpass
Hunden jagades av katten .
NOUN VERB ADP NOUN PUNCT
Definite=Def
Voice=Pass
Definite=Def

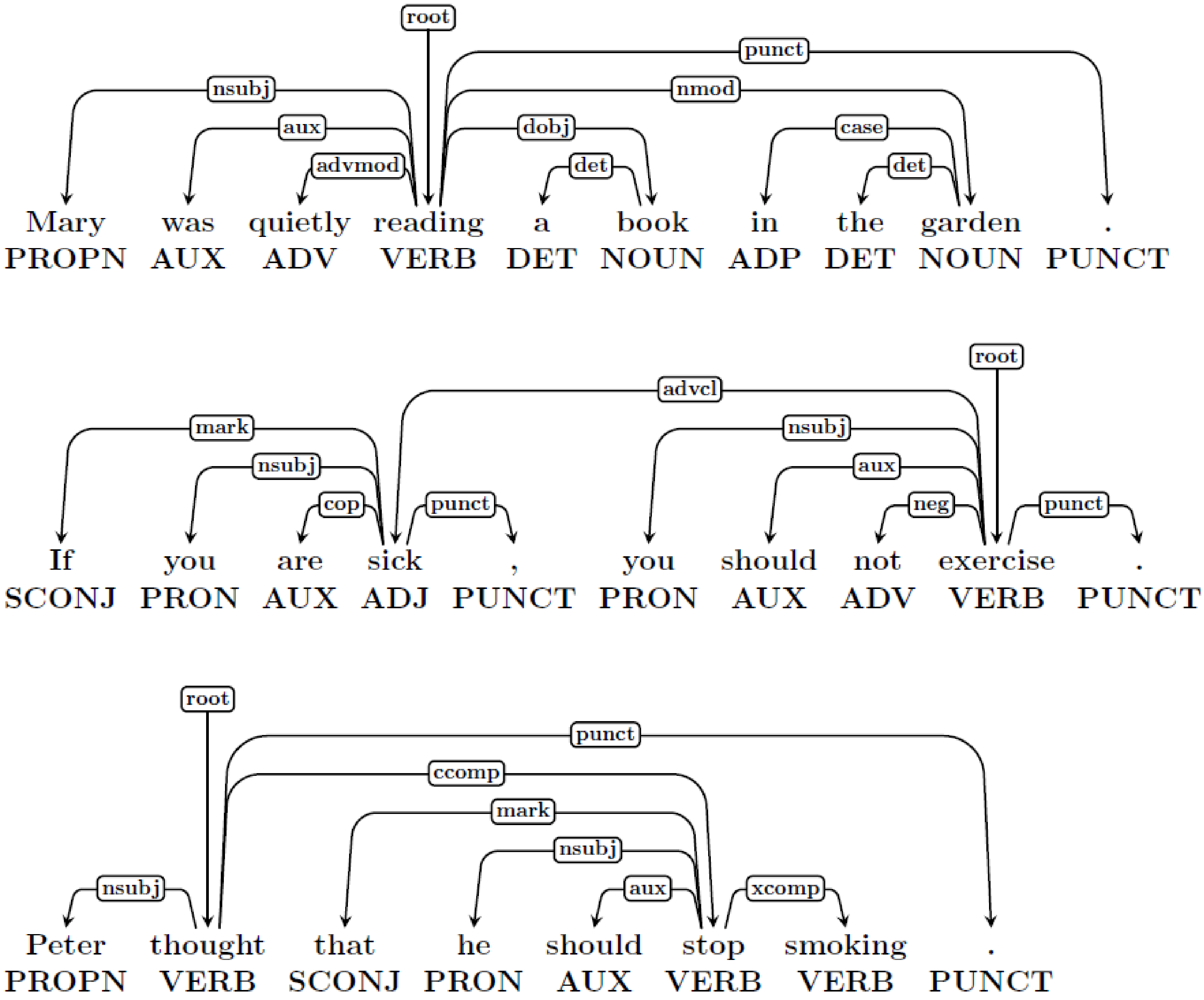

root
punct
nmod
nsubjpass
case
det
auxpass
det
The DET
dog NOUN
was AUX
chased VERB
by ADP
the DET
cat NOUN
. PUNCT
root
punct
nmod
nsubjpass
case
Hunden NOUN Definite=Def
jagades VERB Voice=Pass
av ADP
katten NOUN Definite=Def
. PUNCT

Dependency Relations

- Taxonomy of 37 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)

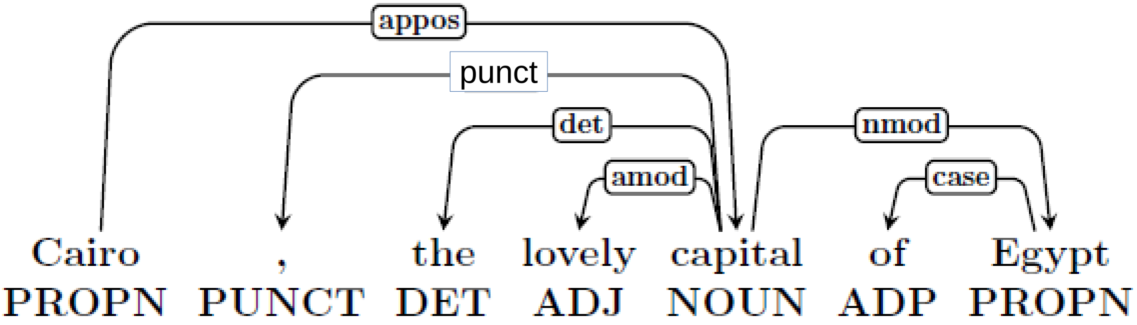
Dependents of Clausal Predicates

	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


root
punct
nsubj
nmod
aux
dobj
case
advmod
det
det
Mary was quietly reading a book in the garden .
PROPN AUX ADV VERB DET NOUN ADP DET NOUN PUNCT
root
advcl
mark
nsubj
nsubj
aux
cop
punct
neg
punct
If you are sick , you should not exercise .
SCONJ PRON AUX ADJ PUNCT PRON AUX ADV VERB PUNCT
root
punct
ccomp
mark
nsubj
nsubj
aux
xcomp
Peter thought that he should stop smoking .
PROPN VERB SCONJ PRON AUX VERB VERB PUNCT

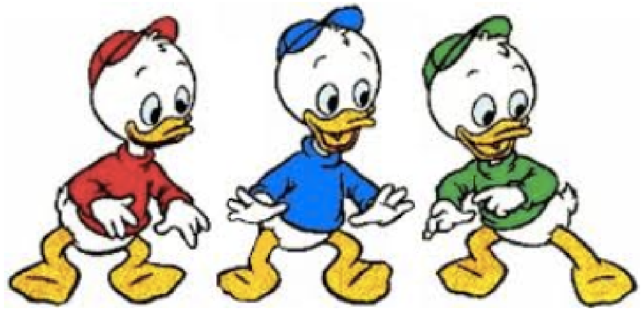
Dependents of Nominals

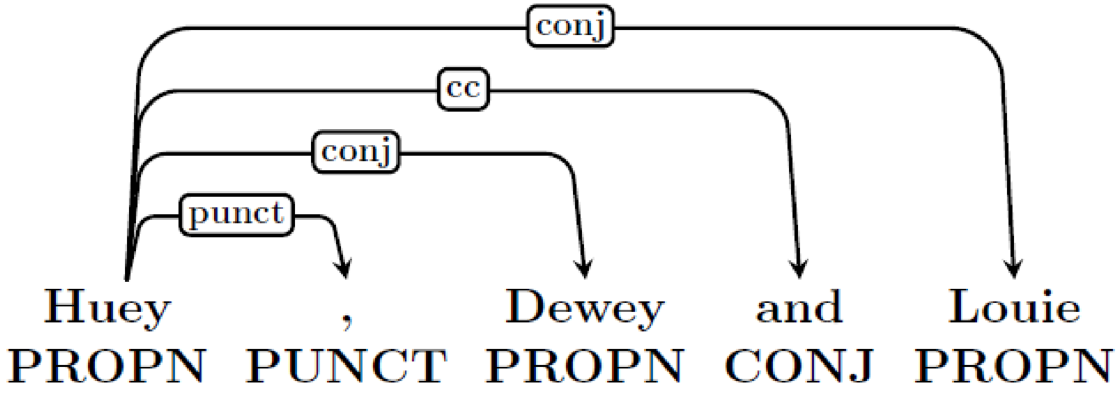
Nominal	Clausal	Other
nmod appos nummod	acl	amod det neg case



Coordination
conj
cc
(punct)

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	*in spite of, as well as, ad hoc*
name	*Roger Bacon, New York*
compound	*phone book, four thousand, dress up*
goeswith	*notwith standing, with out*

- 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
orphan	Orphans in ellipsis linked to promoted head
reparandum	Disfluency linked to (speech) repair
foreign	Elements within opaque stretches of code switching
dep	Unspecified dependency
root	Syntactically independent element of clause/phrase

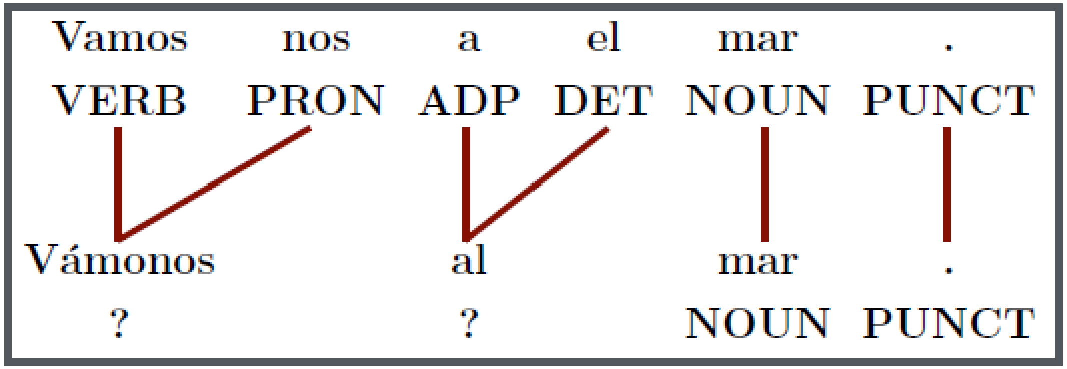
Language-Specific Relations

- 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 *(dress up)*
nmod:poss	Genitive nominal *(Mary ’s book)*
nmod:agent	Agent in passive *(saved by the bell)*
cc:preconj	Preconjunction *(both … and)*
det:predet	Predeterminer *(all those …)*

Word Segmentation

- How do we segment sentences into words?
  - Depends on language and writing system, often non-trivial
  - Segmentation must be reproducible on new data
- Two options provided:
  - Only include words in treebank, but document segmentation
  - Include mapping from low-level tokenization to words in treebank



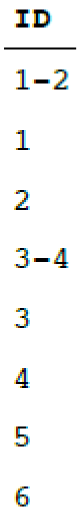
Vamos nos a el mar .

VERB PRON ADP DET NOUN PUNCT

Vámonos al mar .

? ? NOUN PUNCT

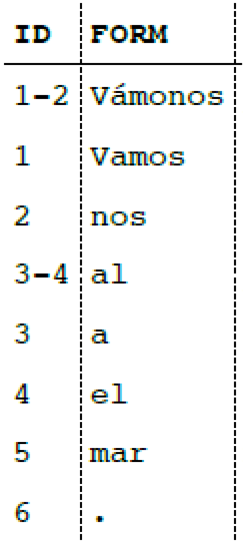
CoNLL-U Format



ID
1-2
1
2
3-4
3
4
5
6

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

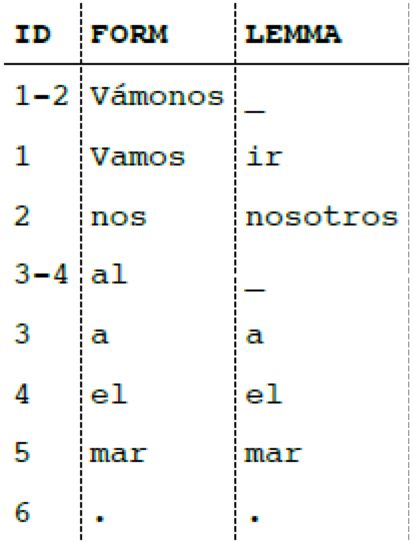
CoNLL-U Format



ID	FORM
1-2	Vámonos
1	Vamos
2	nos
3-4	al
3	a
4	el
5	mar
6	.

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

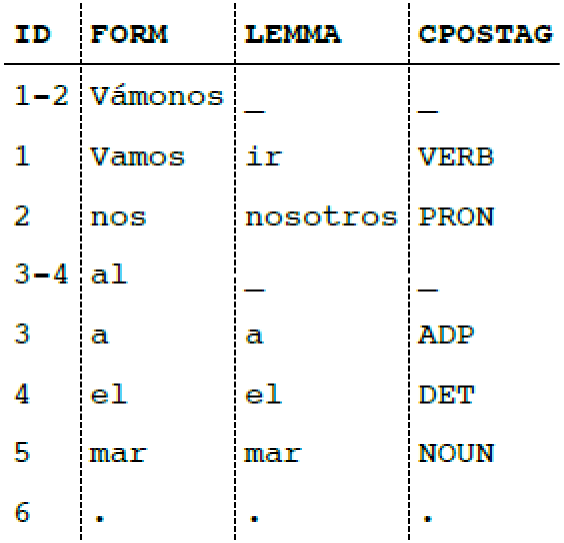
CoNLL-U Format



ID	FORM	LEMMA
1-2	Vámonos	_
1	Vamos	ir
2	nos	nosotros
3-4	al	_
3	a	a
4	el	el
5	mar	mar
6	.	.

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

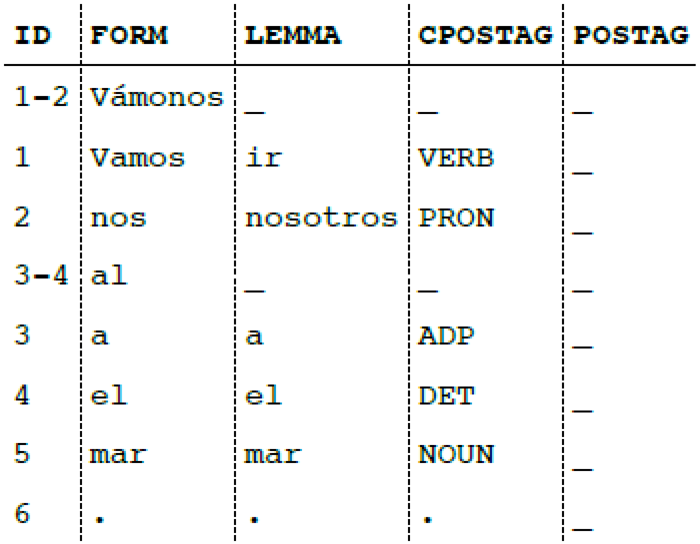
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG
1-2	Vámonos	_	_
1	Vamos	ir	VERB
2	nos	nosotros	PRON
3-4	al	_	_
3	a	a	ADP
4	el	el	DET
5	mar	mar	NOUN
6	.	.	.

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

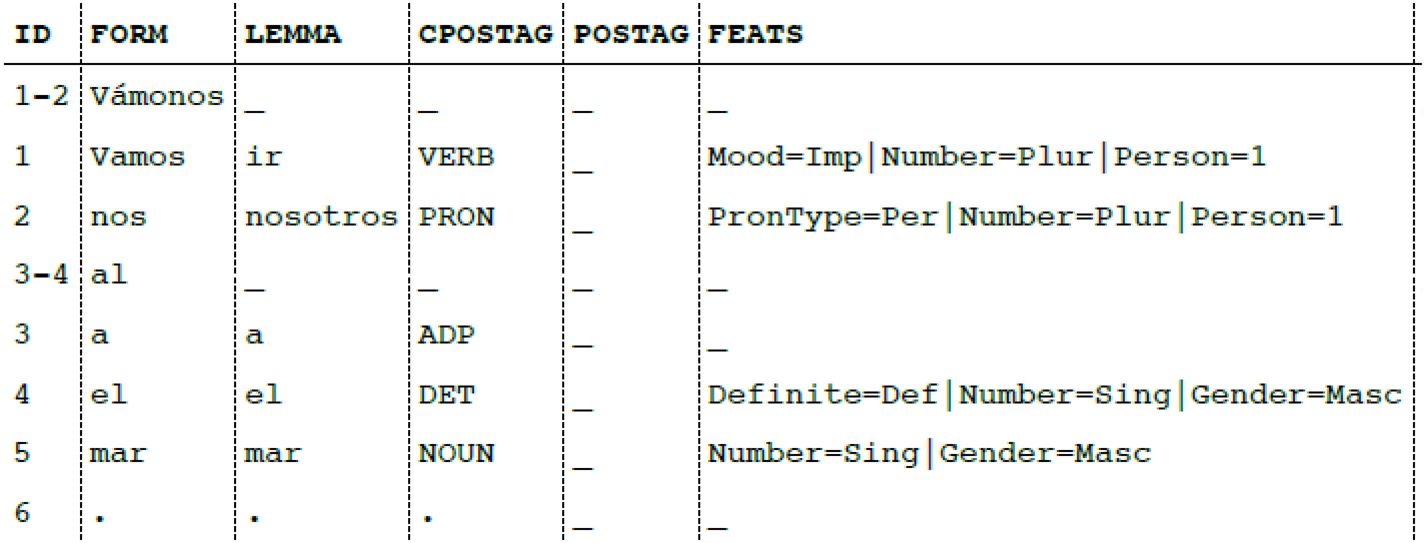
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG
1-2	Vámonos	_	_	_
1	Vamos	ir	VERB	_
2	nos	nosotros	PRON	_
3-4	al	_	_	_
3	a	a	ADP	_
4	el	el	DET	_
5	mar	mar	NOUN	_
6	.	.	.	_

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

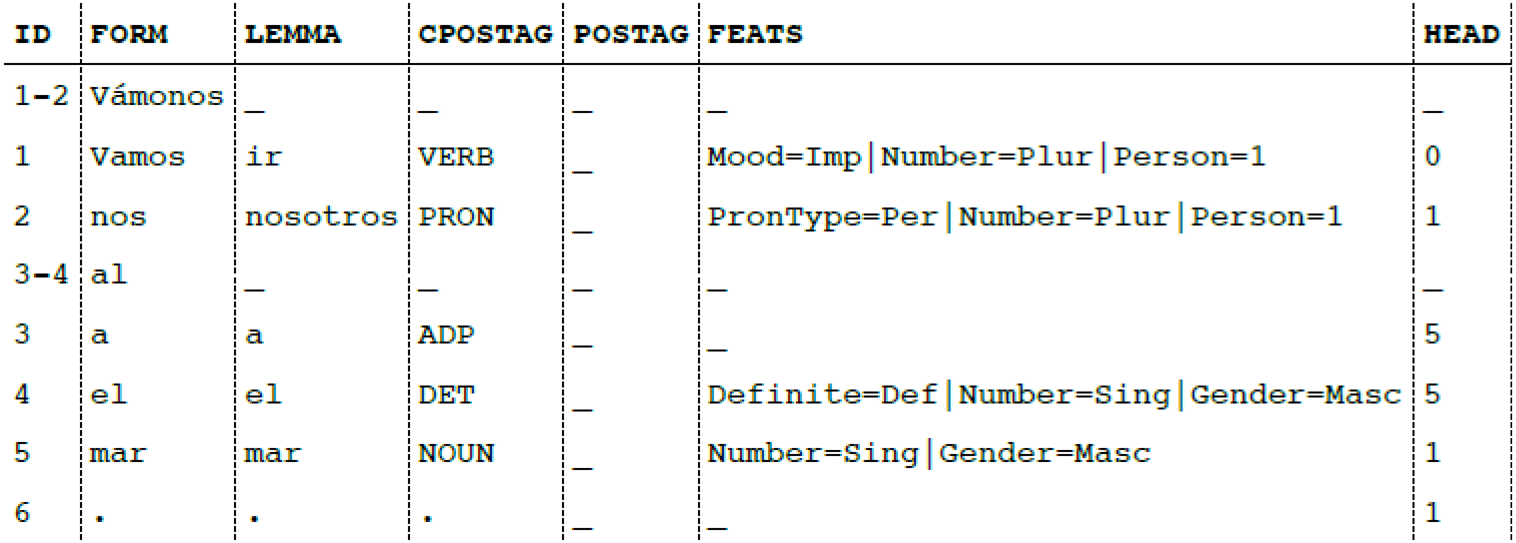
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG	FEATS
1-2	Vámonos	_	_	_	_
1	Vamos	ir	VERB	_	Mood=Imp|Number=Plur|Person=1
2	nos	nosotros	PRON	_	PronType=Per|Number=Plur|Person=1
3-4	al	_	_	_	_
3	a	a	ADP	_	_
4	el	el	DET	_	Definite=Def|Number=Sing|Gender=Masc
5	mar	mar	NOUN	_	Number=Sing|Gender=Masc
6	.	.	.	_	_

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

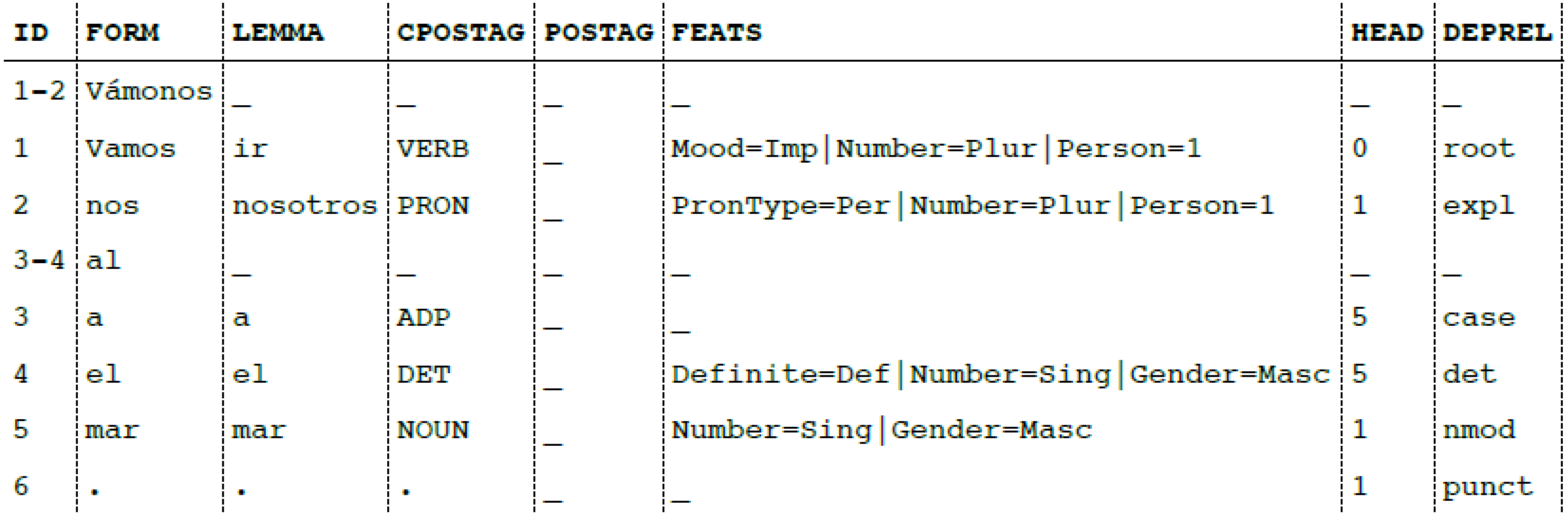
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG	FEATS	HEAD
1-2	Vámonos	_	_	_	_	_
1	Vamos	ir	VERB	_	Mood=Imp|Number=Plur|Person=1	0
2	nos	nosotros	PRON	_	PronType=Per|Number=Plur|Person=1	1
3-4	al	_	_	_	_	_
3	a	a	ADP	_	_	5
4	el	el	DET	_	Definite=Def|Number=Sing|Gender=Masc	5
5	mar	mar	NOUN	_	Number=Sing|Gender=Masc	1
6	.	.	.	_	_	1

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

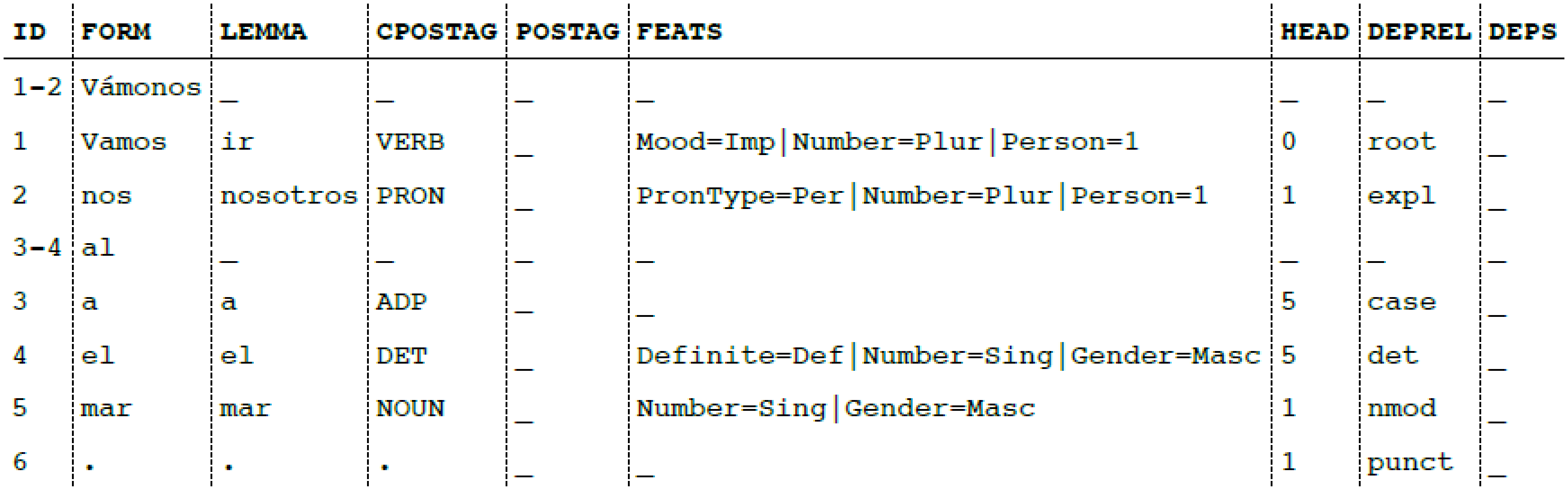
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG	FEATS	HEAD	DEPREL
1-2	Vámonos	_	_	_	_	_	_
1	Vamos	ir	VERB	_	Mood=Imp|Number=Plur|Person=1	0	root
2	nos	nosotros	PRON	_	PronType=Per|Number=Plur|Person=1	1	expl
3-4	al	_	_	_	_	_	_
3	a	a	ADP	_	_	5	case
4	el	el	DET	_	Definite=Def|Number=Sing|Gender=Masc	5	det
5	mar	mar	NOUN	_	Number=Sing|Gender=Masc	1	nmod
6	.	.	.	_	_	1	punct

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

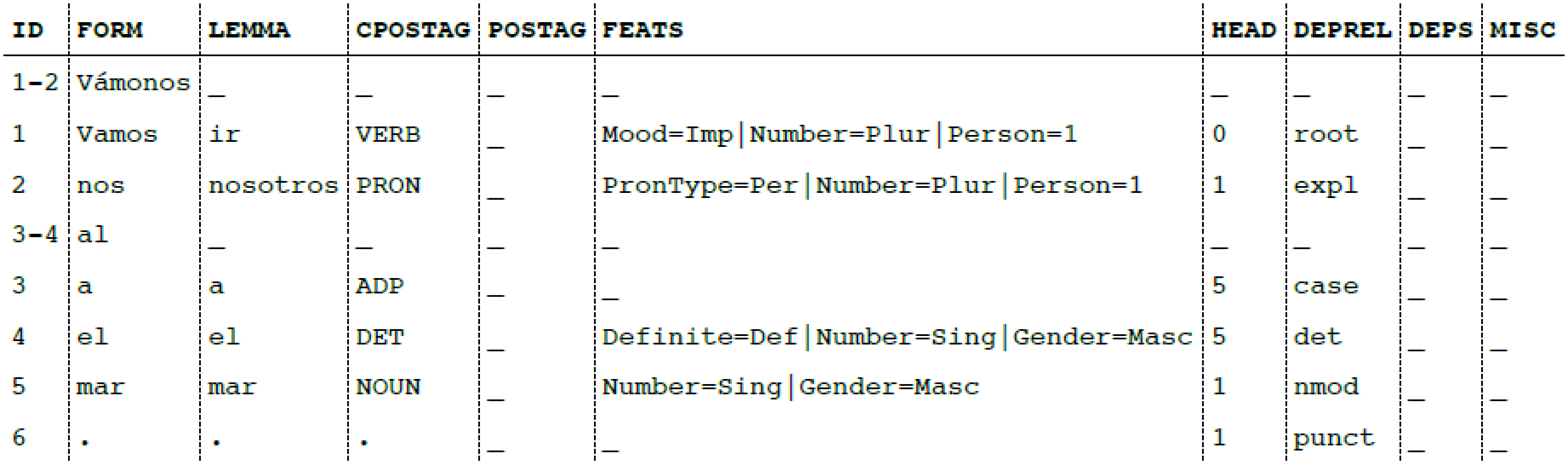
CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG	FEATS	HEAD	DEPREL	DEPS
1-2	Vámonos	_	_	_	_	_	_	_
1	Vamos	ir	VERB	_	Mood=Imp|Number=Plur|Person=1	0	root	_
2	nos	nosotros	PRON	_	PronType=Per|Number=Plur|Person=1	1	expl	_
3-4	al	_	_	_	_	_	_	_
3	a	a	ADP	_	_	5	case	_
4	el	el	DET	_	Definite=Def|Number=Sing|Gender=Masc	5	det	_
5	mar	mar	NOUN	_	Number=Sing|Gender=Masc	1	nmod	_
6	.	.	.	_	_	1	punct	_

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

CoNLL-U Format



ID	FORM	LEMMA	CPOSTAG	POSTAG	FEATS	HEAD	DEPREL	DEPS	MISC
1-2	Vámonos	_	_	_	_	_	_	_	_
1	Vamos	ir	VERB	_	Mood=Imp|Number=Plur|Person=1	0	root	_	_
2	nos	nosotros	PRON	_	PronType=Per|Number=Plur|Person=1	1	expl	_	_
3-4	al	_	_	_	_	_	_	_	_
3	a	a	ADP	_	_	5	case	_	_
4	el	el	DET	_	Definite=Def|Number=Sing|Gender=Masc	5	det	_	_
5	mar	mar	NOUN	_	Number=Sing|Gender=Masc	1	nmod	_	_
6	.	.	.	_	_	1	punct	_	_

- Revised version of the CoNLL-X format
- Two-level segmentation and secondary dependencies

Tools for annotating trees

- **TrEd** (+Treex/EasyTreex extension)
  http://ufal.mff.cuni.cz/tred/
  http://ufal.mff.cuni.cz/treex/install.html
  very powerful & customizable, Perl, old
- **Brat** http://brat.nlplab.org/ online/JS+Python
  UD support (see Cairo mini treebank)
- **EasyTree** https://github.com/alexalittle/easytree
  online demo http://ufallab.ms.mff.cuni.cz/~popel/easytree/
  perhaps too simple
- **GraphAnno** :-) https://github.com/LBierkandt/graph-anno

UDPipe – automatic analysis

- http://ufal.cz/udpipe Try it online/as webservice http://lindat.mff.cuni.cz/services/udpipe/
- End-to-end, batteries included:

  segment, tokenize, tag, morpho, lemma, labelled parsing
- Pretrained models for all the UD (1.2) langs
- User friendly (outputs CoNLL-U, Table, SVG)
- State-of-the-art quality, ultra fast
- Open-source, easy install for Linux, OS X, Win
- Interfaces for C++, C#, Java, Perl, Python
- Easily train on your own data

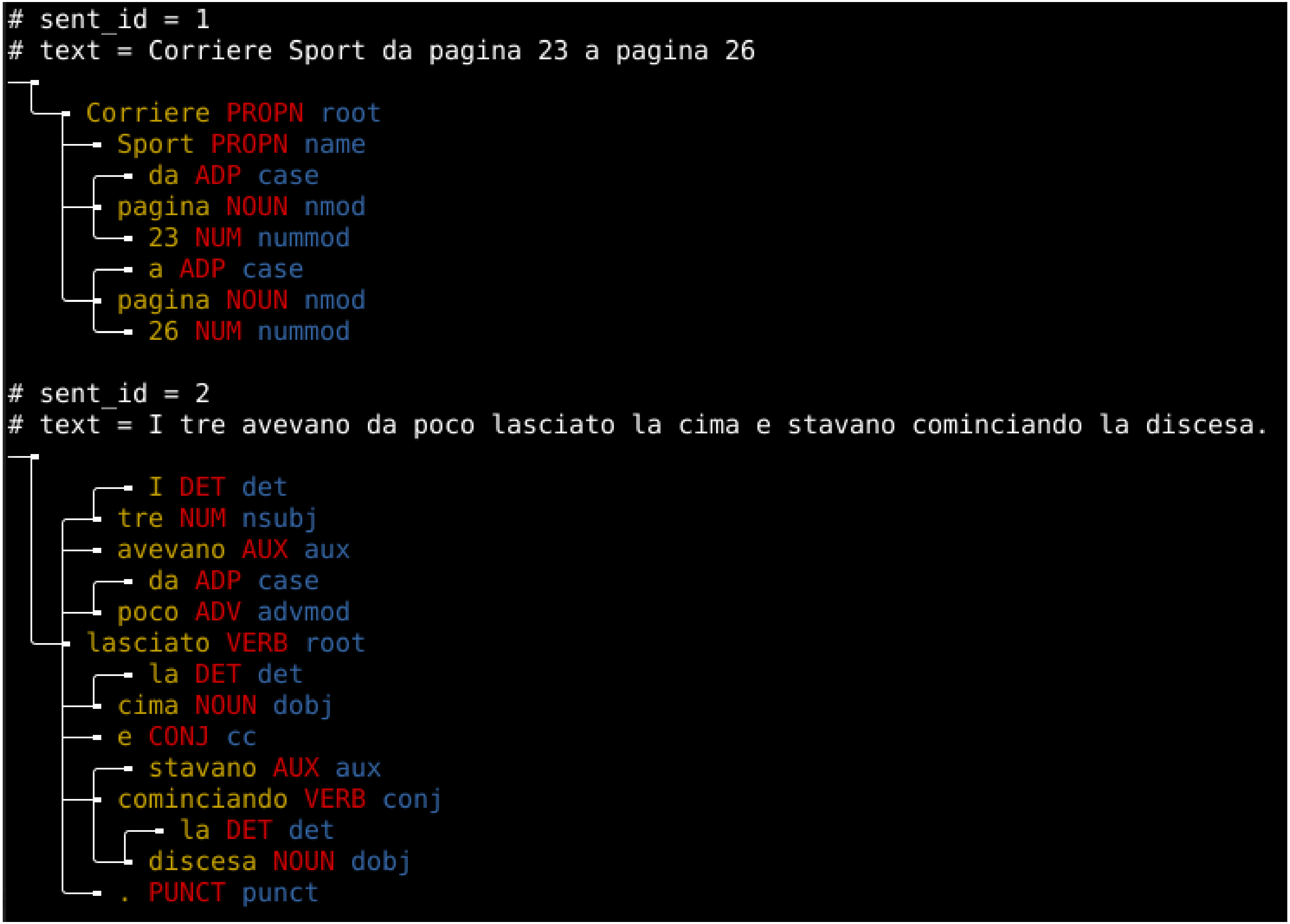
Tools for viewing trees

- **UDPipe** http://lindat.mff.cuni.cz/services/udpipe
- **PML-TQ** tree-query language, UD1.2

  https://lindat.mff.cuni.cz/services/pmltq/
- **Udapi** https://github.com/udapi/udapi-python
  - `udapy Write::HTML < my.conllu > my.html`

    demo: http://ufallab.ms.mff.cuni.cz/~popel/czeng1.6-sample.html
  - `udapy -HA < my.conllu > my.html`

    demo: http://ufallab.ms.mff.cuni.cz/~popel/sv/dev-bugs.html
  - `udapy -T < my.conllu | less -R`



```
# sent_id = 1
# text = Corriere Sport da pagina 23 a pagina 26
─┐
 └─┮ Corriere PROPN root
   ├─╼ Sport PROPN name
   │ ┌─╼ da ADP case
   ├─┾ pagina NOUN nmod
   │ └─╼ 23 NUM nummod
   │ ┌─╼ a ADP case
   └─┾ pagina NOUN nmod
     └─╼ 26 NUM nummod

# sent_id = 2
# text = I tre avevano da poco lasciato la cima e stavano cominciando la discesa.
─┐
 │   ┌─╼ I DET det
 │ ┌─┶ tre NUM nsubj
 │ ├─╼ avevano AUX aux
 │ │ ┌─╼ da ADP case
 │ ├─┶ poco ADV advmod
 └─┾ lasciato VERB root
   │ ┌─╼ la DET det
   ├─┶ cima NOUN dobj
   ├─╼ e CONJ cc
   │ ┌─╼ stavano AUX aux
   ├─┾ cominciando VERB conj
   │ │ ┌─╼ la DET det
   │ └─┶ discesa NOUN dobj
   └─╼ . PUNCT punct
```

Udapi – API+framework for UD

- Available in **Python**, Perl, Java
- History: Treex framework
  - Perl only, slow, XML, tectogrammatical support
  - Deep-syntactic MT for EN↔CS,PT,NL,ES,EU
- Goals:
  - Allows both fast prototyping and full applications
  - Both command-line tool (`udapy`) and library
  - Modularity, reusability, cooperation

Udapi use cases

- Format conversions (CoNLL-U, SDParse, PML)
- Transformations (UD v1 to v2, prepositions up…)
- Validity tests
- Querying
- Automatic parsing, evaluation,…

Hands-on tutorial

- http://ufal.mff.cuni.cz/~popel/udapi/index.html