



# Natural Language Processing

## Syntax

Joakim Nivre

Uppsala University  
Department of Linguistics and Philology  
[joakim.nivre@lingfil.uu.se](mailto:joakim.nivre@lingfil.uu.se)



# What is Syntax?

- ▶ Words do not occur in isolation – they combine into sentences

*Snoopy hugs Woodstock*

*Woodstock hugs Snoopy*

*Snoopy hugs birds*

*\*Birds hugs Snoopy*

*Birds hug Snoopy*

*Snoopy hugs them*

*\*Them hug Snoopy*

*They hug Snoopy*



- ▶ Syntax – the study of sentence structure



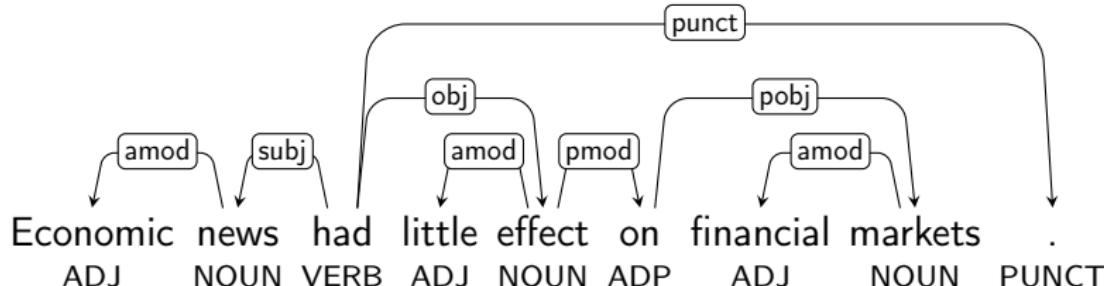
## Two Views of Syntax

- ▶ Dependency:
  - ▶ Syntactic structure resides in relations between words
  - ▶ Focus on the functional roles (**subject, object, ...**)
  - ▶ Long tradition in descriptive grammar going back to Antiquity
- ▶ Constituency:
  - ▶ Syntactic structure consists in the composition of phrases
  - ▶ Focus on structural categories (**noun phrase, verb phrase, ...**)
  - ▶ Introduced in structural linguistics in the 20th century



# Dependency

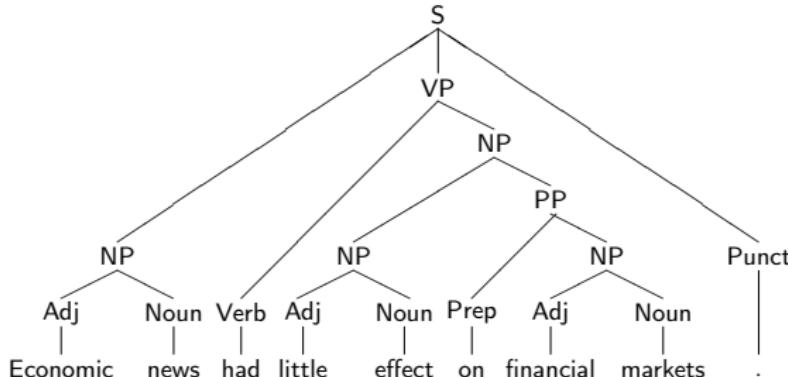
- ▶ Syntactic structure represented by **dependency trees**
  - ▶ Words represented by tree nodes
  - ▶ Dependencies represented by directed arcs between nodes
  - ▶ Functional roles specified by arc labels





# Constituency

- ▶ Syntactic structure represented by **phrase structure trees**
  - ▶ Words represented by terminal tree nodes (leaves)
  - ▶ Phrases represented by internal tree nodes
  - ▶ Phrase types specified by node labels
- ▶ Phrase structure trees can be defined by **context-free grammars**





## Two Complementary Views?

- ▶ Dependency trees explicitly represent
  - ▶ dependency relations (**directed arcs**)
  - ▶ functional categories (**arc labels**)
  - ▶ possibly some structural categories (parts-of-speech)
- ▶ Phrase structure trees explicitly represent
  - ▶ phrases (**internal nodes**)
  - ▶ structural categories (**node labels**)
  - ▶ possibly some functional categories (grammatical functions)
- ▶ Both are widely used in NLP – and hybrid representations exist



# Quiz

- ▶ True or false?
  1. Syntax studies the internal structure of words
  2. Syntax studies the order of words in sentences
  3. Dependency trees show how words are related in sentences
  4. Dependency trees show how sentences are composed of phrases



# Natural Language Processing

## Dependency

Joakim Nivre

Uppsala University  
Department of Linguistics and Philology  
[joakim.nivre@lingfil.uu.se](mailto:joakim.nivre@lingfil.uu.se)



# Dependency

- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**

Economic news had little effect on financial markets .  
ADJ NOUN VERB ADJ NOUN ADP ADJ NOUN PUNCT



# Dependency

- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**

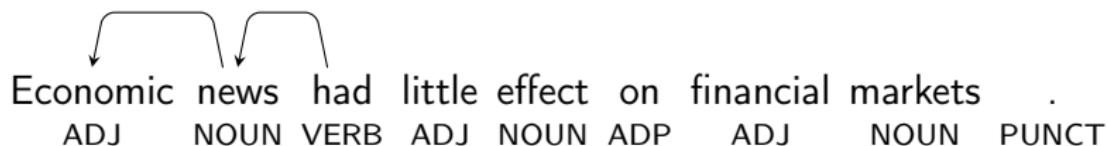
Economic news had little effect on financial markets .  
ADJ NOUN VERB ADJ NOUN ADP ADJ NOUN PUNCT





# Dependency

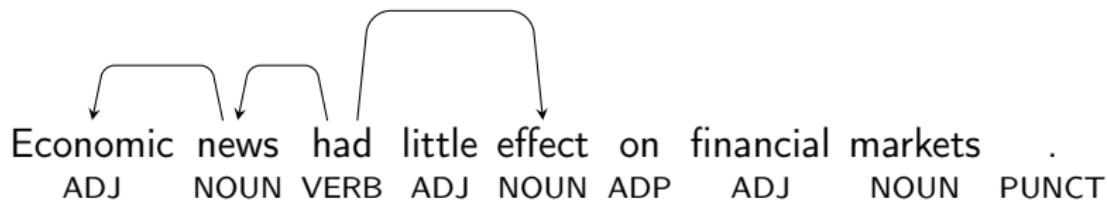
- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**





# Dependency

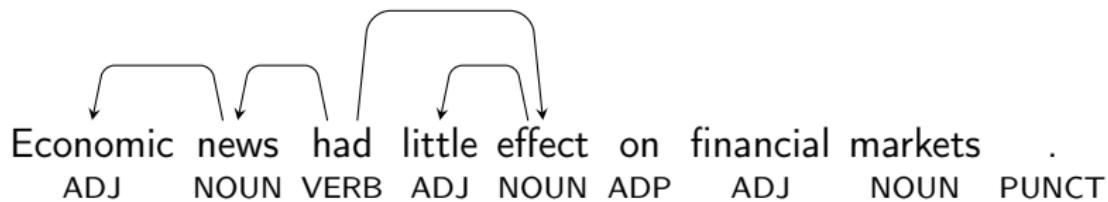
- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**





# Dependency

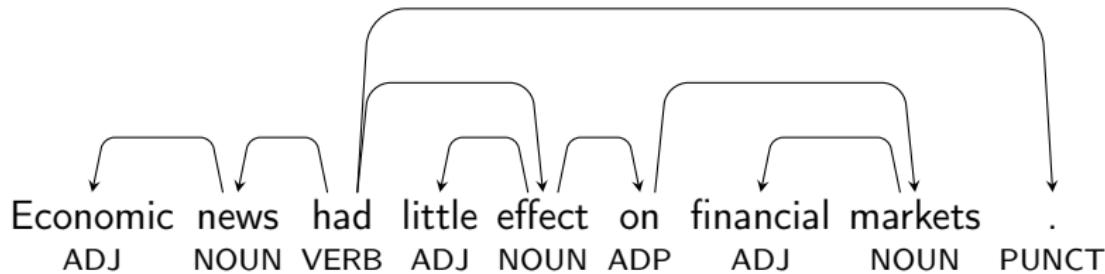
- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**





# Dependency

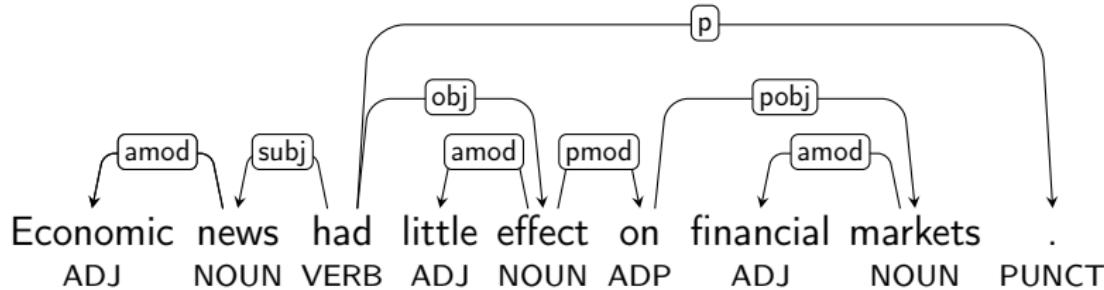
- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**





# Dependency

- ▶ Syntactic structure consists of **words**, related by binary, asymmetric relations called **dependencies**





# Terminology

Higher	Lower
Head	Dependent
Governor	Modifier
Regent	Subordinate
:	:



# Terminology

Higher	Lower
Head	Dependent
Governor	Modifier
Regent	Subordinate
:	:



# Criteria for Heads and Dependents

- ▶ Criteria for a syntactic relation between a head  $H$  and a dependent  $D$  in a construction  $C$ :
  1.  $H$  determines the syntactic category of  $C$ ;  $H$  can replace  $C$ .
  2.  $H$  determines the semantic category of  $C$ ;  $D$  specifies  $H$ .
  3.  $H$  is obligatory;  $D$  may be optional.
  4.  $H$  selects  $D$  and determines whether  $D$  is obligatory.
  5. The form of  $D$  depends on  $H$  (agreement or government).
  6. The linear position of  $D$  is specified with reference to  $H$ .
- ▶ Issues:
  - ▶ Syntactic (and morphological) versus semantic criteria
  - ▶ Exocentric versus endocentric constructions



## Some Clear Cases

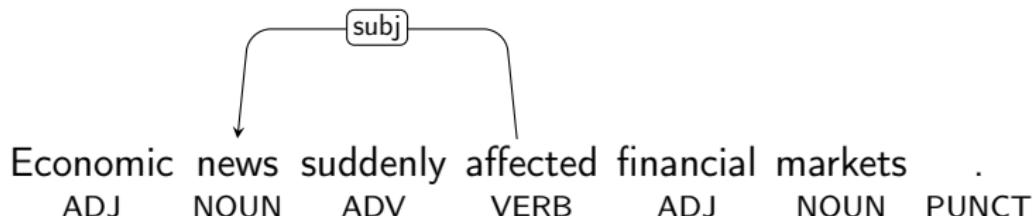
Construction	Head	Dependent
Exocentric	Verb	Subject ( <b>subj</b> )
	Verb	Object ( <b>obj</b> )
Endocentric	Verb	Adverbial ( <b>advmod</b> )
	Noun	Attribute ( <b>amod</b> )

Economic news suddenly affected financial markets .  
ADJ NOUN ADV VERB ADJ NOUN PUNCT



## Some Clear Cases

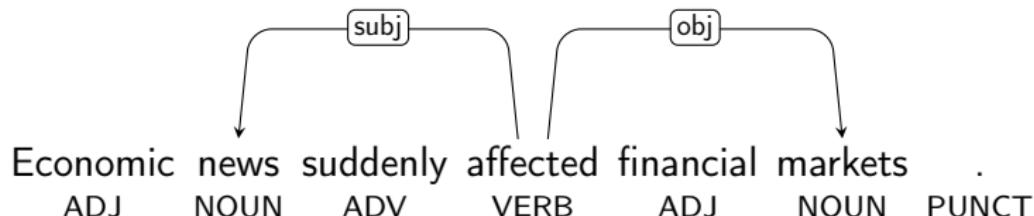
Construction	Head	Dependent
Exocentric	Verb	Subject ( <b>subj</b> )
	Verb	Object ( <b>obj</b> )
Endocentric	Verb	Adverbial ( <b>advmod</b> )
	Noun	Attribute ( <b>amod</b> )





## Some Clear Cases

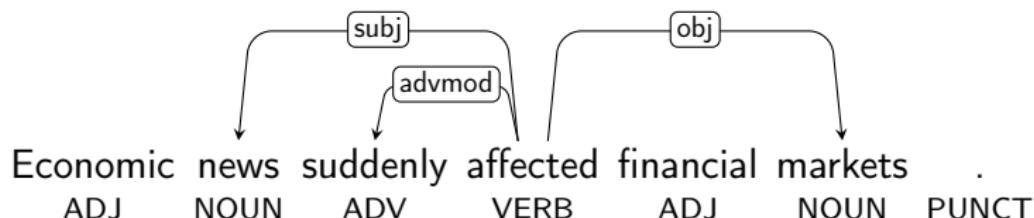
Construction	Head	Dependent
Exocentric	Verb	Subject ( <b>subj</b> )
	Verb	Object ( <b>obj</b> )
Endocentric	Verb	Adverbial ( <b>advmod</b> )
	Noun	Attribute ( <b>amod</b> )





## Some Clear Cases

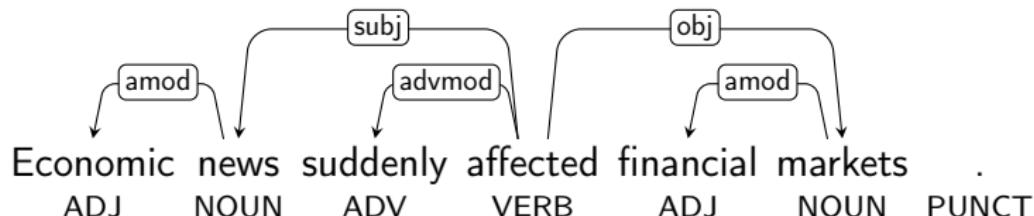
Construction	Head	Dependent
Exocentric	Verb	Subject ( <b>subj</b> )
	Verb	Object ( <b>obj</b> )
Endocentric	Verb	Adverbial ( <b>advmod</b> )
	Noun	Attribute ( <b>amod</b> )





## Some Clear Cases

Construction	Head	Dependent
Exocentric	Verb	Subject ( <b>subj</b> )
	Verb	Object ( <b>obj</b> )
Endocentric	Verb	Adverbial ( <b>advmod</b> )
	Noun	Attribute ( <b>amod</b> )





## Some Tricky Cases

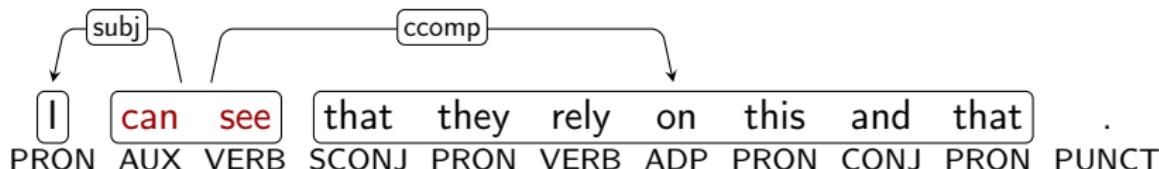
- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation

I can see that they rely on this and that .  
PRON AUX VERB SCONJ PRON VERB ADP PRON CONJ PRON PUNCT



## Some Tricky Cases

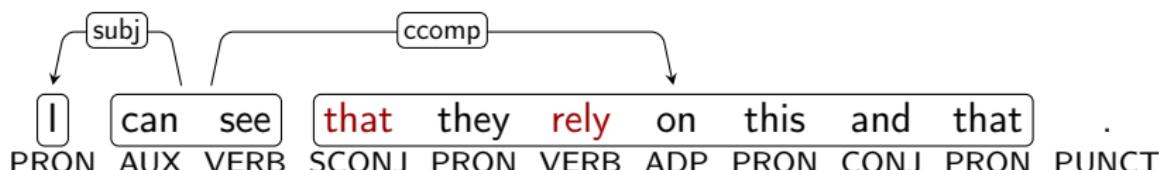
- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation





## Some Tricky Cases

- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation





## Some Tricky Cases

- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation

I can see that they rely on **this and that** .  
PRON AUX VERB SCONJ PRON VERB ADP PRON CONJ PRON PUNCT



## Some Tricky Cases

- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation

I can see that they rely on this and that .  
PRON AUX VERB SCONJ PRON VERB ADP PRON CONJ PRON PUNCT



## Some Tricky Cases

- ▶ Complex verb groups (auxiliary ↔ main verb)
- ▶ Subordinate clauses (complementizer ↔ verb)
- ▶ Coordination (coordinator ↔ conjuncts)
- ▶ Prepositional phrases (preposition ↔ nominal)
- ▶ Punctuation

I can see that they rely on this and that .

PRON AUX VERB SCONJ PRON VERB ADP PRON CONJ PRON PUNCT



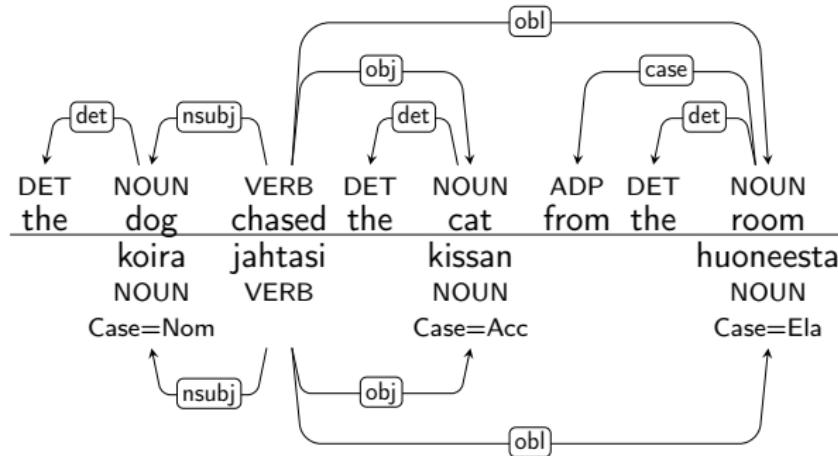
# Treebanks

- ▶ Treebanks
  - ▶ Syntactically annotated corpora are called **treebanks**
  - ▶ Treebanks can be used to train and evaluate **syntactic parsers**
- ▶ Dependency treebanks
  - ▶ Treebanks with dependency-based annotation
  - ▶ Example: Prague Dependency Treebank of Czech
  - ▶ Annotation schemes can vary considerably across languages



# Universal Dependencies (UD)

- ▶ Standardized framework for dependency annotation
- ▶ Consistent analysis across typologically different languages





# Quiz

The big bear scared the little dog.

- ▶ True or false
  - 1. The word **dog** is a dependent of the word **bear**
  - 2. The word **bear** is a dependent of the word **scared**
  - 3. The word **scared** is a dependent of the word **little**
  - 4. The word **little** is a dependent of the word **dog**



# Natural Language Processing

## Constituency

Joakim Nivre

Uppsala University  
Department of Linguistics and Philology  
[joakim.nivre@lingfil.uu.se](mailto:joakim.nivre@lingfil.uu.se)



# Constituency

- ▶ Word groups can act as single units

*Los Angeles*

*a high-class spot such as Mindy's  
three parties from Brooklyn  
they*

- ▶ Such groups of words are called **constituents**
- ▶ Constituents have similar internal structure and behavior



# Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.



# Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.

The girl in the corner + wears a yellow hat and dark sunglasses.



## Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.

The girl in the corner + wears a yellow hat and dark sunglasses.

[The girl + in the corner] [wears + a yellow hat and dark sunglasses].



# Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.

The girl in the corner + wears a yellow hat and dark sunglasses.

[The girl + in the corner] [wears + a yellow hat and dark sunglasses].

[The + girl] [in + the corner] wears [a yellow hat + and + dark sunglasses].



# Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.

The girl in the corner + wears a yellow hat and dark sunglasses.

[The girl + in the corner] [wears + a yellow hat and dark sunglasses].

[The + girl] [in + the corner] wears [a yellow hat + and + dark sunglasses].

The girl in [the + corner] wears [a + yellow hat] and [dark + sunglasses].



# Immediate Constituency Analysis

- ▶ We can find constituents by recursive decomposition:

The girl in the corner wears a yellow hat and dark sunglasses.

The girl in the corner + wears a yellow hat and dark sunglasses.

[The girl + in the corner] [wears + a yellow hat and dark sunglasses].

[The + girl] [in + the corner] wears [a yellow hat + and + dark sunglasses].

The girl in [the + corner] wears [a + yellow hat] and [dark + sunglasses].

The girl in the corner wears a [yellow + hat] and dark sunglasses.



## Test for Constituents

- ▶ Substitution: similar constituents can replace each other

*The girl in the corner wears a yellow hat and dark sunglasses*

*The girl wears a yellow hat and dark sunglasses*

*She wears a yellow hat and dark sunglasses*

*\*In the corner wears a yellow hat and dark sunglasses*

- ▶ Movement: words in a constituent move together

*She went to Paris on Monday.*

*On Monday she went to Paris.*

*\*Paris on Monday she went to.*

*\*On she went to Paris Monday.*



# Constituent Types

- ▶ Noun phrase (NP)

*she*

*the house*

*Robin Hood and his merry men*

- ▶ Verb phrase (VP)

*blushed*

*loves Mary*

*was told to sit down and be quiet*

- ▶ Prepositional phrase (PP)

*on it*

*with the telescope*

*through the foggy dew*



# Context-Free Grammar

- ▶ A context-free grammar (CFG) consists of
  - ▶ a finite set of **nonterminal** symbols
  - ▶ a finite set of **terminal** symbols
  - ▶ a distinguished nonterminal symbol **S** (for Start)
  - ▶ a finite set of **rules** of the form

$$A \rightarrow \alpha$$

where  $A$  is a nonterminal and  $\alpha$  is a (possibly empty) sequence of nonterminal and terminal symbols



## Example Grammar

S	→	NP VP Punct	Verb	→	had
VP	→	VP PP	Noun	→	news
VP	→	Verb NP	Noun	→	effect
NP	→	NP PP	Noun	→	markets
NP	→	Adj Noun	Adj	→	Economic
PP	→	Prep NP	Adj	→	little
			Adj	→	financial
			Prep	→	on
			Punct	→	.



## Example Grammar

### Grammar

S → NP VP Punct  
VP → VP PP  
VP → Verb NP  
NP → NP PP  
NP → Adj Noun  
PP → Prep NP

### Lexicon

Verb → had  
Noun → news  
Noun → effect  
Noun → markets  
Adj → Economic  
Adj → little  
Adj → financial  
Prep → on  
Punct → .



# Derivations

S



# Derivations

S  
NP VP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct  
Economic Noun VP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct  
Economic Noun VP Punct  
Economic news VP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct  
Economic Noun VP Punct  
Economic news VP Punct  
Economic news Verb NP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct  
Economic Noun VP Punct  
Economic news VP Punct  
Economic news Verb NP Punct  
Economic news had NP Punct



# Derivations

S  
NP VP Punct  
Adj Noun VP Punct  
Economic Noun VP Punct  
Economic news VP Punct  
Economic news Verb NP Punct  
Economic news had NP Punct  
Economic news had NP PP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct

Economic news had little effect on NP Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct

Economic news had little effect on NP Punct

Economic news had little effect on Adj Noun Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct

Economic news had little effect on NP Punct

Economic news had little effect on Adj Noun Punct

Economic news had little effect on financial Noun Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct

Economic news had little effect on NP Punct

Economic news had little effect on Adj Noun Punct

Economic news had little effect on financial Noun Punct

Economic news had little effect on financial markets Punct



# Derivations

S

NP VP Punct

Adj Noun VP Punct

Economic Noun VP Punct

Economic news VP Punct

Economic news Verb NP Punct

Economic news had NP Punct

Economic news had NP PP Punct

Economic news had Adj Noun PP Punct

Economic news had little Noun PP Punct

Economic news had little effect PP Punct

Economic news had little effect Prep NP Punct

Economic news had little effect on NP Punct

Economic news had little effect on Adj Noun Punct

Economic news had little effect on financial Noun Punct

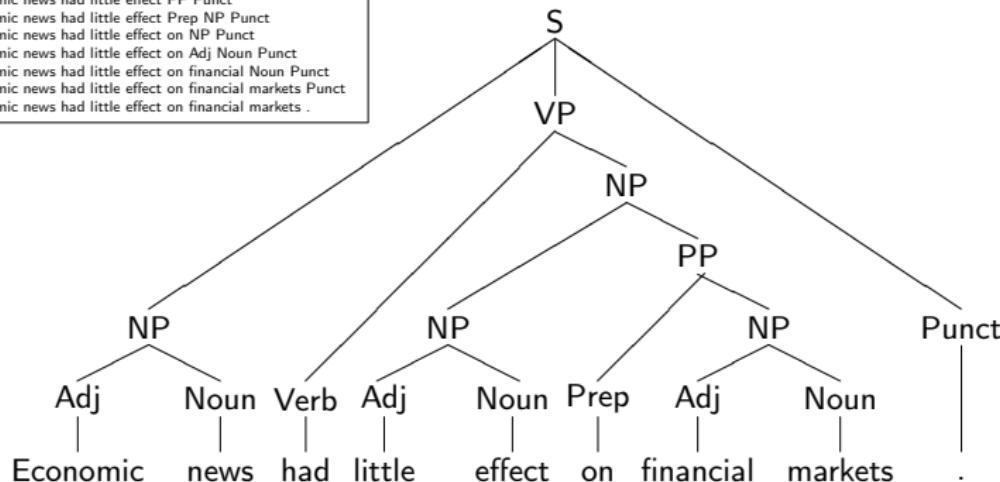
Economic news had little effect on financial markets Punct

Economic news had little effect on financial markets .



# Phrase Structure Trees

```
S
NP VP Punct
Adj Noun VP Punct
Economic Noun VP Punct
Economic news VP Punct
Economic news Verb NP Punct
Economic news had NP Punct
Economic news had NP PP Punct
Economic news had Adj Noun PP Punct
Economic news had little Noun PP Punct
Economic news had little effect PP Punct
Economic news had little effect Prep NP Punct
Economic news had little effect on NP Punct
Economic news had little effect on Adj Noun Punct
Economic news had little effect on financial Noun Punct
Economic news had little effect on financial markets Punct
Economic news had little effect on financial markets .
```





# Treebanks and Grammars

- ▶ Constituency treebanks
  - ▶ Treebanks with constituency-based annotation
  - ▶ Example: Penn Treebank of English
- ▶ Treebank grammars
  - ▶ We can extract CFGs from constituency treebanks
  - ▶ Treebank grammars can be used to build syntactic parsers



# Quiz

The big bear scared the little dog.

- ▶ True or false
  - 1. The substring **dog** is a noun phrase
  - 2. The substring **little dog** is a noun phrase
  - 3. The substring **the little dog** is a noun phrase
  - 4. The substring **scared the little dog** is a noun phrase
  - 5. The substring **scared the little dog** is a verb phrase



# Natural Language Processing

## Universal Dependencies

Joakim Nivre

Uppsala University  
Department of Linguistics and Philology  
[joakim.nivre@lingfil.uu.se](mailto:joakim.nivre@lingfil.uu.se)



# Universal Dependencies (UD)

- ▶ Framework for multilingual grammatical annotation
- ▶ Morphological layer:
  - ▶ Part-of-speech tags
  - ▶ Morphological features
  - ▶ Lemmas
- ▶ Syntactic layer:
  - ▶ Dependency tree



# Predicates and Arguments

nsubj  
obj  
iobj

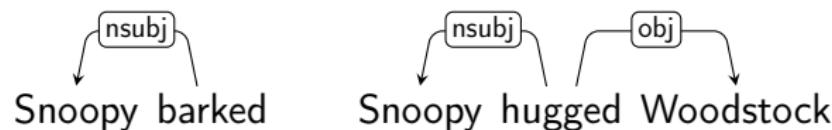
nsubj  
Snoopy barked

A bracket labeled "nsubj" is drawn above the word "Snoopy" in the sentence "Snoopy barked".



# Predicates and Arguments

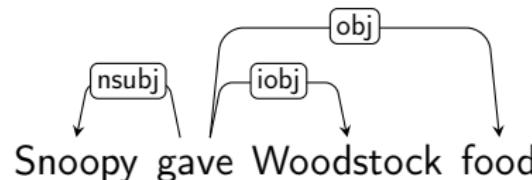
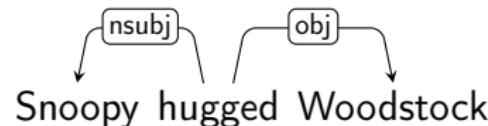
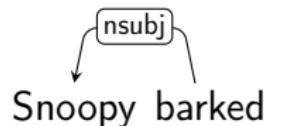
nsubj  
obj  
iobj





# Predicates and Arguments

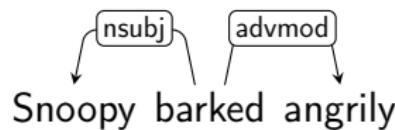
nsubj  
obj  
iobj





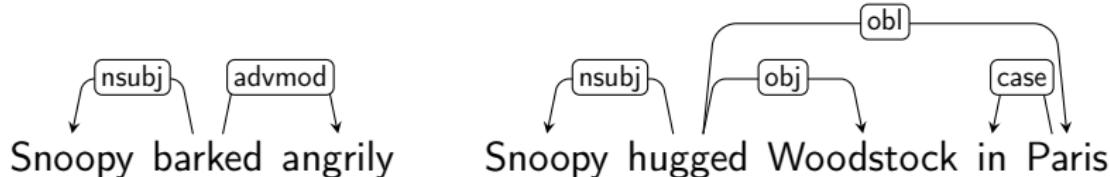
# Predicates and Modifiers

advmmod  
obl





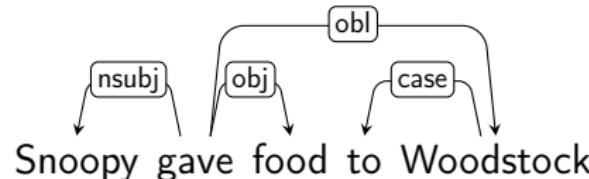
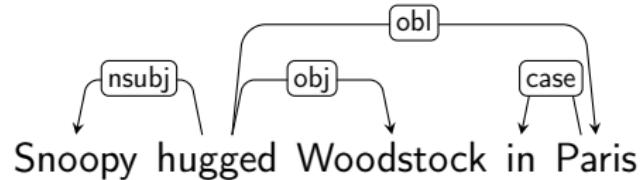
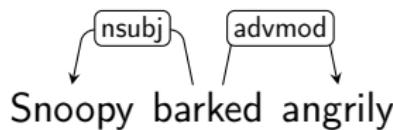
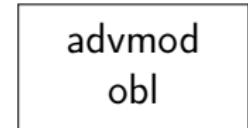
# Predicates and Modifiers



advmmod  
obl



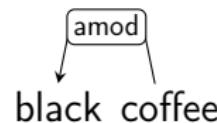
# Predicates and Modifiers





# Noun Phrases

amod  
nmod





# Noun Phrases

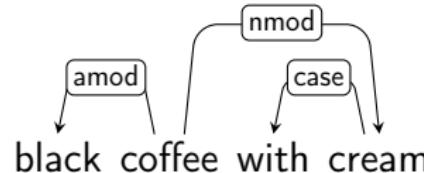
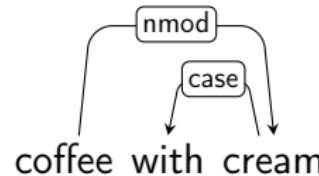
amod  
nmod





# Noun Phrases

amod  
nmod





# Quiz

Snoopy likes tasty bagels

- ▶ Which words are dependents of **likes**?
  1. Snoopy
  2. likes
  3. tasty
  4. bagels



# Quiz

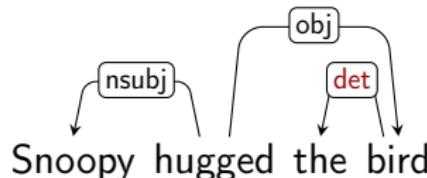
Snoopy likes tasty bagels

- ▶ Which dependency relation does **bagels** have to its head?
  1. nsubj
  2. obj
  3. obl
  4. nmod



# Function Words

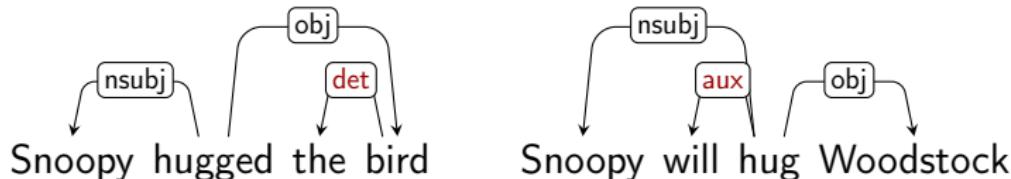
aux  
case  
det





# Function Words

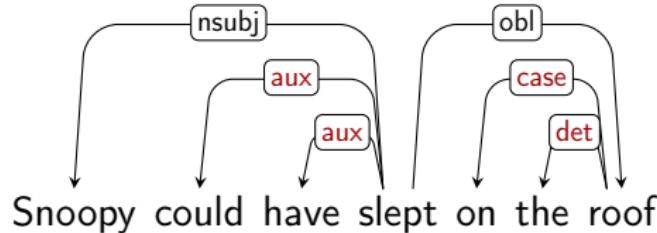
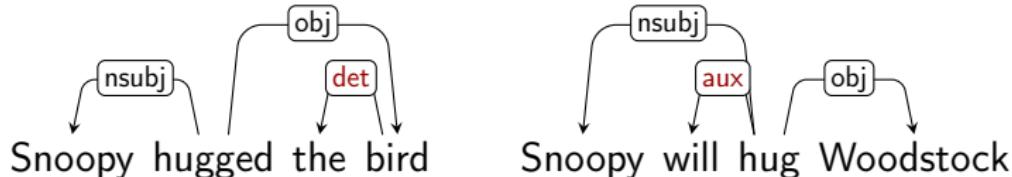
aux  
case  
det





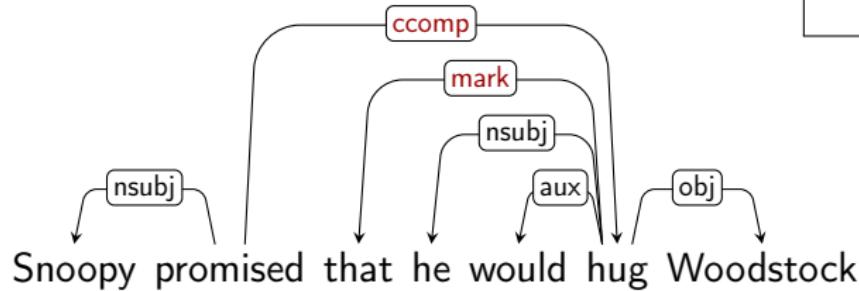
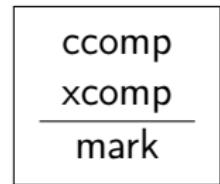
# Function Words

aux  
case  
det



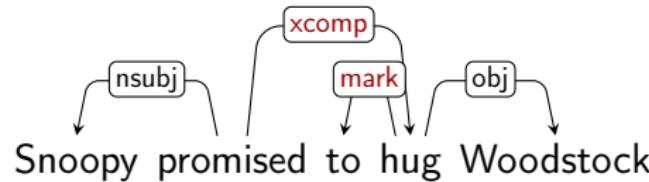
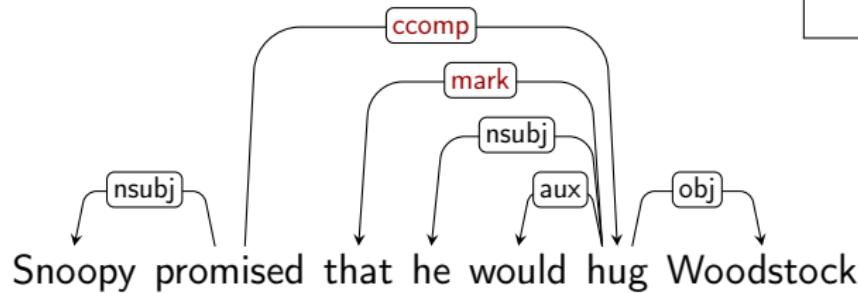
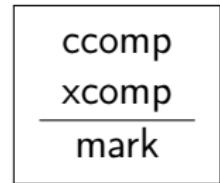


# Subordinate Clauses



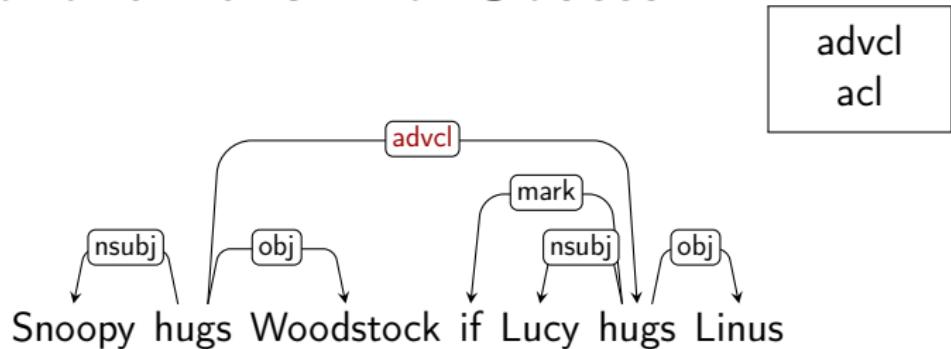


# Subordinate Clauses



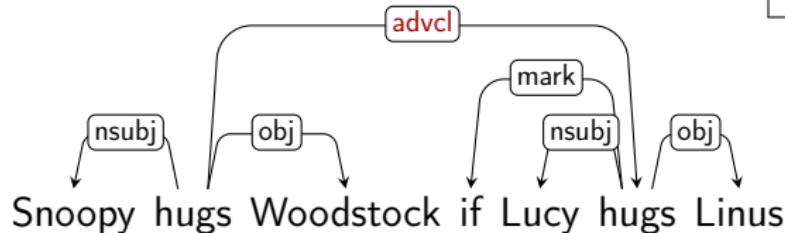


# Adverbial and Adnominal Clauses

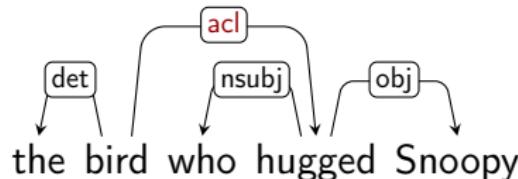




# Adverbial and Adnominal Clauses

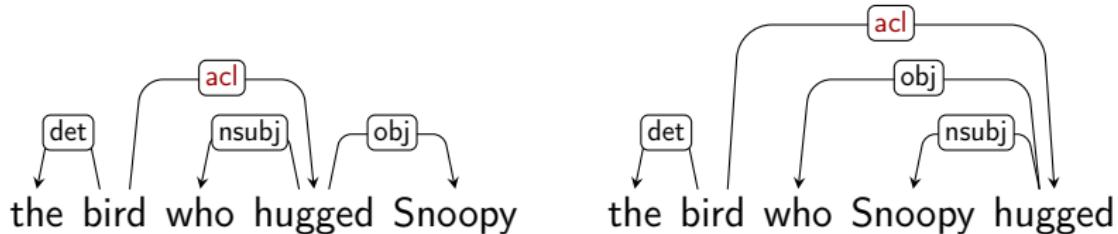
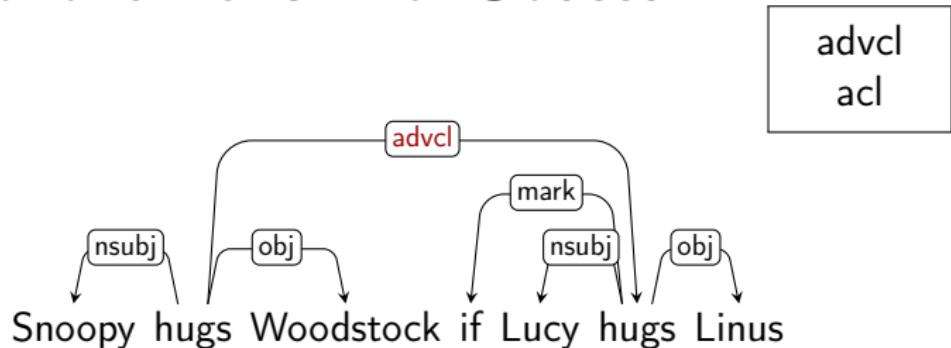


advcl  
acl



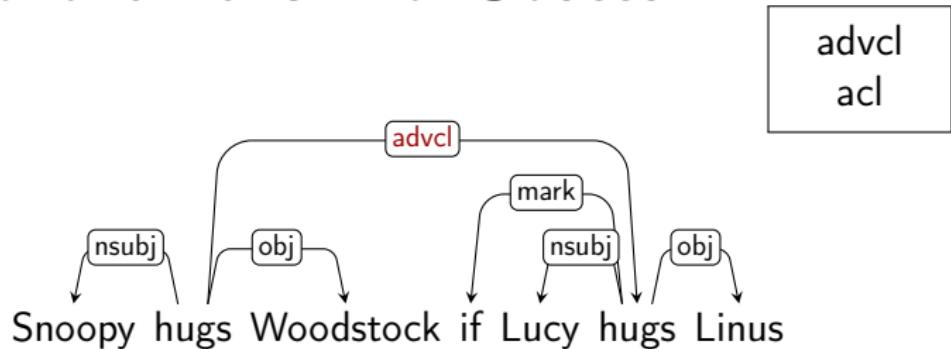


# Adverbial and Adnominal Clauses

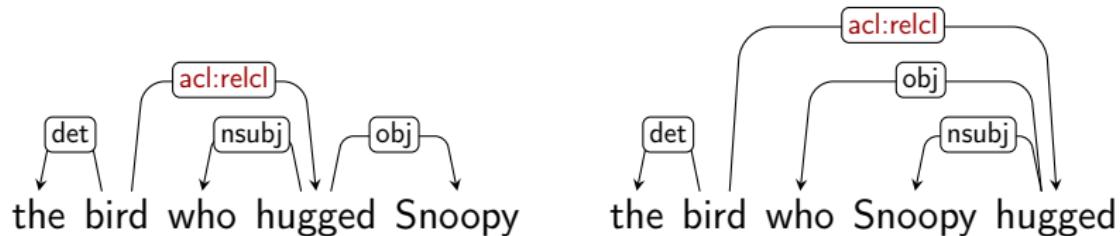




# Adverbial and Adnominal Clauses



advcl  
acl

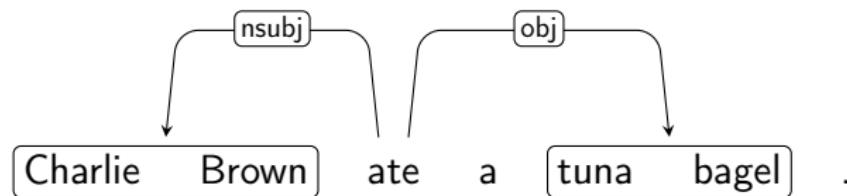


acl:relcl



## Other Relations

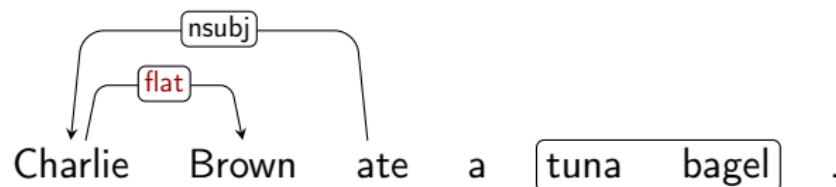
flat  
compound  
punct





## Other Relations

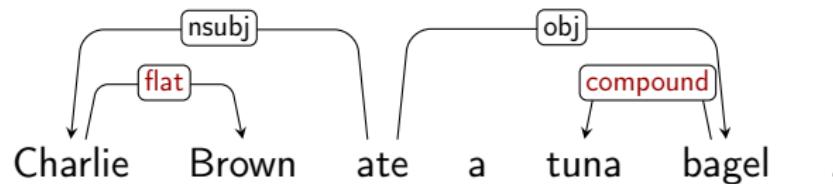
flat  
compound  
punct





## Other Relations

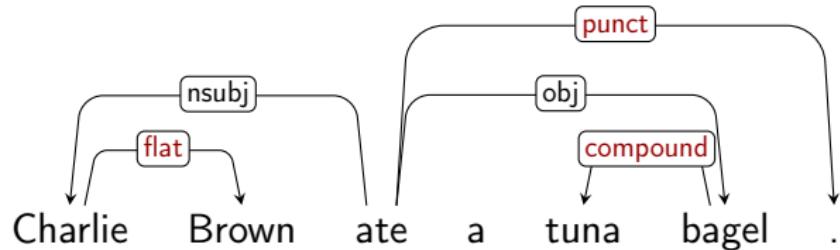
flat  
compound  
punct





## Other Relations

flat  
compound  
punct





## Quiz

Snoopy could have tried to hug Woodstock

- ▶ Which verb is the root of the dependency tree?
  1. could
  2. have
  3. tried
  4. hug



## Quiz

Snoopy could have tried to hug Woodstock

- ▶ Which verbs have the dependency relation **aux** to its head?
  1. could
  2. have
  3. tried
  4. hug