

---

# **The SUBTLE NL Parsing Pipeline: A Complete Parser for English**

**Mitch Marcus**  
**University of Pennsylvania**



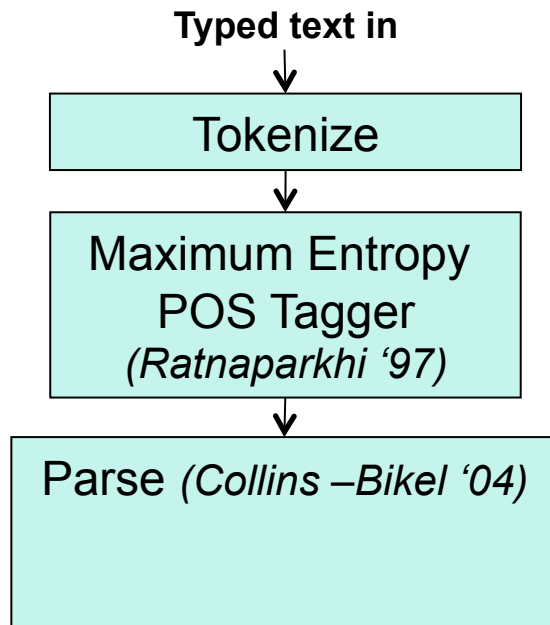
# PICTURE OF ANALYSIS PIPELINE

---

- **Tokenize**
- **Maximum Entropy POS tagger MXPOST  
Ratnaparkhi**
- **Core Parser Collins Generative Prob. CFG**
- **Functional Tag Replacement – Kulick**
- **Null Element Placement – Ryan Gabbard PhD Dec  
2010 (just submitted)**



# NLP Analysis Pipeline



Junior, defuse any bombs immediately when going into a room.

Junior\_NNP ,\_, defuse\_VBP any\_DT bombs\_NNS immediately\_RB  
when\_WRB going\_VBG into\_IN a\_DT room\_NN .\_.

( (S (NP (NNP Junior))  
(, ,)

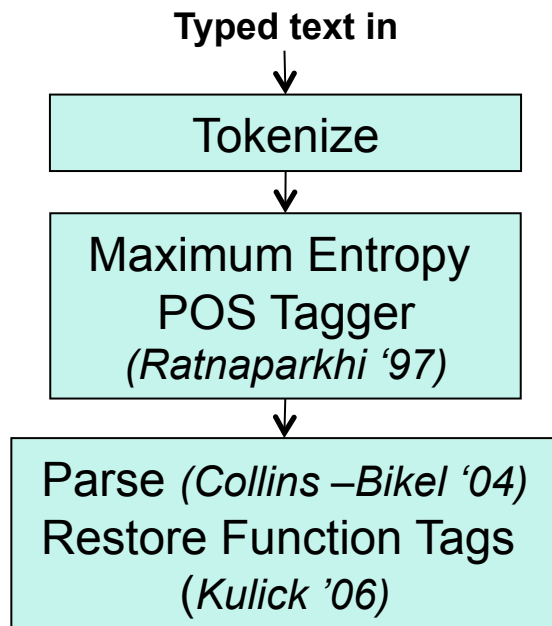
(VP (VB defuse)  
(NP (DT any)  
(NNS bombs))  
(ADVP (RB immediately))  
(SBAR (WHADVP (WRB when))

(S  
(VP (VBG going)  
(PP (IN into)  
(NP (DT a)  
(NN room))))  
)))))

(. .))



# NLP Analysis Pipeline



Junior, defuse any bombs immediately when going into a room.

Junior\_NNP ,\_, defuse\_VBP any\_DT bombs\_NNS immediately\_RB  
when\_WRB going\_VBG into\_IN a\_DT room\_NN .\_.

( (S (NP (NNP Junior))  
(, ,)

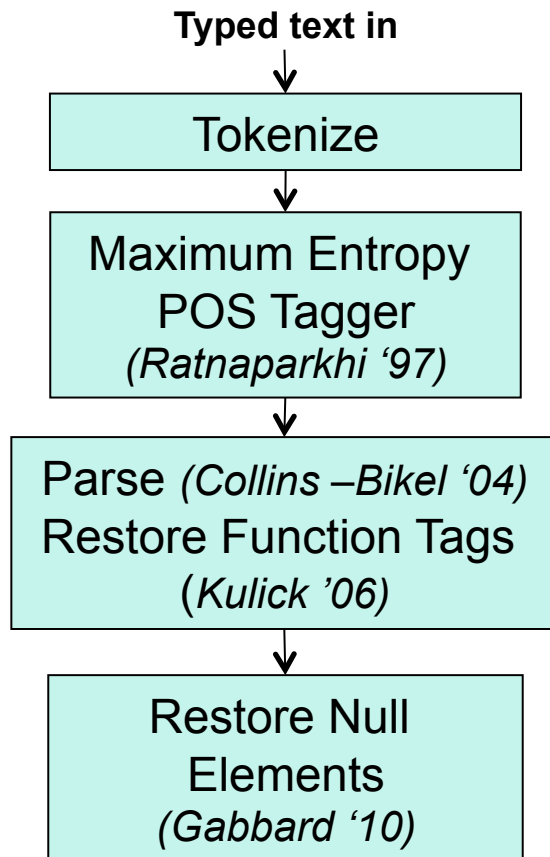
(VP (VB defuse)  
(NP (DT any)  
(NNS bombs))  
(ADVP (RB immediately))  
(SBAR (WHADVP (WRB when))

(S  
(VP (VBG going)  
(PP (IN into)  
(NP (DT a)  
(NN room)))  
)))))

(. .))



# NLP Analysis Pipeline



Junior, defuse any bombs immediately when going into a room.

Junior\_NNP ,\_, defuse\_VBP any\_DT bombs\_NNS immediately\_RB  
when\_WRB going\_VBG into\_IN a\_DT room\_NN .\_.

( (S (NP-**VOC** (NNP Junior))  
(, ,)

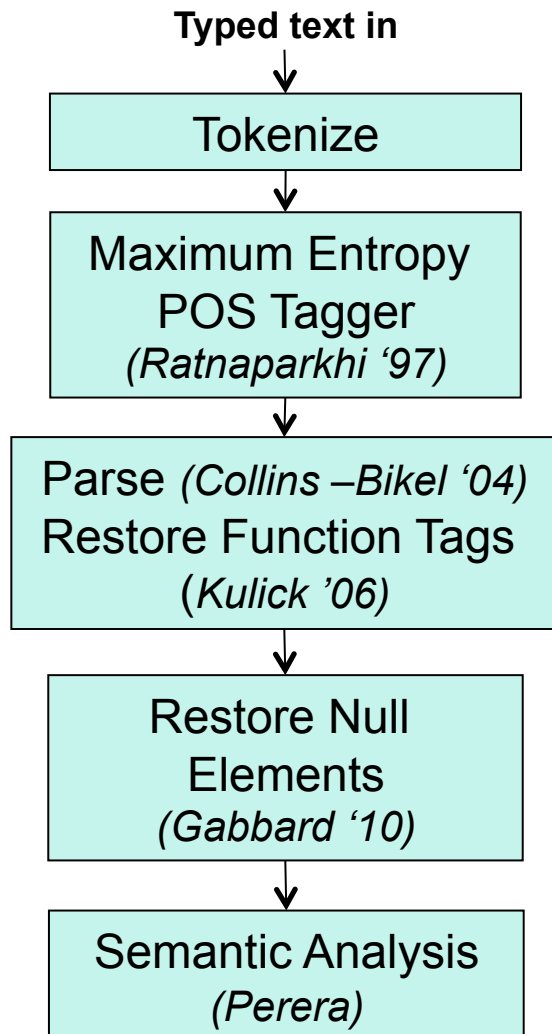
(VP (VB defuse)  
(NP (DT any)  
(NNS bombs))  
(ADVP-**TMP** (RB immediately))  
(SBAR-**TMP** (WHADVP (WRB when))

(S  
(VP (VBG going)  
(PP-**DIR** (IN into)  
(NP (DT a)  
(NN room)))  
)))))

(. .))



# NLP Analysis Pipeline



Junior, defuse any bombs immediately when going into a room.

Junior\_NNP ,\_, defuse\_VBP any\_DT bombs\_NNS immediately\_RB  
when\_WRB going\_VBG into\_IN a\_DT room\_NN .\_.

( (S (NP-**VOC** (NNP Junior))  
(, ,)  
(NP-**SBJ-0** (-NONE- \*))  
(VP (VB defuse)  
(NP (DT any)  
(NNS bombs))  
(ADVP-**TMP** (RB immediately))  
(SBAR-**TMP** (WHADVP-1 (WRB when))  
(S (NP-**SBJ-0** (-NONE- \*))  
(VP (VBG going)  
(PP-**DIR** (IN into)  
(NP (DT a)  
(NN room))))  
(ADVP-1 (-NONE- \*T\*))))))  
(. .)))

11/4/2010

SUBTLE Annual Review

6





# Generative Parsing Results (<40 w)

---

Parser	F <sub>1</sub>
Magerman 95	84.7
Collins 96	86.0
Charniak 97	87.4
Collins 99 (Bikel 04)	88.6
Liang 08 (extending Charniak 05)	91.7

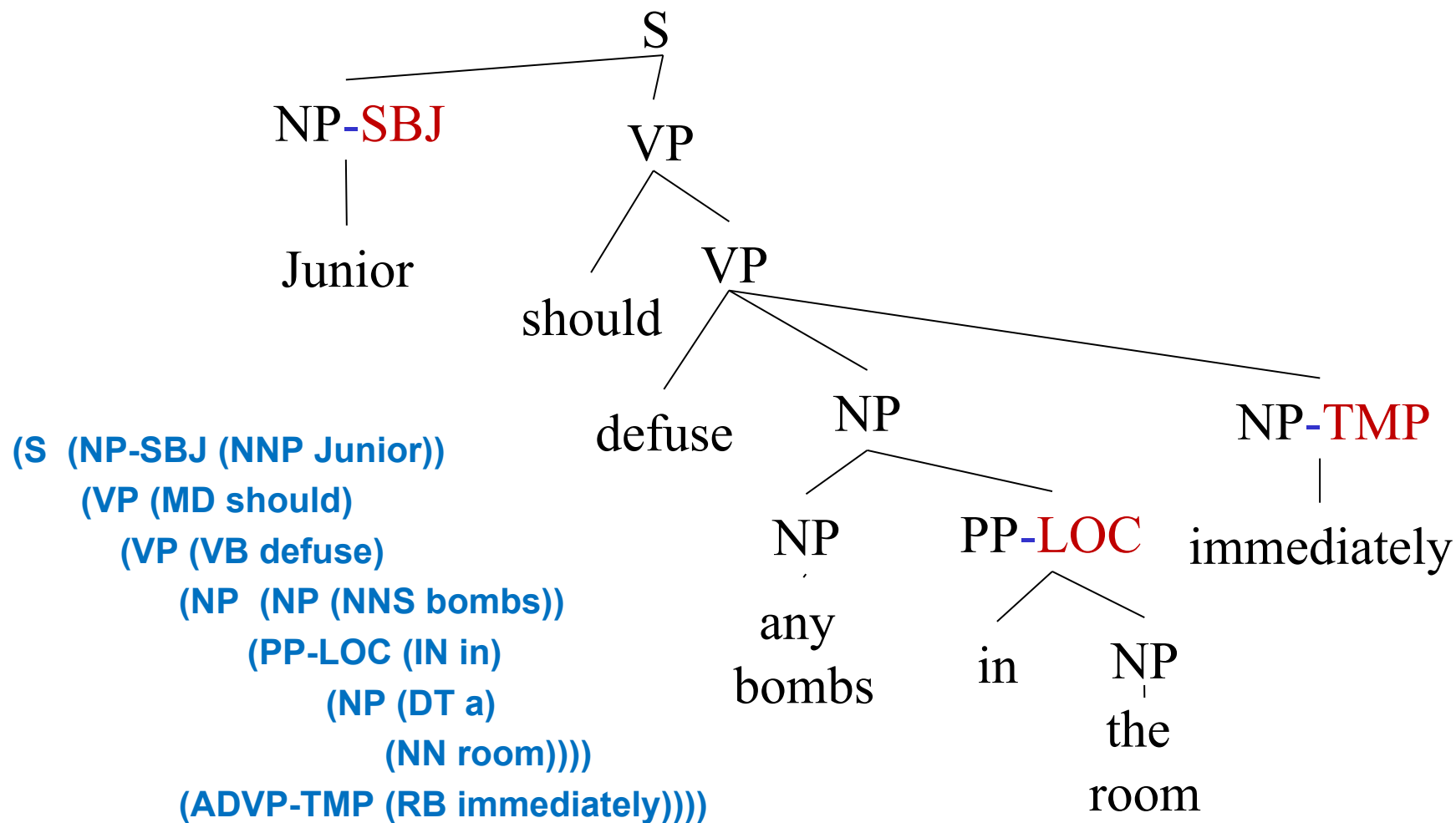


---

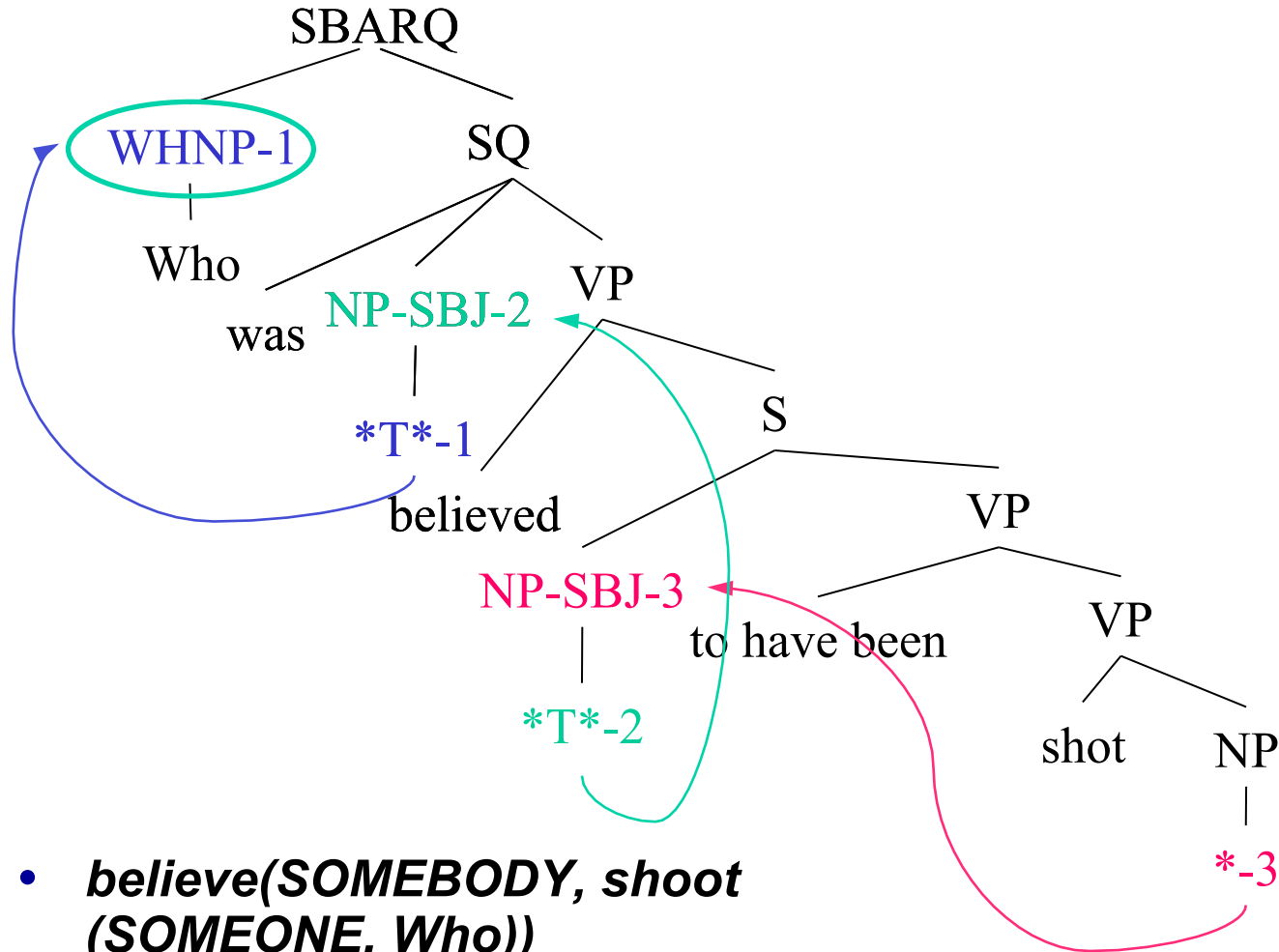
# Meaningless Progress.....



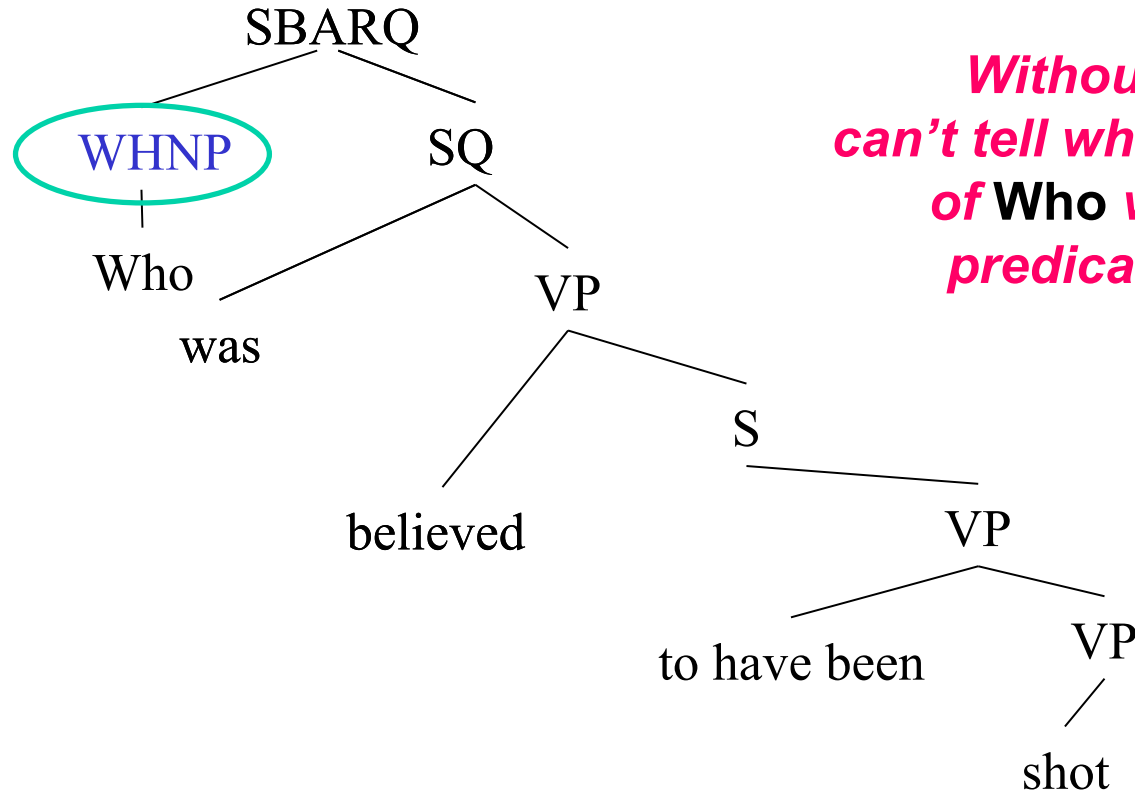
# Treebank Tree with Function Tags



# Treebank Tree with Null Elements



# A “Perfect” Parseval Analysis..



*Without null element,  
can't tell what the semantic role  
of **Who** was, or even the  
predication it's part of.*

- ***believe(SOMEBODY, shoot (SOMEONE, Who))***



# A Complete Parsing System

---

- **Ryan Gabbard, Seth Kulick, and Mitch Marcus. Fully Parsing the Penn Treebank. HLT/NAACL 2006**
- **Ryan Gabbard, Null Element Restoration, PhD Dissertation, Nov 2010**



---

# Step 1: Recovering Function Tags



# Treebank II Function Tag Set

<i>Syntactic</i>		<i>Semantic</i>		<i>Miscellaneous</i>	
<b>SBJ</b>	<b>Subject</b>	<b>TMP</b>	<b>Temporal</b>	<b>CLF</b>	<b>It-cleft</b>
<b>LGS</b>	<b>Logical Subject</b>	<b>LOC</b>	<b>Location</b>	<b>HLN</b>	<b>Headline</b>
<b>PRD</b>	<b>Predicate</b>	<b>DIR</b>	<b>Direction</b>	<b>TTL</b>	<b>Title</b>
<b>DTV</b>	<b>Dative</b>	<b>EXT</b>	<b>Extent</b>	<b><i>Topicalization</i></b>	
<b>PUT</b>	<b>Locative of “put”</b>	<b>BNF</b>	<b>Benefactive</b>	<b>TPC</b>	<b>Topic</b>
<b>VOC</b>	<b>Vocative</b>	<b>MNR</b>	<b>Manner</b>		
		<b>PRP</b>	<b>Purpose</b>		
		<b>NOM</b>	<b>Nominal</b>	<b><i>CLR</i></b>	
		<b>ADV</b>	<b>Non-specific adverbial</b>	<b>CLR</b>	<b>Closely-Related</b>



# A Sample PCFG

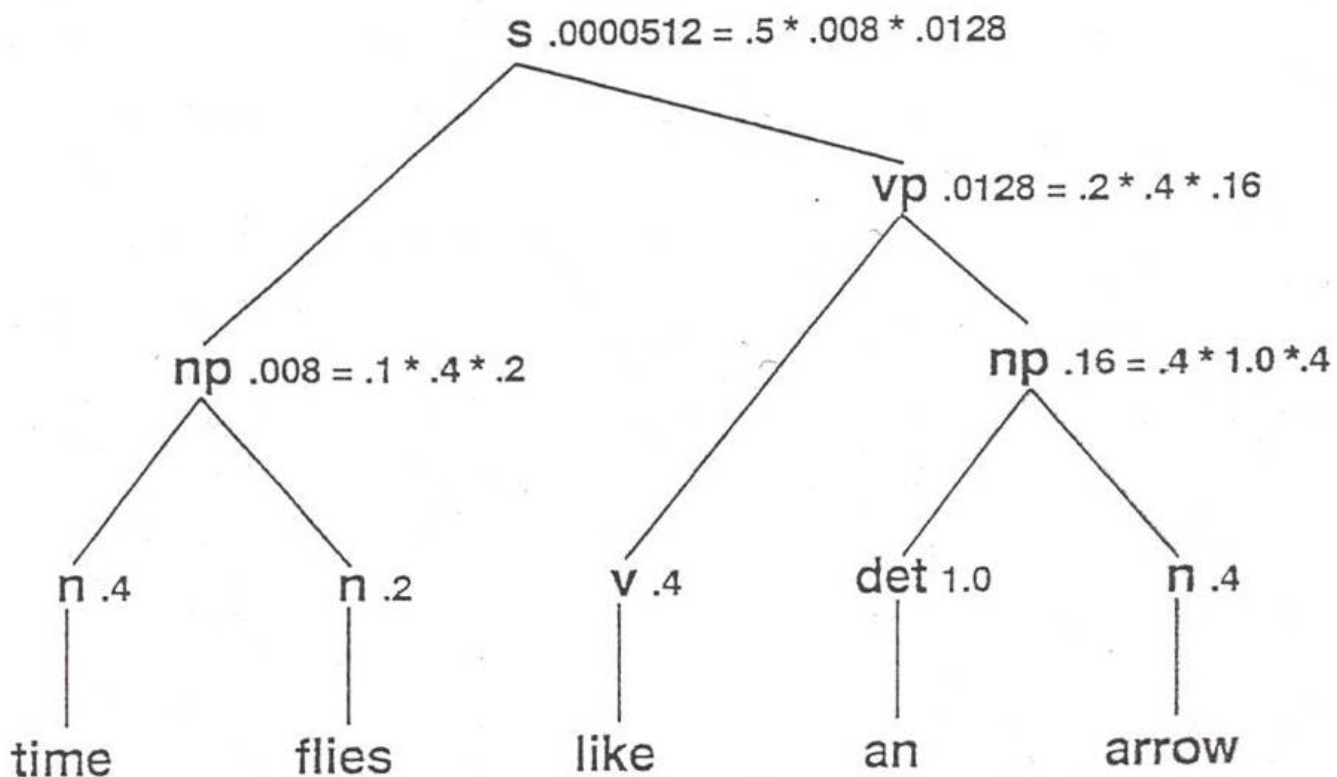
---

1.  $s \xrightarrow{.5} np\ vp$
2.  $s \xrightarrow{.25} s\ pp$
3.  $s \xrightarrow{.15} v\ np$
4.  $s \xrightarrow{.10} v\ pp$
5.  $vp \xrightarrow{.2} v\ np$
6.  $vp \xrightarrow{.4} v\ pp$
7.  $vp \xrightarrow{.4} v$
8.  $np \xrightarrow{.1} n\ n$
9.  $np \xrightarrow{.4} det\ n$
10.  $np \xrightarrow{.3} np\ pp$
11.  $np \xrightarrow{.2} n$
12.  $n \xrightarrow{.4} arrow$
13.  $n \xrightarrow{.2} flies$
14.  $n \xrightarrow{.4} time$
15.  $v \xrightarrow{.4} flies$
16.  $v \xrightarrow{.4} like$
17.  $v \xrightarrow{.2} time$
18.  $p \xrightarrow{1.0} like$
19.  $det \xrightarrow{1.0} a$
20.  $pp \xrightarrow{1.0} p\ np$



# An Example Derivation

---

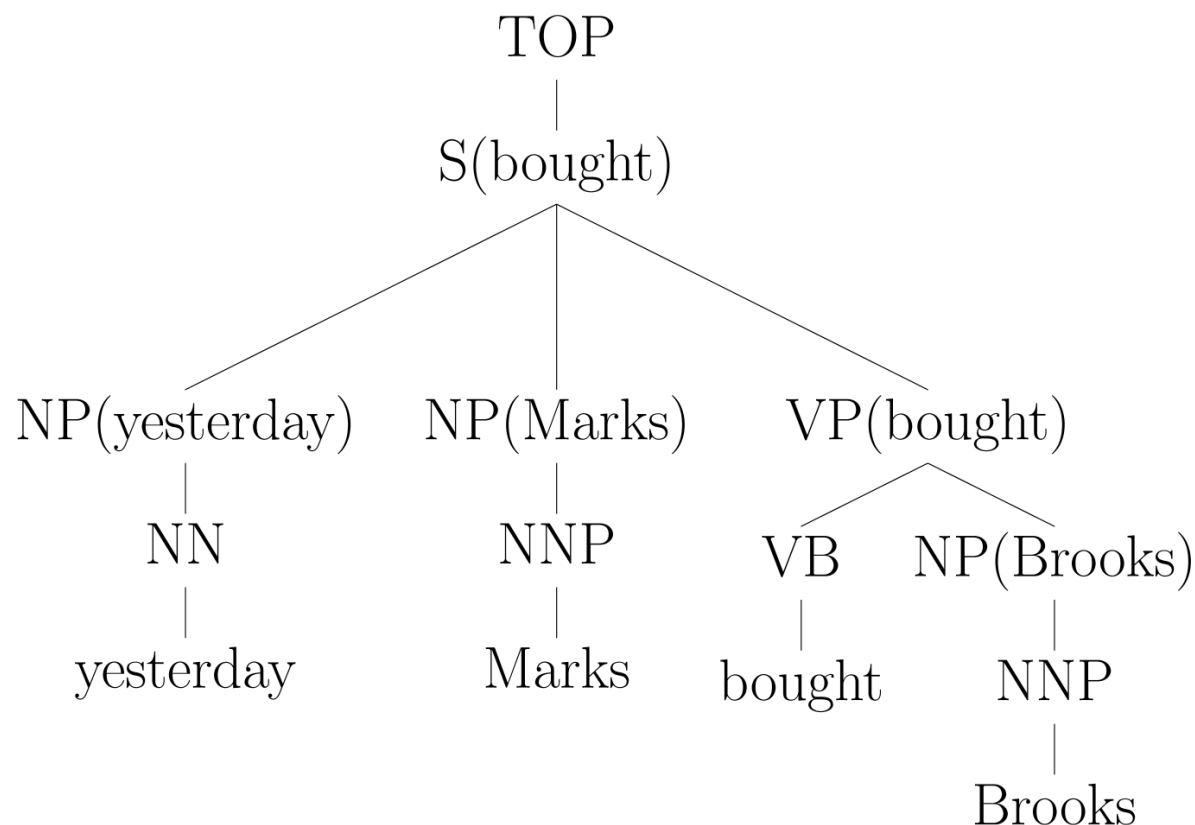


*This example adapted from Jelinek et. al. 1991.*



# The Key Trick in '95: Lexicalizing the CFG

---



New problem: Estimating the probability of  
 $S(\text{bought}) \rightarrow NP(\text{yesterday}) NP(\text{Marks}) VP(\text{bought})$



# Solving the Sparse Data Problem with Independence Assumptions

---

Instead of

$S(\text{bought}) \rightarrow NP(\text{yesterday}) NP(\text{Marks}) VP(\text{bought})$

1. Generate Head, given lexicalized parent:

$S(\text{bought}) \rightarrow \dots VP(\text{bought}) \dots$

2. Generate other children given parent & head:

$S(\text{bought}) \rightarrow NP( ) NP( ) VP(\text{bought})$

3. Lexicalize other children, given parent, head & node type:

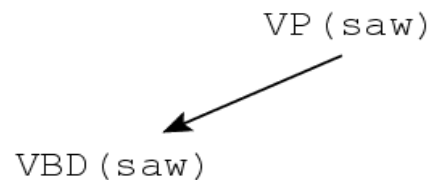
$S(\text{bought}) \rightarrow NP(\text{yesterday}) NP(\text{Marks}) VP(\text{bought})$



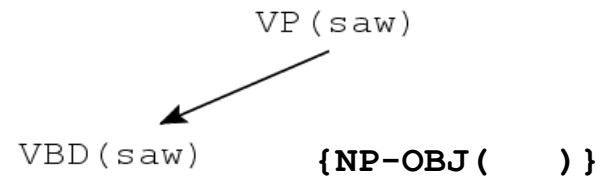
# To augment the parser output with function tags

---

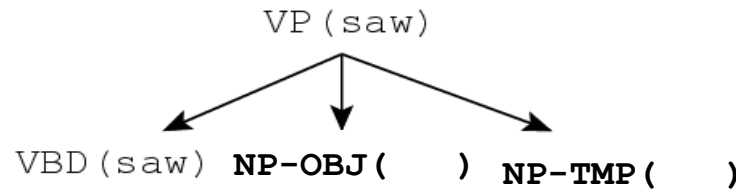
1. Find the code in Collins Model II that removes function tags
2. Remove it!



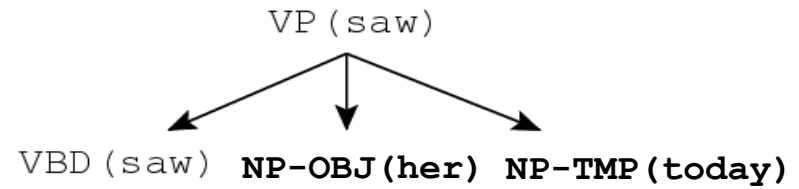
(1)



(2)



(3)



(4)

*(From Collins & Koo '05)*



## Results for Function Tag Groups & Overall

---

	Overall w/ CLR	Overall w/o CLR
<i>Collins 2 +FTags</i>	<b>88.95</b>	<b>90.78</b>
<b>Blaheta, 2003</b>		<b>88.28</b>
<b>Jijkoun &amp; de Rijke, 2004</b>	<b>88.50</b>	

- **With no drop in basic parser F score!**



## Results for Function Tag Groups & Overall

---

	Overall w/ CLR	Overall w/o CLR	Syn 56%	Sem 36%
<b><i>Collins 2 +FTags</i></b>	<b>88.95</b>	<b>90.78</b>	<b>95.76</b>	<b>84.56</b>
<b>Blaheta, 2003</b>		<b>88.28</b>		
<b>Jijkoun &amp; de Rijke, 2004</b>	<b>88.50</b>			
<b>Musillo &amp; Merlo, Oct 2005</b>			<b>96.5</b>	<b>85.6</b>

- **With no drop in basic parser F score!**

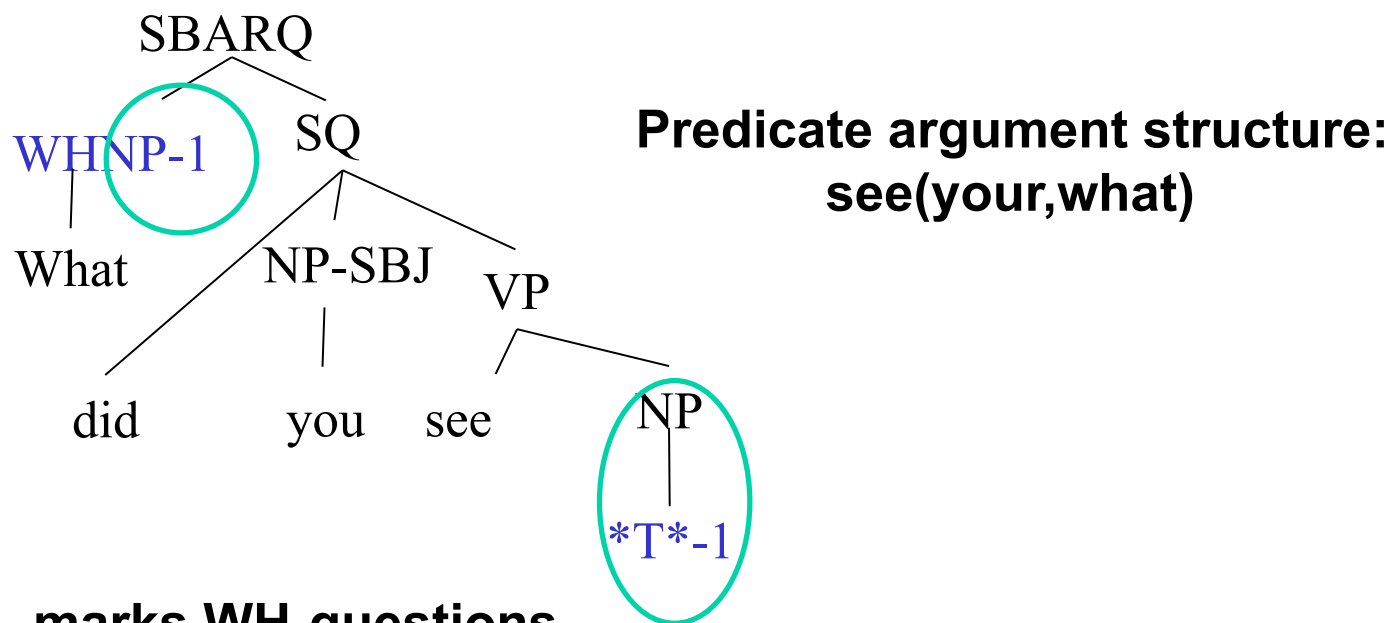


---

## Step 2: Recovering Null Elements



# Null Elements for “Wh-movement”



- **\*T\*** - marks WH-questions
- Null elements are co-indexed using integer tags.
- To recover Pred-arg structure:
  - Replace null element with co-indexed lexical material
- ***Similar approaches for other grammatical phenomenon***



# Previous Work Forward

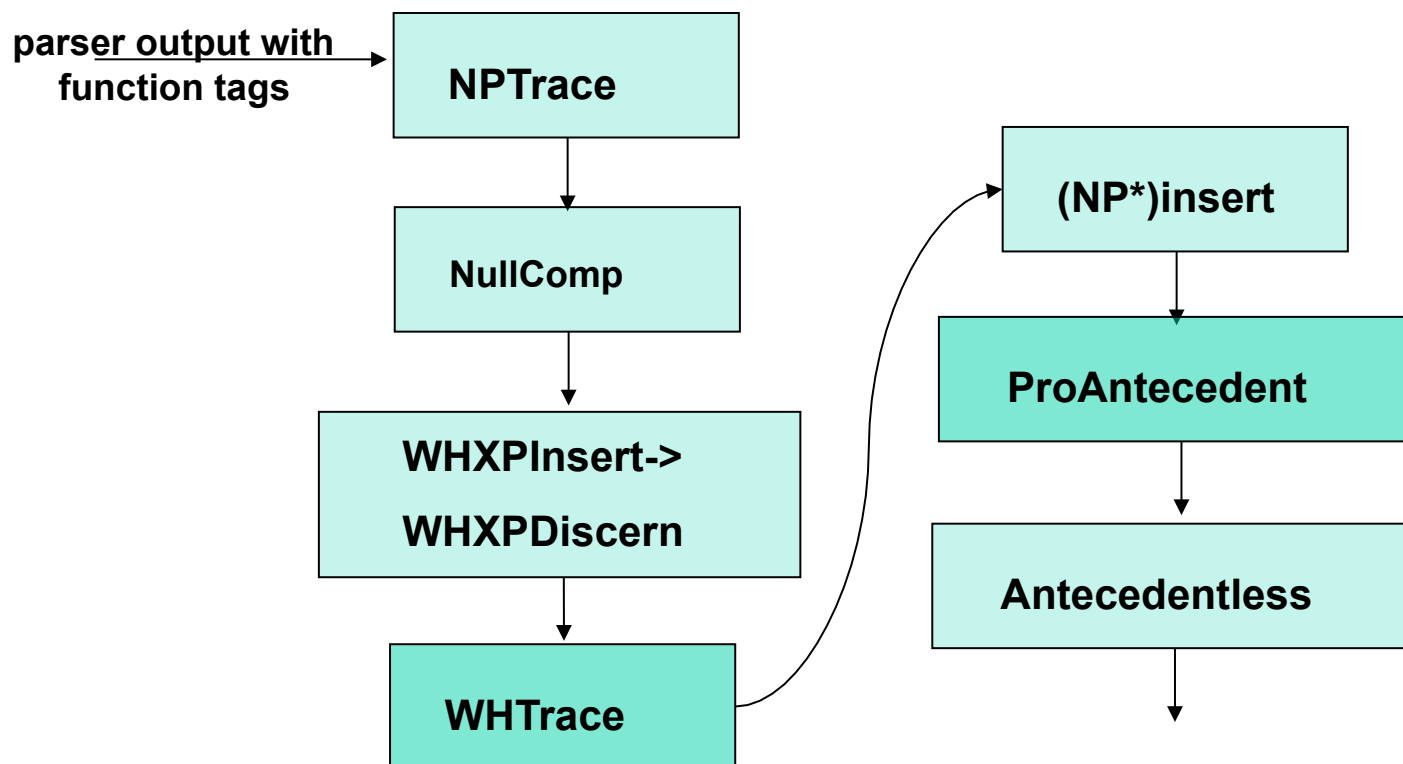
---

- **Integrated with Parsing** (Collins Model 3, Dienes&Dubey 03)
- **Postprocessing Parser Output**
  - **Hand Built Rule-based** (Campbell 04)
  - **Statistical Machine Learning Approaches**  
(Johnson 02, Levy&Manning 04, Jijkoun&de Rijke 04)
- **Our approach combines:**
  - **Campbell's Problem Decomposition**
  - **Machine Learning**
- **Into:**
  - **A pipeline of linear classifiers**



# The Null Element Replacement Algorithm

- Pipeline & rich linguistic features á la Campbell
- Pipeline of linear classifiers (McCallum's Mallet)



## Aggregate results over all recovered categories

---

System	Precision	Recall	F1
D&D	<b>78.50</b>	68.08	72.92
Pres	74.70	<b>74.62</b>	<b>74.66</b>

- **Automatically parsed trees from section 23**



## Present vs. Campbell's Rule Based System

---

- Test on gold--standard trees from section 23 using Campbell's metric

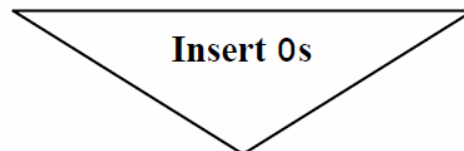
Category	Present	Campbell
NP *	<b>88.8</b>	86.9
NP *T*	<b>96.3</b>	96.0
ADV P *T*	<b>82.2</b>	79.9
0	<b>99.8</b>	98.5



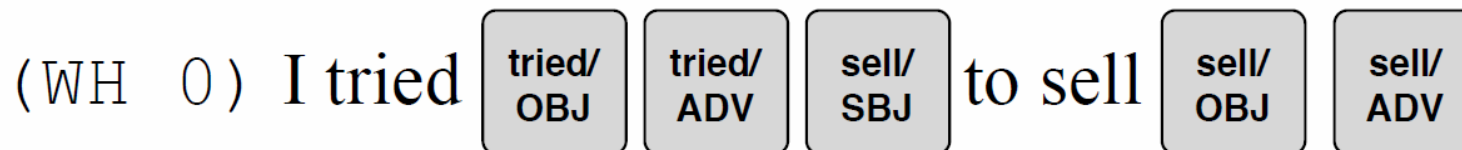
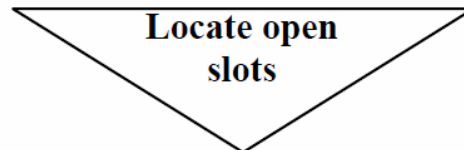
# Wh – placement with factor graphs

---

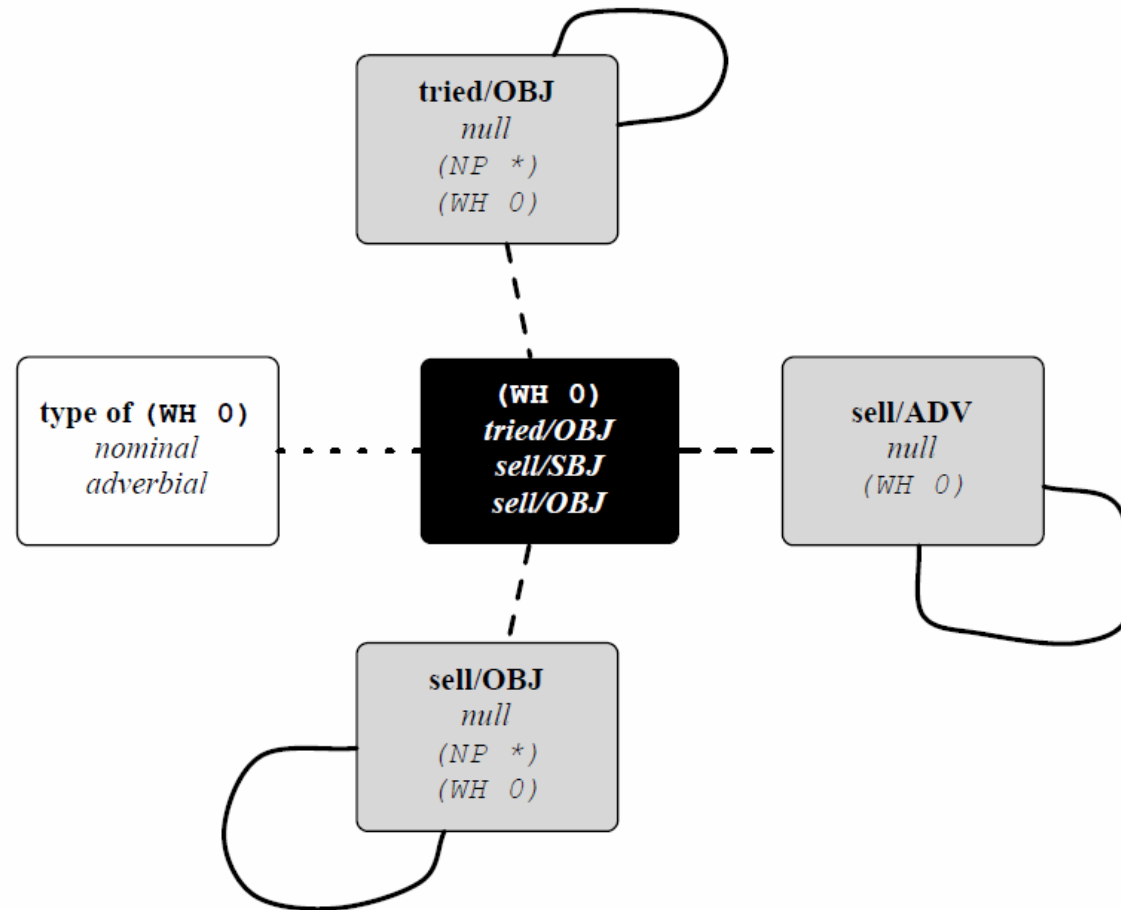
That is the book I tried to sell



That is the book (WH 0) I tried to sell



# Wh – placement with factor graphs



# Results: two systems for Wh-placement

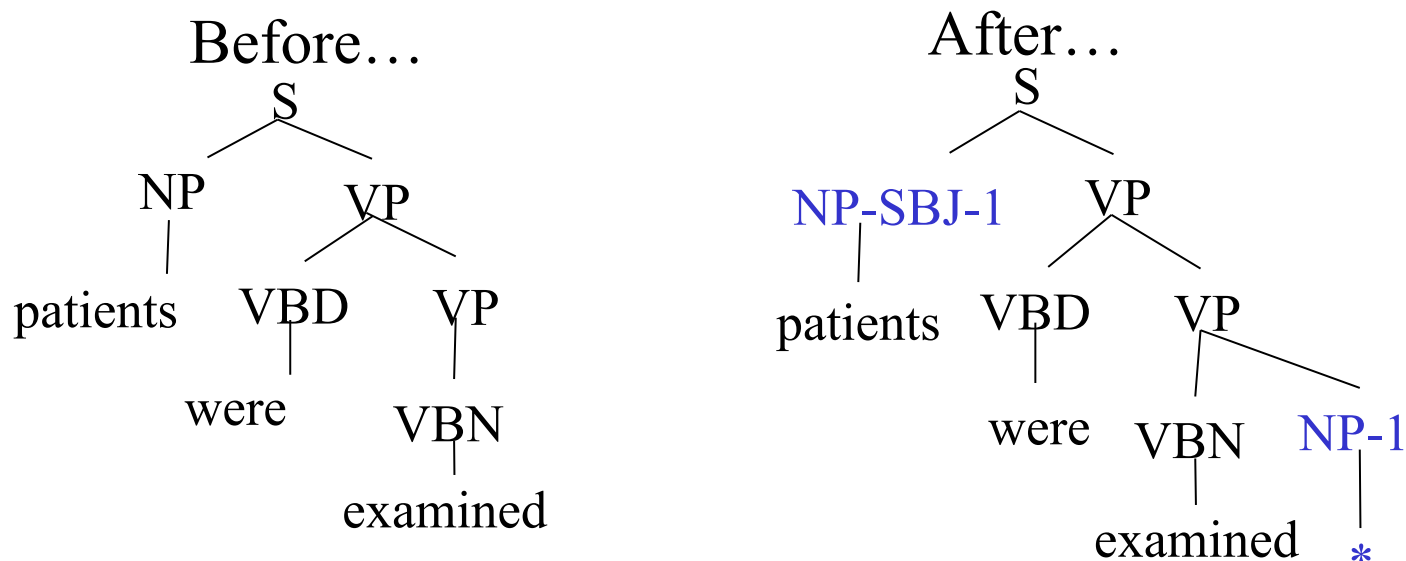
---

Type	Gold		Parsed	
	Old	New	Old	New
NP *T*	92.8	<b>96.5</b>	85.9	<b>87.9</b>
ADVP *T*	79.2	<b>86.7</b>	77.7	<b>79.8</b>
NP *	78.6	<b>82.7</b>	<b>72.2</b>	71.6
<i>NP * (na)</i>	95.8	<b>96.6</b>	<b>88.3</b>	88.0
WHNP 0	<b>92.0</b>	90.2	<b>61.5</b>	59.6
WHADVP 0	<b>71.0</b>	66.7	<b>68.9</b>	61.8



# Summary: Full Parsing!

---



- **20 Function Tags – Mostly syntactic or semantic**
  - **Syntactic Tags: 95% recovery accuracy**
  - **Semantic Tags: 84% recovery accuracy**
- **Traces**
  - **Passivization, control, raising: ~70% recovered**
  - **Wh-traces: ~80% recovered**

