

# L95: Natural Language Syntax and Parsing

## 8) Unification-based Grammars and Parsing

Paula Buttery

Dept of Computer Science & Technology, University of Cambridge

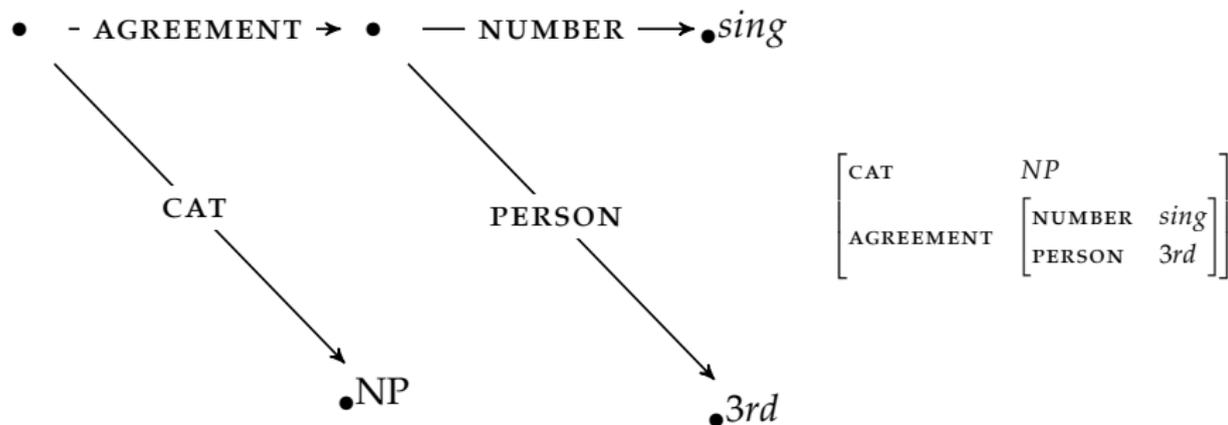
## Reminder...

Last time we looked at lexicalisation and features to help us with:

- modelling structural dependency across the tree as a whole
  - e.g. correctly modelling *NP* expansion
- modelling the structural behaviour specific to a lexical item:
  - pp-attachment
  - subcategorisation
  - co-ordination

# Alternative approach represents features in **DAGs**

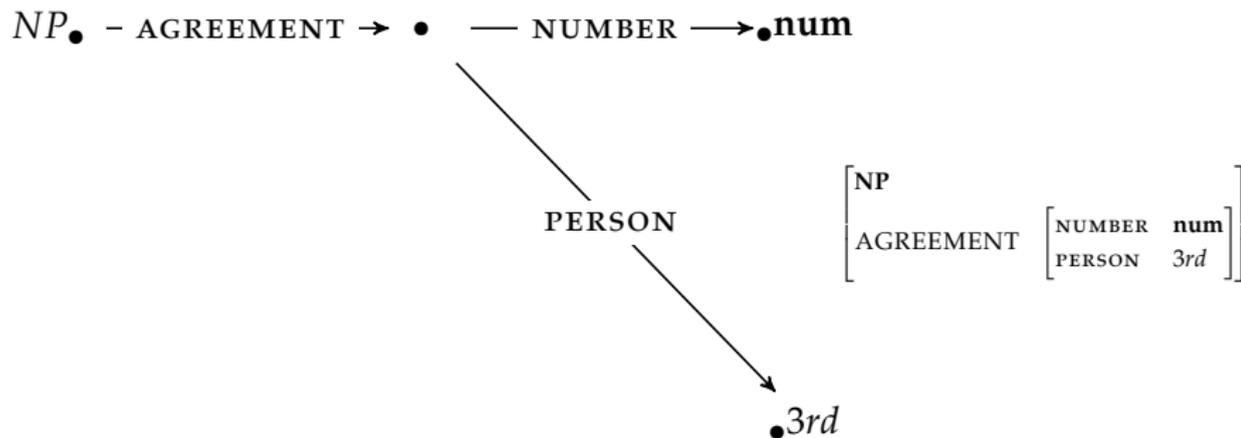
Re-conceptualise words, non-terminal nodes and parses as **Directed Acyclic Graphs** which may be represented as **Attribute Value Matrices**



We have **atomic values** at each of the terminal nodes and another **AVM/DAG** at all other nodes

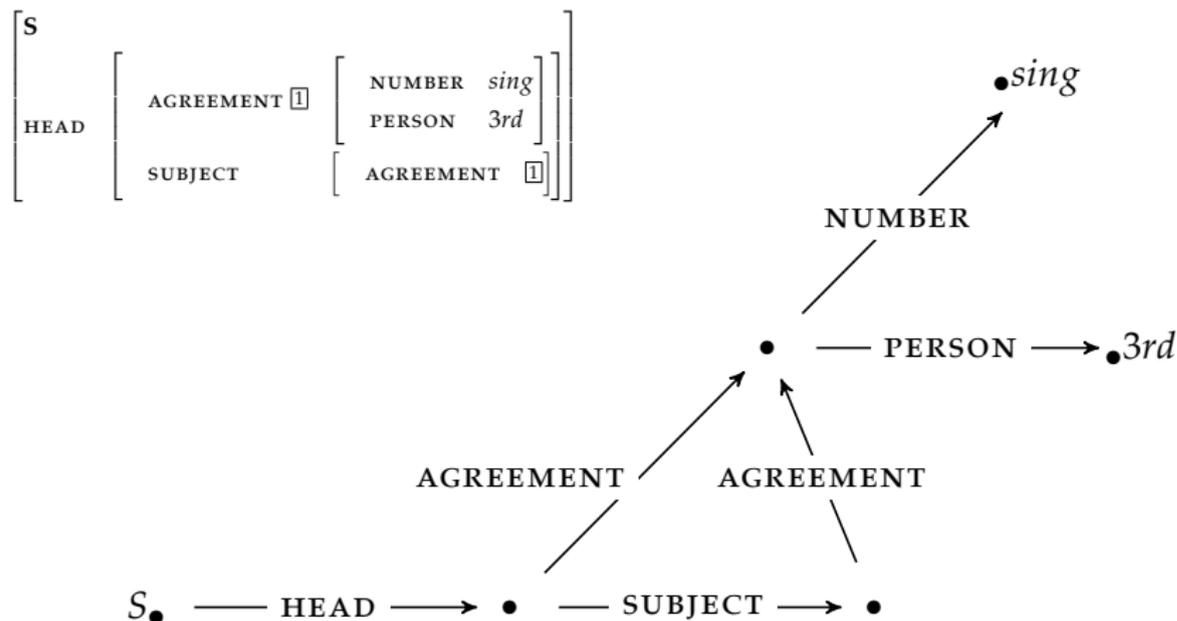
## Some grammars allow the **AVMs** to be **typed**

Typing facilitates grammar building. Hierarchies of AVM types can be used to automatically populate attributes



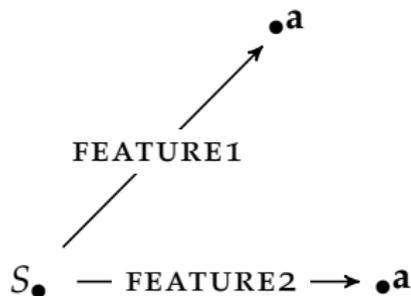
An shorthand notation uses angle bracket notation to indicate attribute paths: e.g. **<NP AGREEMENT PERSON>** would represent the attribute path leading to the atomic value *3rd*.

# DAGs and AVMs may exhibit re-entrancy

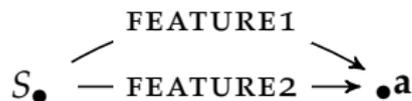


# DAGs and AVMs may exhibit re-entrancy

1. Non re-entrant:  $\begin{bmatrix} \text{FEATURE1} & \mathbf{a} \\ \text{FEATURE2} & \mathbf{a} \end{bmatrix}$



2. Re-entrant:  $\begin{bmatrix} \text{FEATURE1} & \boxed{1} & \mathbf{a} \\ \text{FEATURE2} & \boxed{1} & \end{bmatrix}$



## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\left[ \begin{array}{l} \text{PERSON} \\ 3rd \end{array} \right] \sqcup \left[ \begin{array}{l} \text{NUMBER} \\ plural \end{array} \right] = \left[ \begin{array}{l} \text{PERSON} \\ 3rd \\ \text{NUMBER} \\ plural \end{array} \right]$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\begin{bmatrix} \text{PERSON} & 3rd \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & plural \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & plural \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \mathbf{num} \end{bmatrix} =$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & \text{plural} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \text{plural} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & \text{plural} \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & \text{plural} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \mathbf{num} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & \text{plural} \end{bmatrix}$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & \text{plural} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \text{plural} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & \text{plural} \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & \text{plural} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \mathbf{num} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & \text{plural} \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & \text{sing} \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \text{plural} \end{bmatrix} =$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & plural \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & plural \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \mathbf{num} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & plural \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & sing \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & plural \end{bmatrix} = \text{unification fails}$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\begin{bmatrix} \text{PERSON} & 3rd \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & plural \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & plural \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & \mathbf{num} \end{bmatrix} = \begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & plural \end{bmatrix}$$

$$\begin{bmatrix} \text{PERSON} & 1st \\ \text{NUMBER} & sing \end{bmatrix} \sqcup \begin{bmatrix} \text{NUMBER} & plural \end{bmatrix} = \text{unification fails}$$

$$\begin{bmatrix} \text{FEATURE1} & \begin{bmatrix} \text{FEATURE2} & \boxed{1} \end{bmatrix} \\ \text{FEATURE3} & \boxed{1} \end{bmatrix} \sqcup \begin{bmatrix} \text{FEATURE3} & a \end{bmatrix}$$

## Parsing with DAGs involves **Unification**

- The unification of two DAGs is the most specific DAG which contains all the information in both of the original attribute-value structures.
- Unification fails if the two DAGs contain conflicting information.

$$\left[ \begin{array}{l} \text{PERSON} \quad 3rd \\ \text{NUMBER} \end{array} \right] \sqcup \left[ \begin{array}{l} \text{NUMBER} \quad plural \end{array} \right] = \left[ \begin{array}{l} \text{PERSON} \quad 3rd \\ \text{NUMBER} \quad plural \end{array} \right]$$

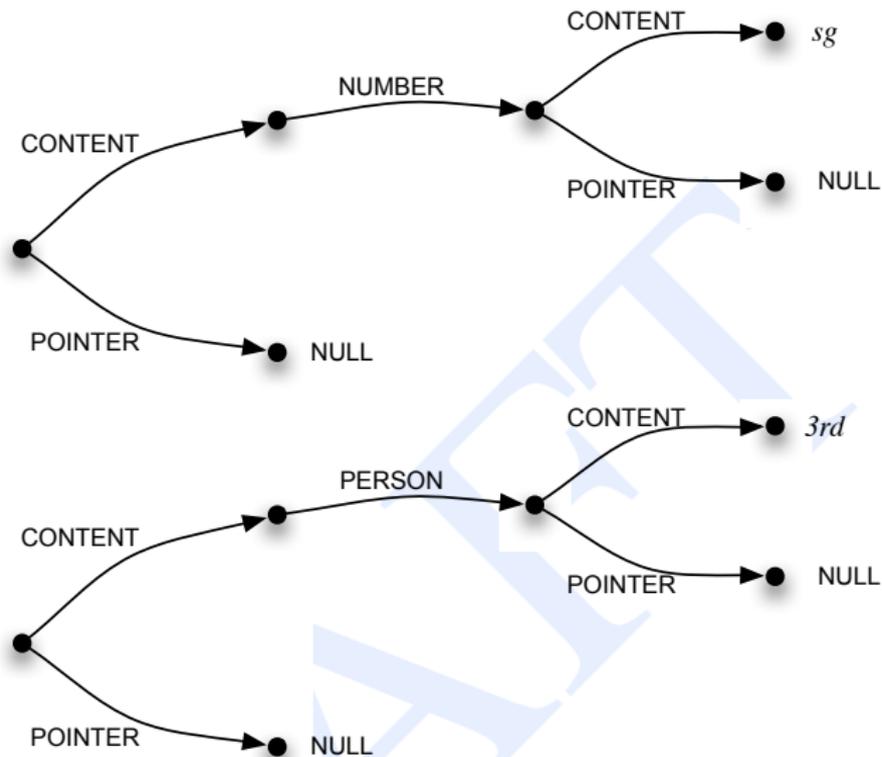
$$\left[ \begin{array}{l} \text{PERSON} \quad 1st \\ \text{NUMBER} \quad plural \end{array} \right] \sqcup \left[ \begin{array}{l} \text{NUMBER} \quad \mathbf{num} \end{array} \right] = \left[ \begin{array}{l} \text{PERSON} \quad 1st \\ \text{NUMBER} \quad plural \end{array} \right]$$

$$\left[ \begin{array}{l} \text{PERSON} \quad 1st \\ \text{NUMBER} \quad sing \end{array} \right] \sqcup \left[ \begin{array}{l} \text{NUMBER} \quad plural \end{array} \right] = \text{unification fails}$$

$$\left[ \begin{array}{l} \text{FEATURE1} \quad \left[ \begin{array}{l} \text{FEATURE2} \quad \boxed{1} \end{array} \right] \\ \text{FEATURE3} \quad \boxed{1} \end{array} \right] \sqcup \left[ \begin{array}{l} \text{FEATURE3} \quad a \end{array} \right] = \left[ \begin{array}{l} \text{FEATURE1} \quad \left[ \begin{array}{l} \text{FEATURE2} \quad \boxed{1} \quad a \end{array} \right] \\ \text{FEATURE3} \quad \boxed{1} \quad a \end{array} \right]$$

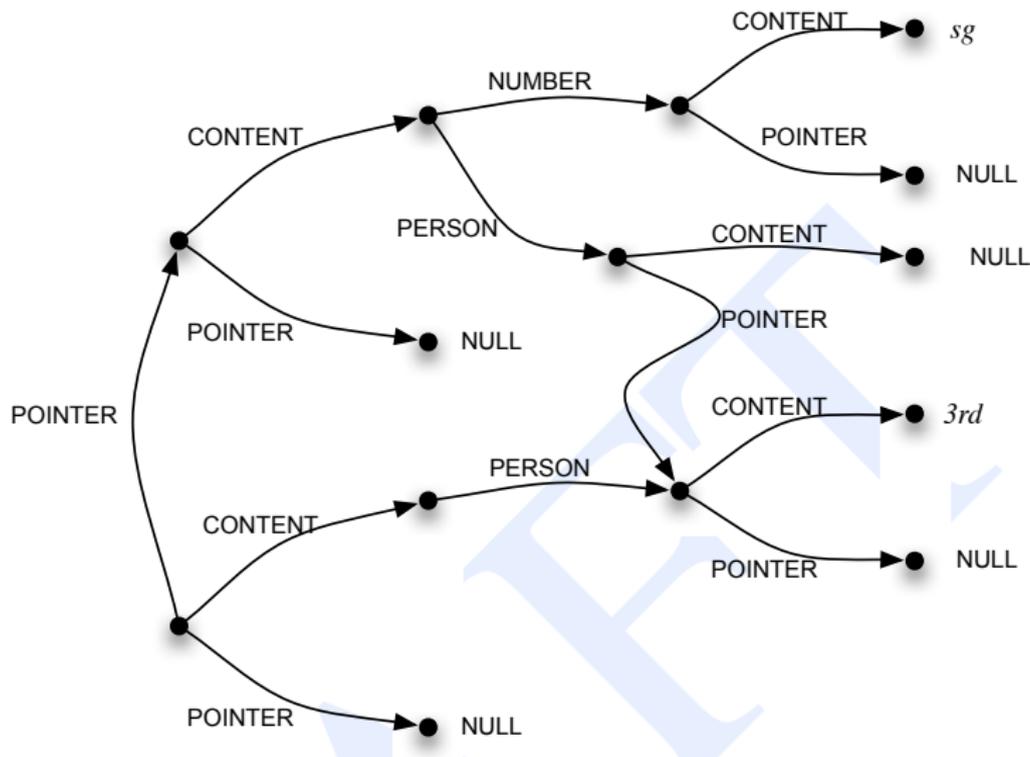
## Unification examples...

# Unification **algorithm** requires **extra** graph structure



From Jurafsky and Martin version 2

# Unification algorithm requires **extra** graph structure



From Jurafsky and Martin version 2

# DAGs can be straightforwardly associated with the lexicon

$\left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \end{array} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{array} \right] \right] \rightarrow \{\text{fish, rivers, pools, they}\}$

$V \rightarrow \{\text{cans, fishes}\}$   
 $\langle V \text{ AGREEMENT PERSON} \rangle = 3rd$   
 $\langle V \text{ AGREEMENT NUMBER} \rangle = sing$

$\left\langle \text{they,} \left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \end{array} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & sing \end{array} \right] \right] \right\rangle$

# DAGs can be straightforwardly associated with the lexicon

$\left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{array} \right] \end{array} \right] \rightarrow \{\text{fish, rivers, pools, they}\}$

$\mathbf{V} \rightarrow \{\text{cans, fishes}\}$   
 $\langle \mathbf{V} \text{ AGREEMENT PERSON} \rangle = 3rd$   
 $\langle \mathbf{V} \text{ AGREEMENT NUMBER} \rangle = sing$

$\left\langle \text{they,} \left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & sing \end{array} \right] \end{array} \right] \right\rangle$

# DAGs can be straightforwardly associated with the lexicon

$\left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & plural \end{array} \right] \end{array} \right] \rightarrow \{\text{fish, rivers, pools, they}\}$

$V \rightarrow \{\text{cans, fishes}\}$   
 $\langle V \text{ AGREEMENT PERSON} \rangle = 3rd$   
 $\langle V \text{ AGREEMENT NUMBER} \rangle = sing$

$\left\langle \text{they}, \left[ \begin{array}{l} \mathbf{N} \\ \text{AGREEMENT} \left[ \begin{array}{ll} \text{PERSON} & 3rd \\ \text{NUMBER} & sing \end{array} \right] \end{array} \right] \right\rangle$

## We can **modify** CFG algorithms to **parse** with DAGs

- We can use any CFG parsing algorithm if:
  - associate attribute paths with CFG rules
  - unify DAGs in the states

$S \rightarrow NP VP$

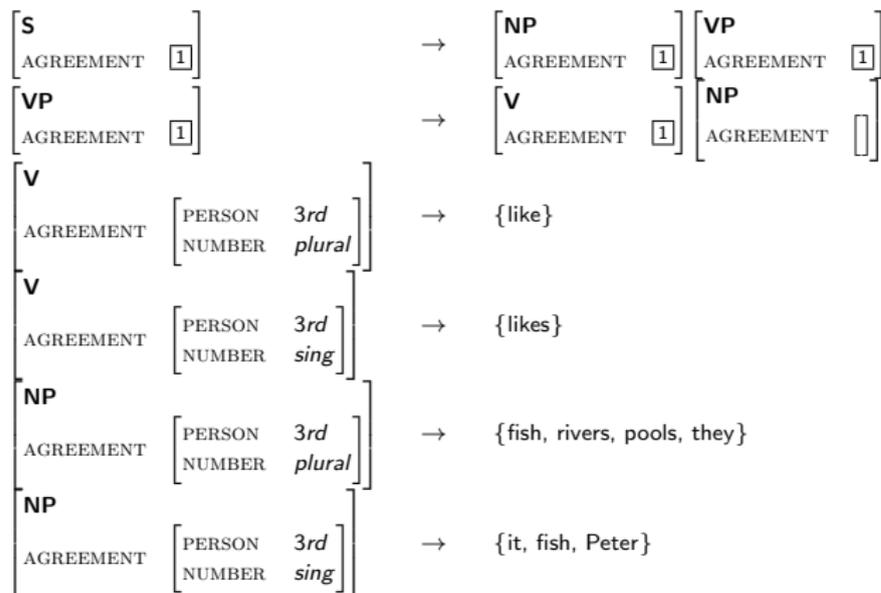
$\langle NP \text{ HEAD AGREEMENT} \rangle = \langle VP \text{ HEAD AGREEMENT} \rangle$

$\langle S \text{ HEAD} \rangle = \langle VP \text{ HEAD} \rangle$

- We would have items like  $[X, [a, b], DAG]$  on the agenda or at each cell

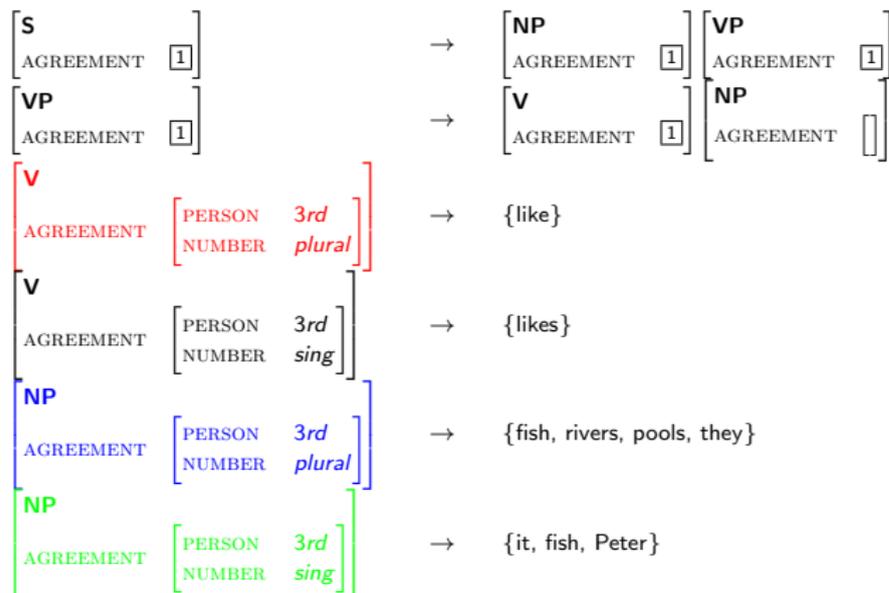
# Parsing example...

They like Peter



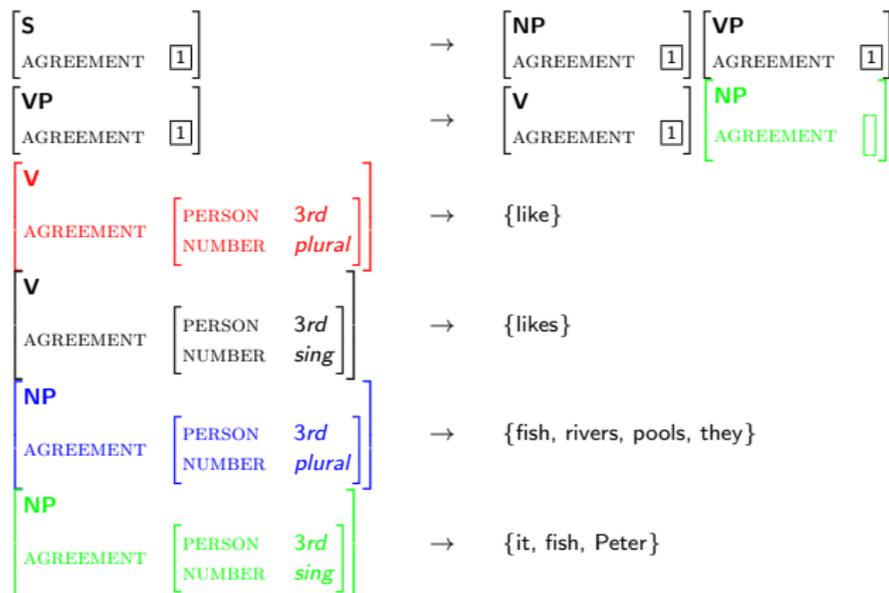
# Parsing example...

They like Peter



# Parsing example...

They like Peter



# Parsing example...

They like Peter

[S  
AGREEMENT 1]

→ [NP AGREEMENT 1] [VP AGREEMENT 1]

[VP  
AGREEMENT 1]

→ [V AGREEMENT 1] [NP AGREEMENT [PERSON 3rd  
NUMBER sing]]

[V  
AGREEMENT [PERSON 3rd  
NUMBER plural]]

→ {like}

[V  
AGREEMENT [PERSON 3rd  
NUMBER sing]]

→ {likes}

[NP  
AGREEMENT [PERSON 3rd  
NUMBER plural]]

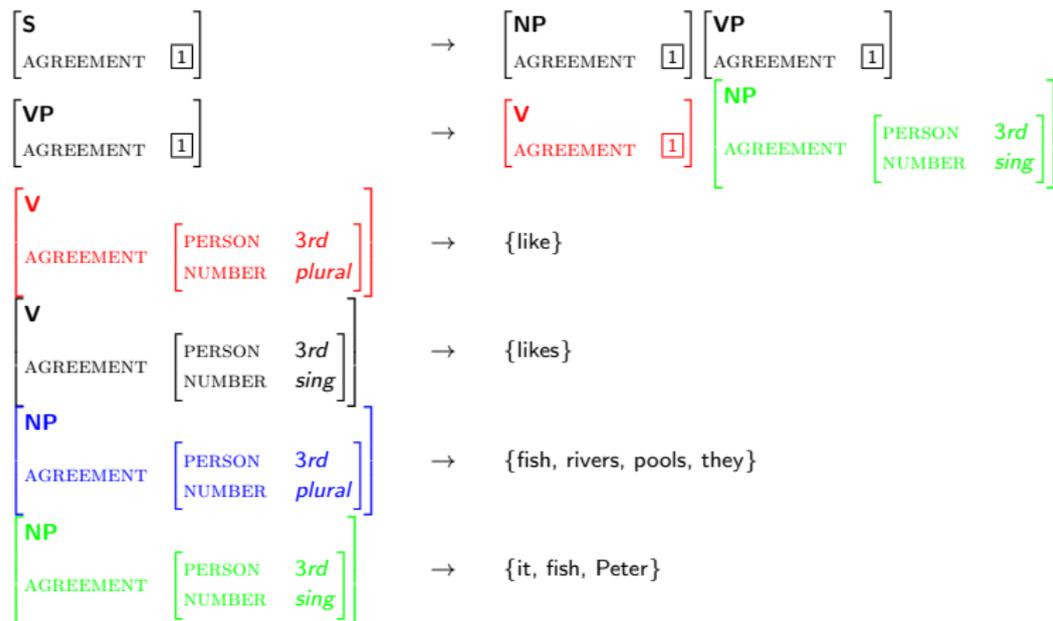
→ {fish, rivers, pools, they}

[NP  
AGREEMENT [PERSON 3rd  
NUMBER sing]]

→ {it, fish, Peter}

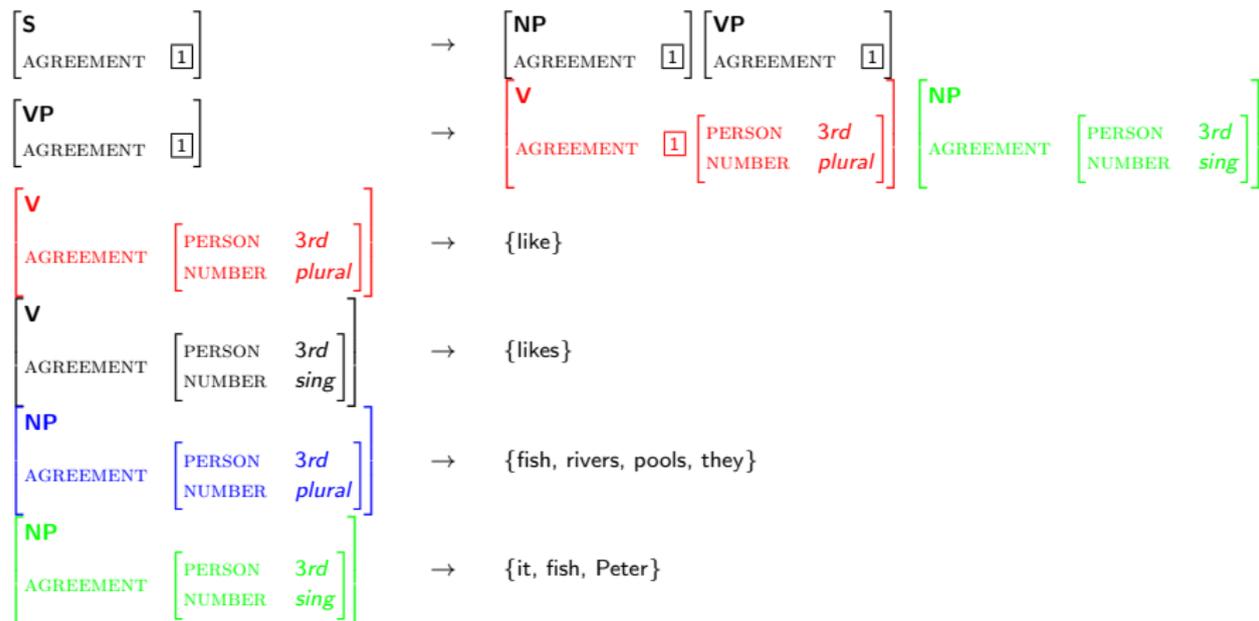
# Parsing example...

They like Peter



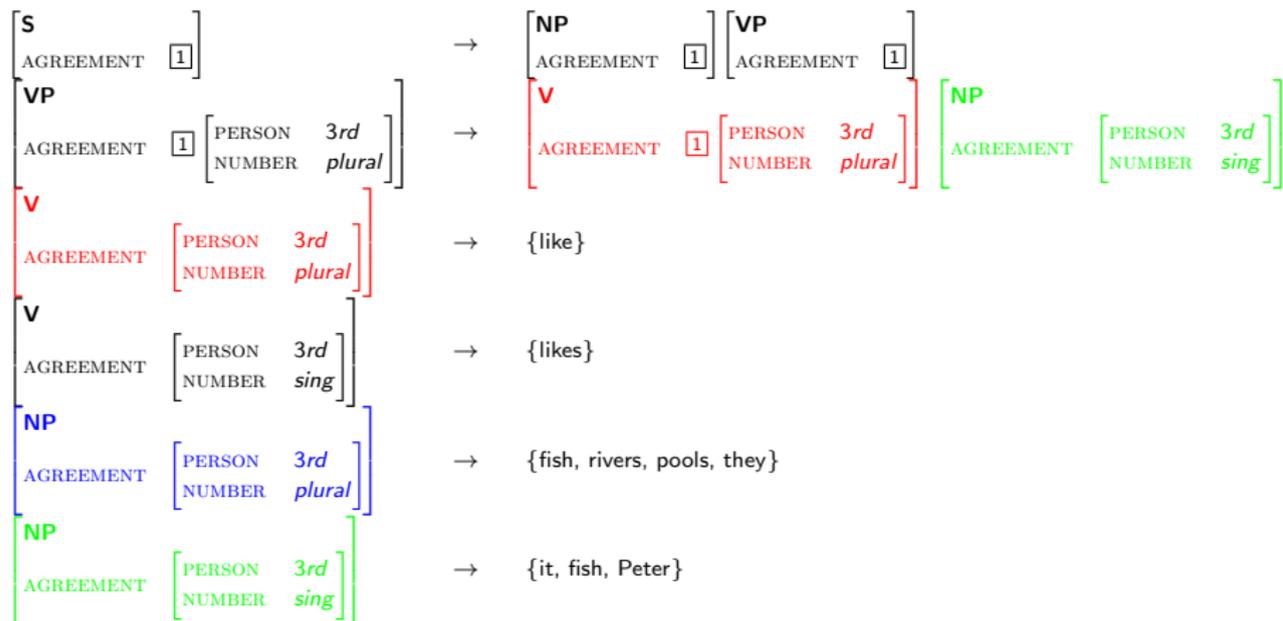
# Parsing example...

They like Peter



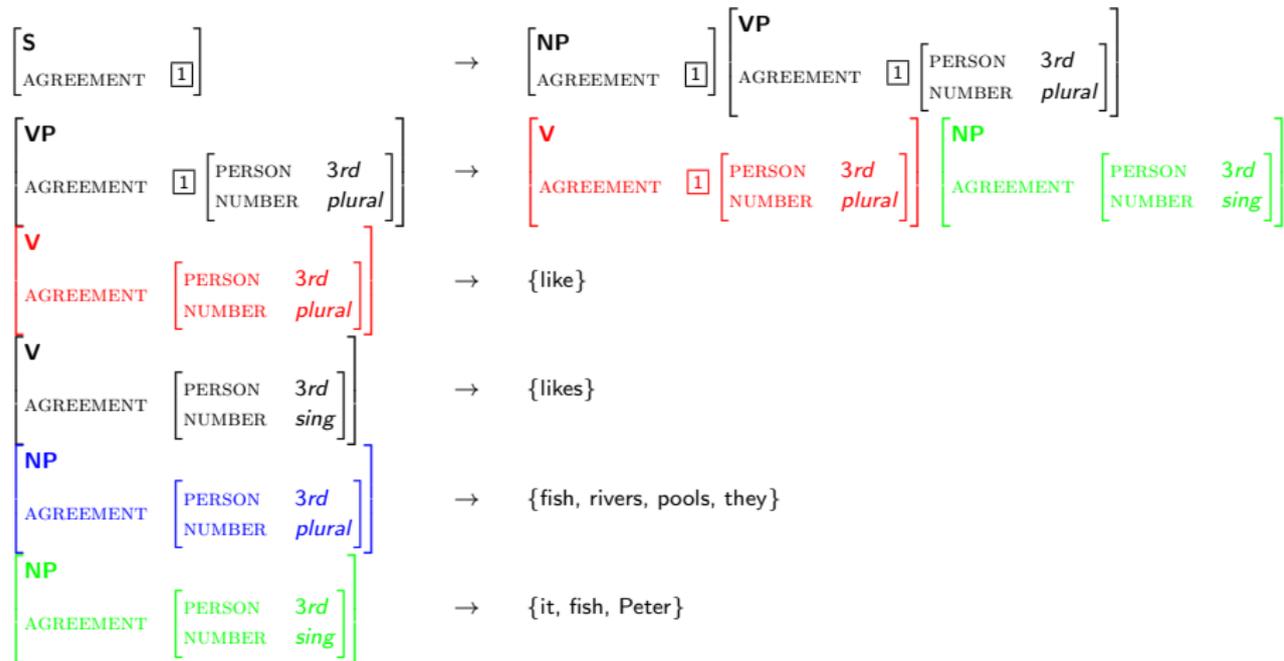
# Parsing example...

They like Peter



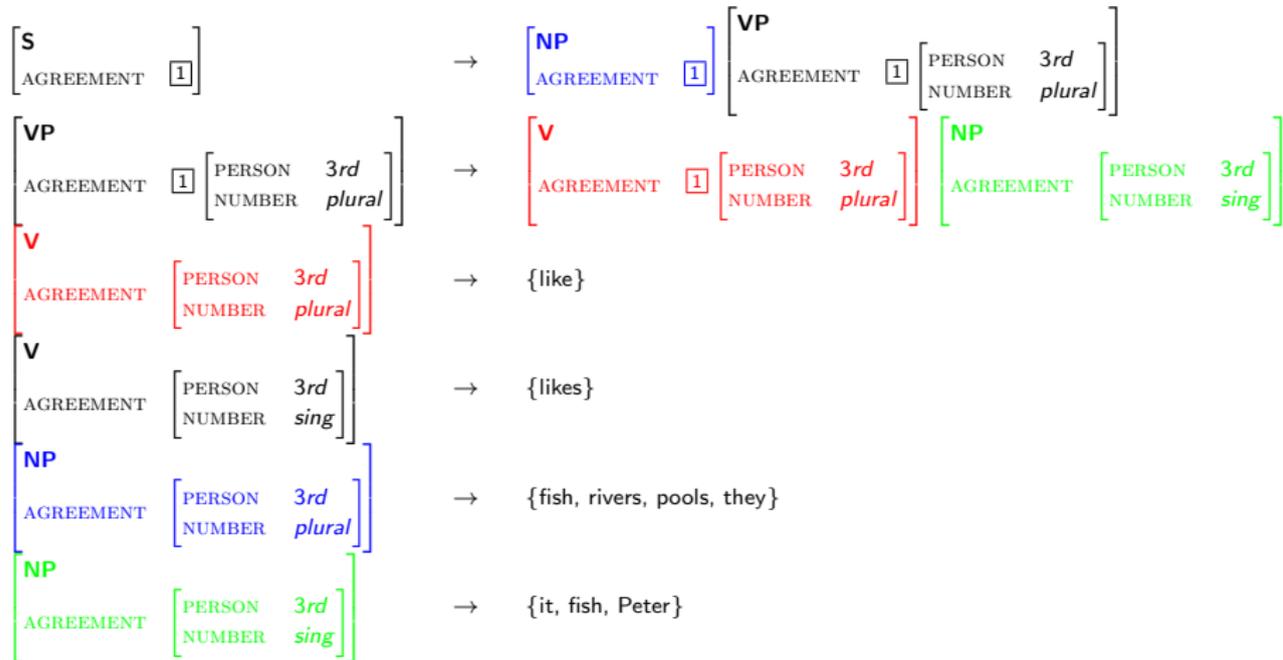
# Parsing example...

They like Peter



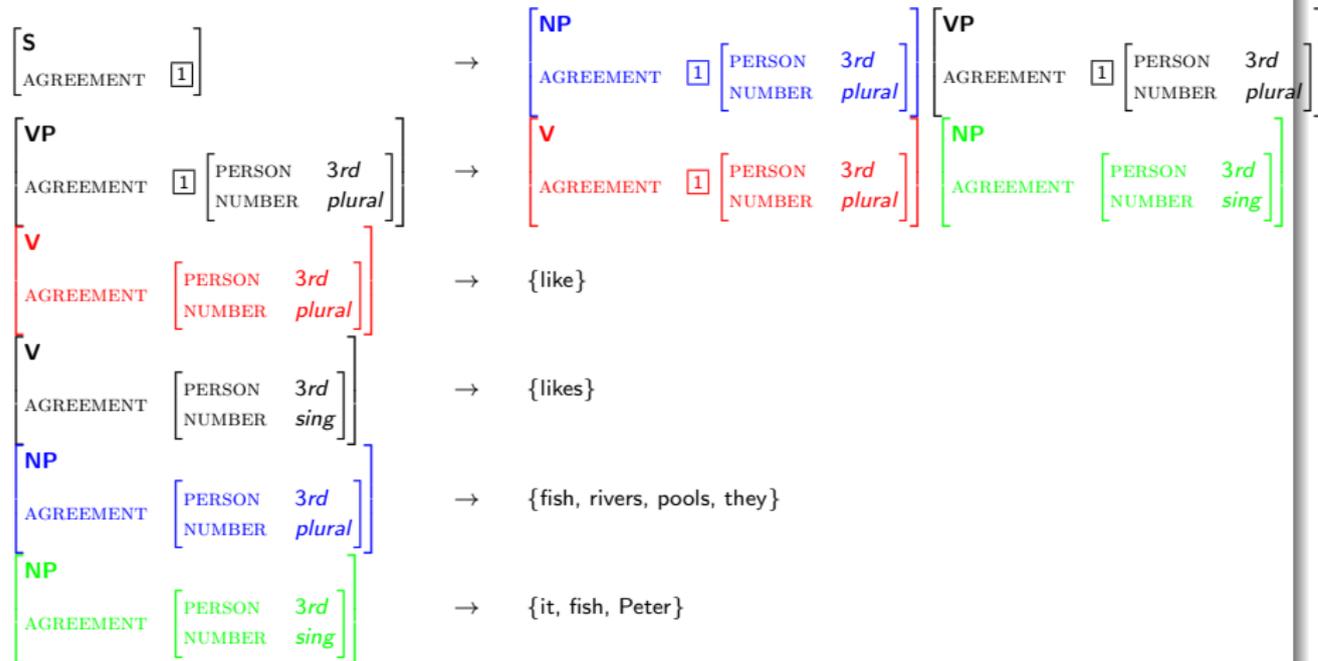
# Parsing example...

They like Peter



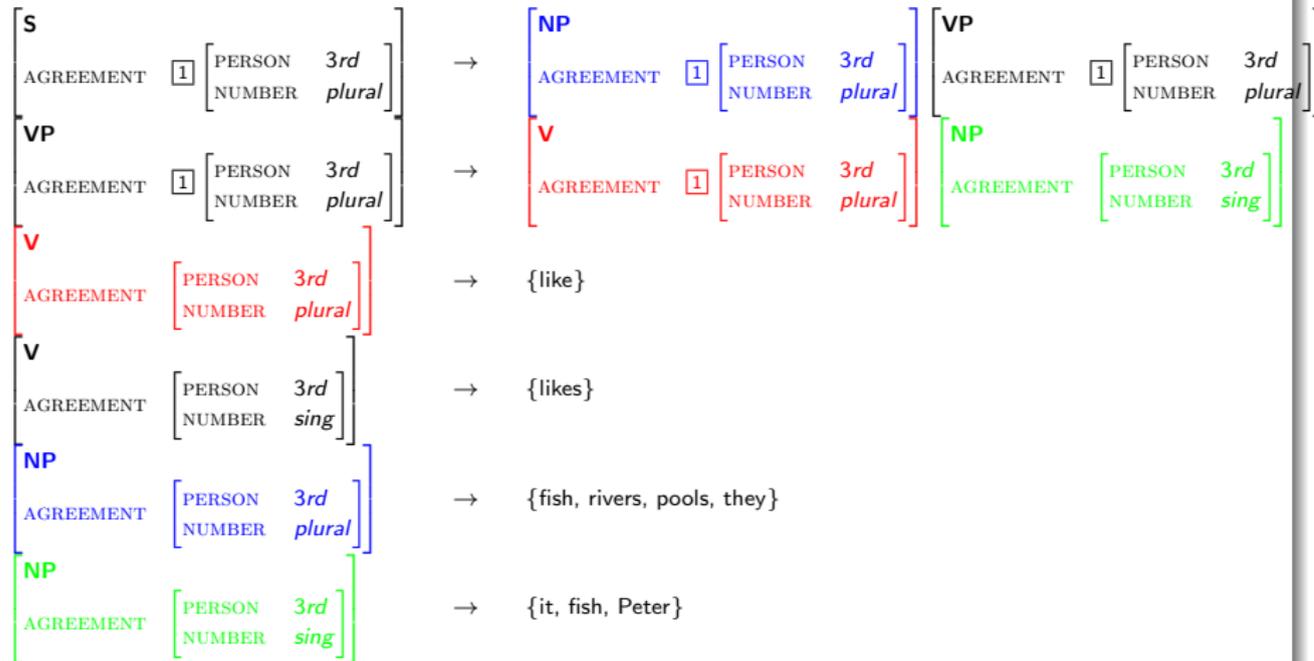
# Parsing example...

They like Peter



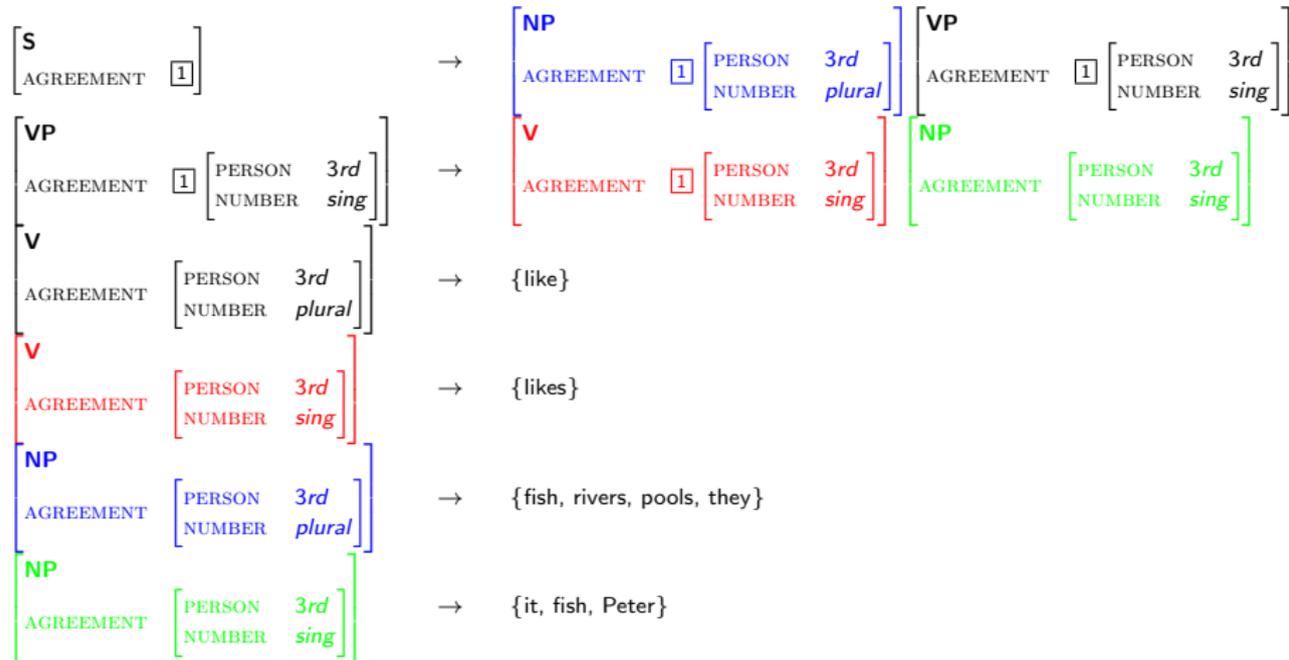
# Parsing example...

They like Peter



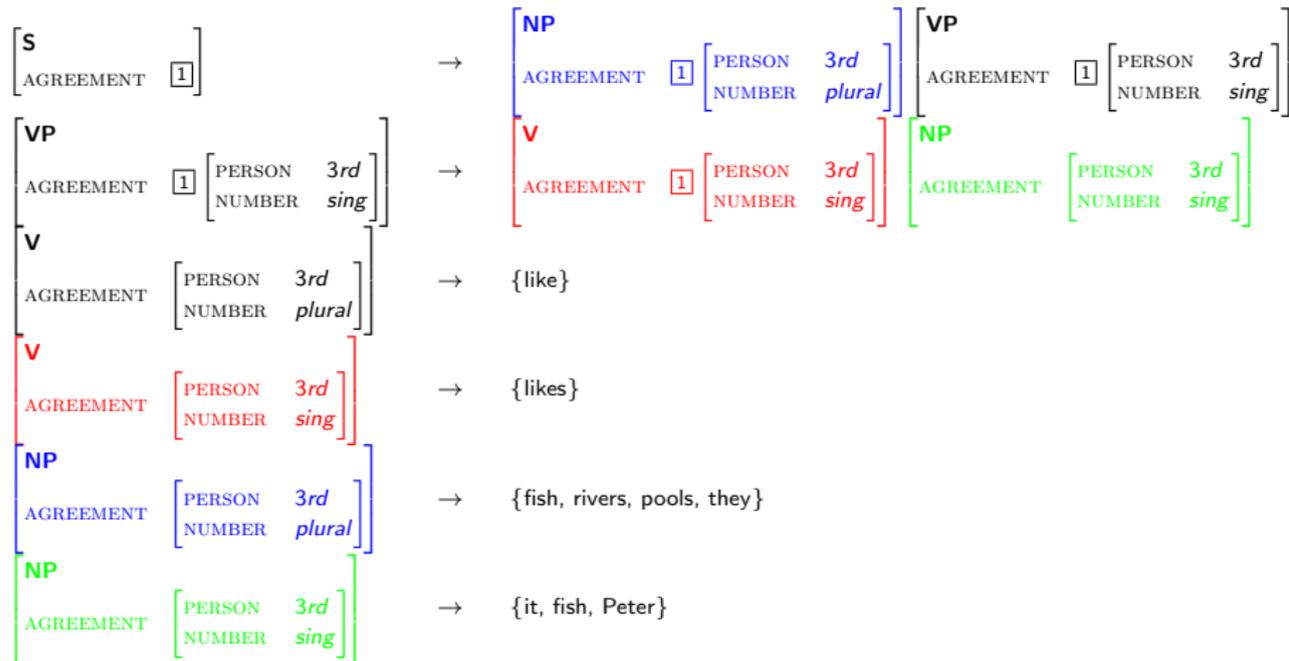
# Parsing example...

They likes Peter

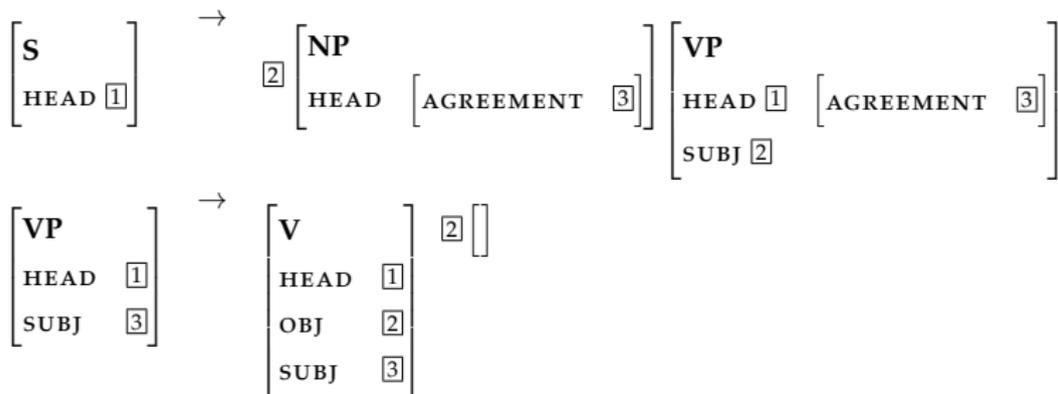


# Parsing example...

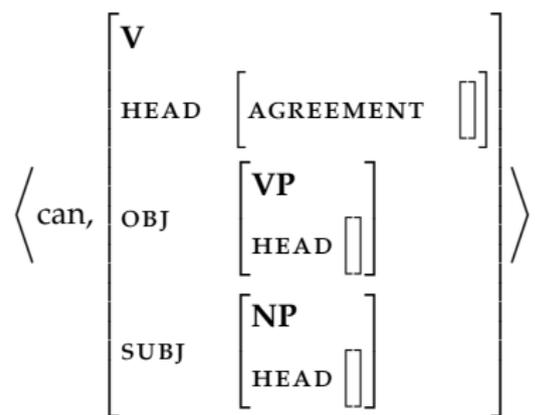
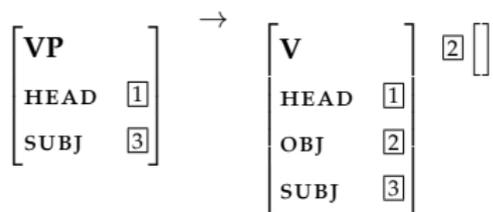
They likes Peter — UNIFICATION FAILS BECAUSE OF CO-INDEXATION



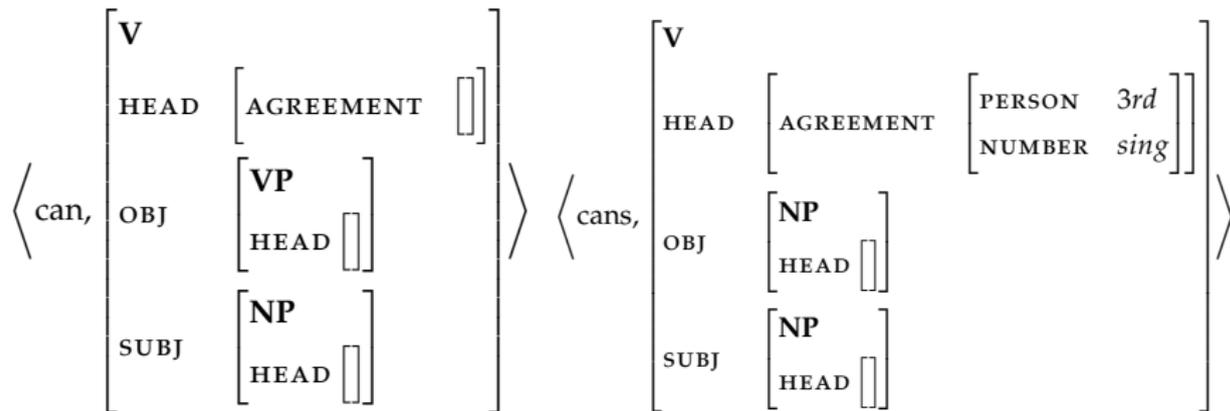
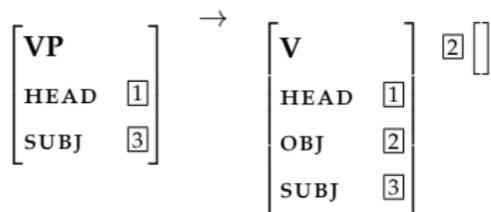
# Subcategorization is captured by the feature constraints



# Subcategorization is captured by the feature constraints



# Subcategorization is captured by the feature constraints



## Alternatively use **unification** as the parsing operation

Alternatively we can use **unification** as the parsing operation instead of just for feature checking:

- $X_0 \rightarrow X_1 X_2$   
 $\langle X_1 \text{ HEAD AGREEMENT} \rangle = \langle X_2 \text{ HEAD AGREEMENT} \rangle$   
 $\langle X_0 \text{ HEAD} \rangle = \langle X_1 \text{ HEAD} \rangle$
- $X_0 \rightarrow X_1 X_2$   
 $\langle X_0 \text{ HEAD} \rangle \langle X_1 \text{ HEAD} \rangle$   
 $\langle X_2 \text{ CAT} \rangle = PP$
- $X_0 \rightarrow X_1 \text{ and } X_2$   
 $\langle X_0 \text{ CAT} \rangle \langle X_1 \text{ CAT} \rangle$   
 $\langle X_1 \text{ CAT} \rangle \langle X_2 \text{ CAT} \rangle$

## Alternatively use **unification** as the parsing operation

Alternatively we can use **unification** as the parsing operation instead of just for feature checking:

- $X_0 \rightarrow X_1 X_2$   
 $\langle X_1 \text{ HEAD AGREEMENT} \rangle = \langle X_2 \text{ HEAD AGREEMENT} \rangle$   
 $\langle X_0 \text{ HEAD} \rangle = \langle X_1 \text{ HEAD} \rangle$
- $X_0 \rightarrow X_1 X_2$   
 $\langle X_0 \text{ HEAD} \rangle \langle X_1 \text{ HEAD} \rangle$   
 $\langle X_2 \text{ CAT} \rangle = PP$
- $X_0 \rightarrow X_1 \text{ and } X_2$   
 $\langle X_0 \text{ CAT} \rangle \langle X_1 \text{ CAT} \rangle$   
 $\langle X_1 \text{ CAT} \rangle \langle X_2 \text{ CAT} \rangle$

## Alternatively use **unification** as the parsing operation

Alternatively we can use **unification** as the parsing operation instead of just for feature checking:

- $X_0 \rightarrow X_1 X_2$   
 $\langle X_1 \text{ HEAD AGREEMENT} \rangle = \langle X_2 \text{ HEAD AGREEMENT} \rangle$   
 $\langle X_0 \text{ HEAD} \rangle = \langle X_1 \text{ HEAD} \rangle$
- $X_0 \rightarrow X_1 X_2$   
 $\langle X_0 \text{ HEAD} \rangle \langle X_1 \text{ HEAD} \rangle$   
 $\langle X_2 \text{ CAT} \rangle = PP$
- $X_0 \rightarrow X_1 \text{ and } X_2$   
 $\langle X_0 \text{ CAT} \rangle \langle X_1 \text{ CAT} \rangle$   
 $\langle X_1 \text{ CAT} \rangle \langle X_2 \text{ CAT} \rangle$

# Lexical AVMs may be derived through unification

- We have assumed we have a lexicon entry for all the inflected forms of a word.
- With a morphological analysis step we can return a word its stem and affixes and then build the AVM from the pieces: *foxes*  $\rightarrow$  *fox*<sup>s</sup>

$$\left\langle \text{^s}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } pl \right] \right\rangle$$

$$\left\langle \text{fox}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } \boxed{\phantom{pl}} \right] \right\rangle$$

$$\left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right] \sqcup \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } \boxed{\phantom{pl}} \right] = \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right]$$

# Lexical AVMs may be derived through unification

- We have assumed we have a lexicon entry for all the inflected forms of a word.
- With a morphological analysis step we can return a word its stem and affixes and then build the AVM from the pieces: *foxes*  $\rightarrow$  *fox*<sup>s</sup>

$$\left\langle \text{^s}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } pl \right] \right\rangle$$

$$\left\langle \text{fox}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } \left[ \right] \right] \right\rangle$$

$$\left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right] \sqcup \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } \left[ \right] \right] = \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right]$$

# Lexical AVMs may be derived through unification

- We have assumed we have a lexicon entry for all the inflected forms of a word.
- With a morphological analysis step we can return a word its stem and affixes and then build the AVM from the pieces: *foxes*  $\rightarrow$  *fox*<sup>s</sup>

$$\left\langle \text{^s}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } pl \right] \right\rangle$$

$$\left\langle \text{fox}, \left[ \text{HEAD} \left[ \text{N} \right] \right. \right. \\ \left. \left. \text{AGREEMENT } \left[ \right] \right] \right\rangle$$

$$\left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right] \sqcup \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } \left[ \right] \right] = \left[ \text{HEAD} \left[ \text{N} \right] \right. \left. \text{AGREEMENT } pl \right]$$

## Unification based parsing in the wild...

- Focus on adequacy for a wide range of languages as well as tractable for parsing
- Examples include **Lexical Functional Grammar, LFG** (Bresnan and Kaplan) and **Head-driven Phrase Structure Grammar, HPSG** (Pollard and Sag)
- Grammars tend to incorporate aspects of morphology, syntax and compositional semantics:

If you are interested see: <http://www.delph-in.net>