

Syntactic Features

Morphology, heads, gaps,
etc.

600.465 - Intro to NLP - J. Eisner

1

3 views of a context-free rule

- generation (production): $S \rightarrow NP VP$
- parsing (comprehension): $S \leftarrow NP VP$
- verification (checking): $S = NP VP$
- Today you should keep the third, declarative perspective in mind.
- Each phrase has
 - an interface (S) saying where it can go
 - an implementation ($NP VP$) saying what's in it
- To let the parts of the tree coordinate more closely with one another, enrich the interfaces:
 $S[\text{features}...] = NP[\text{features}...] VP[\text{features}...]$

600.465 - Intro to NLP - J. Eisner

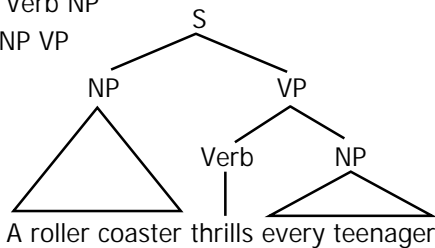
2

Examples

Verb \rightarrow thrills

VP \rightarrow Verb NP

S \rightarrow NP VP



A roller coaster thrills every teenager

600.465 - Intro to NLP - J. Eisner

3

3 common ways to use features

morphology of a single word:

Verb[head=thrill, tense=present, num=sing, person=3,...] \rightarrow thrills

projection of features up to a bigger phrase

VP[head= α , tense= β , num= γ ...] \rightarrow V[head= α , tense= β , num= γ ...] NP
provided α is in the set TRANSITIVE-VERBS

agreement between sister phrases:

S[head= α , tense= β] \rightarrow NP[num= γ ...] VP[head= α , tense= β , num= γ ...]

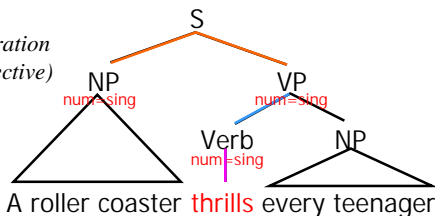
600.465 - Intro to NLP - J. Eisner

4

3 Common Ways to Use Features

Verb[head=thrill, tense=present, num=sing, person=3,...] \rightarrow thrills
 VP[head= α , tense= β , num= γ ...] \rightarrow V[head= α , tense= β , num= γ ...] NP
 S[head= α , tense= β] \rightarrow NP[num= γ ...] VP[head= α , tense= β , num= γ ...]

(generation
perspective)



A roller coaster thrills every teenager

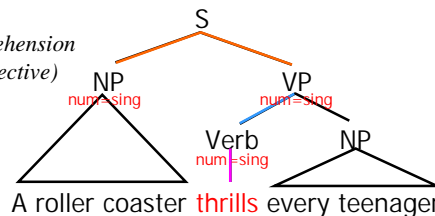
600.465 - Intro to NLP - J. Eisner

5

3 Common Ways to Use Features

Verb[head=thrill, tense=present, num=sing, person=3,...] \rightarrow thrills
 VP[head= α , tense= β , num= γ ...] \rightarrow V[head= α , tense= β , num= γ ...] NP
 S[head= α , tense= β] \rightarrow NP[num= γ ...] VP[head= α , tense= β , num= γ ...]

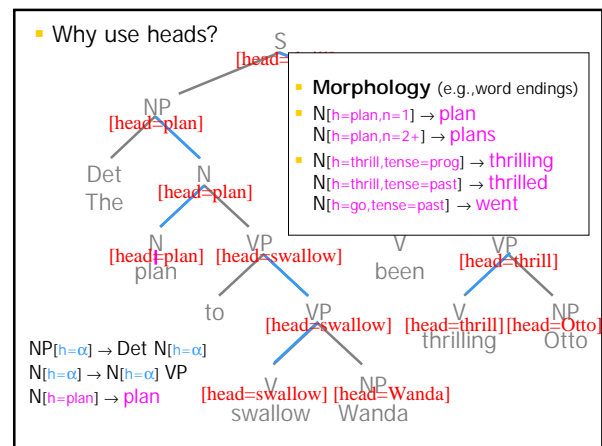
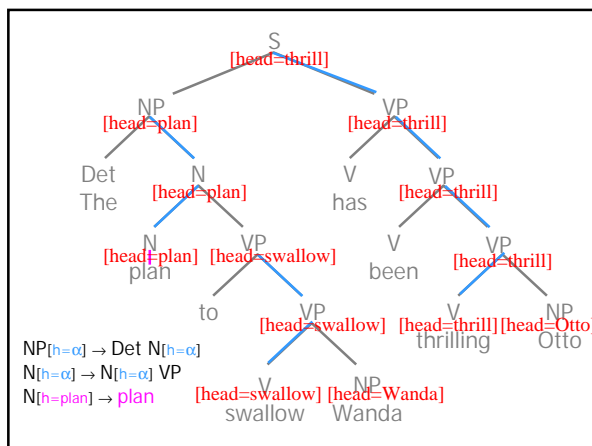
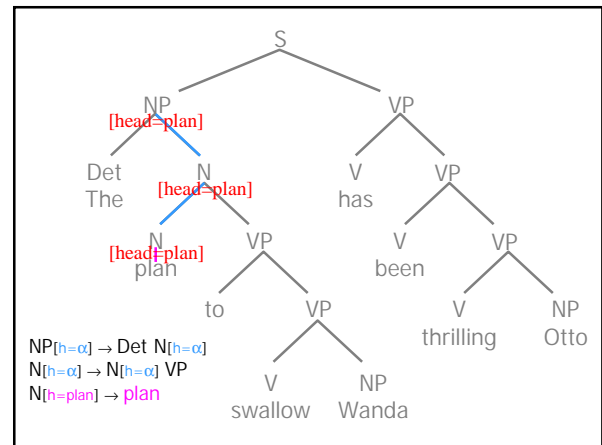
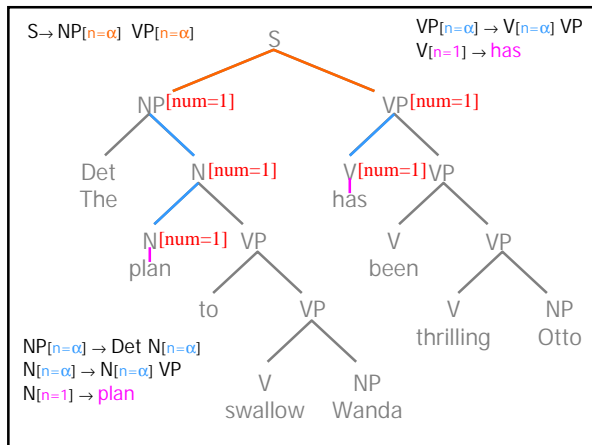
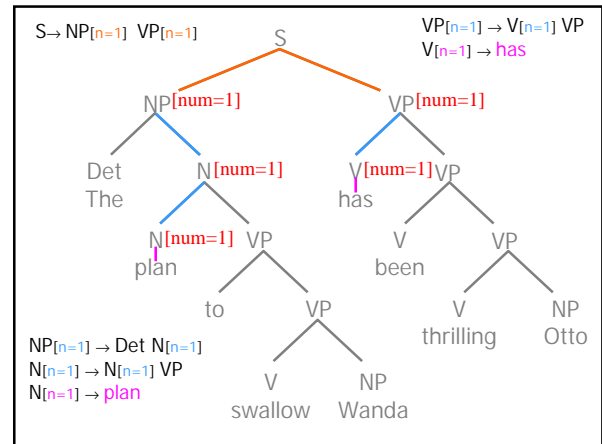
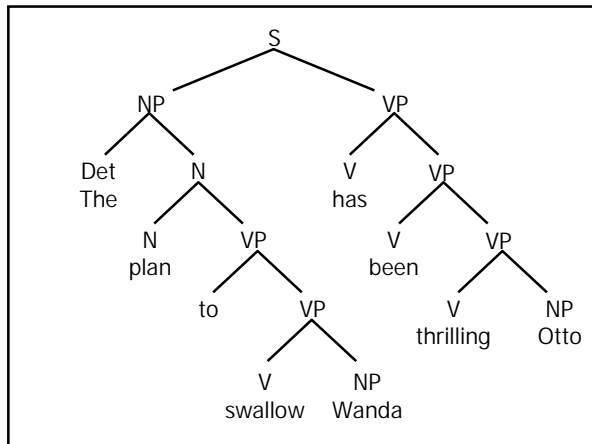
(comprehension
perspective)



A roller coaster thrills every teenager

600.465 - Intro to NLP - J. Eisner

6



Why use heads?

Subcategorization (i.e., transitive vs. intransitive)

- When is $VP \rightarrow V NP$ ok?
 $VP[h=\alpha] \rightarrow V[h=\alpha] NP$
restrict to $\alpha \in \text{TRANSITIVE_VERBS}$
- When is $N \rightarrow N VP$ ok?
 $N[h=\alpha] \rightarrow N[h=\alpha] VP$
restrict to $\alpha \in \{\text{plan, plot, hope, ...}\}$

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

Why use heads?

Equivalently: keep the template but make prob depend on α, β

Selectional restrictions

- $VP[h=\alpha] \rightarrow V[h=\alpha] NP$
- I.e. $VP[h=\alpha] \rightarrow V[h=\alpha] NP[h=\beta]$
- Don't fill template in all ways:
 $VP[h=\text{thrill}] \rightarrow V[h=\text{thrill}] NP[h=\text{Otto}]$
 $VP[h=\text{thrill}] \rightarrow V[h=\text{thrill}] NP[h=\text{plan}]$
leave out, or low prob

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

NP[$h=\alpha$] \rightarrow Det N[$h=\alpha$]
N[$h=\alpha$] \rightarrow N[$h=\alpha$] VP
N[$h=\text{plan}$] \rightarrow plan

Part of the English Tense System

	Present	Past	Future	Infinitive
Simple	eats	ate	will eat	to eat
Perfect	has eaten	had eaten	will have eaten	to have eaten
progressive	is eating	was eating	will be eating	to be eating
Perfect+ progressive	has been eating	had been eating	will have been eating	to have been eating

Let's distinguish the different kinds of VP by tense ...

NP[$h=\text{plan}$] \rightarrow The plan ...

VP[$tense=\text{pres}, head=\text{thrill}$] \rightarrow has

VP[$tense=\text{perf}, head=\text{thrill}$] \rightarrow been

VP[$tense=\text{prog}, head=\text{thrill}$] \rightarrow thrilling

NP[$h=\text{Otto}$] \rightarrow Otto

Past

NP[$h=\text{plan}$] \rightarrow The plan ...

VP[$tense=\text{pres}, head=\text{thrill}$] \rightarrow thrills

VP[$tense=\text{past}, head=\text{thrill}$] \rightarrow thrilled

NP[$h=\text{Otto}$] \rightarrow Otto

Past

NP[$h=\text{plan}$] \rightarrow The plan ...

VP[$tense=\text{pres}, head=\text{thrill}$] \rightarrow thrills

VP[$tense=\text{past}, head=\text{thrill}$] \rightarrow thrilled

NP[$h=\text{Otto}$] \rightarrow Otto

Past

NP[$h=\text{plan}$] \rightarrow The plan ...

VP[$tense=\text{pres}, head=\text{thrill}$] \rightarrow thrills

VP[$tense=\text{past}, head=\text{thrill}$] \rightarrow thrilled

NP[$h=\text{Otto}$] \rightarrow Otto

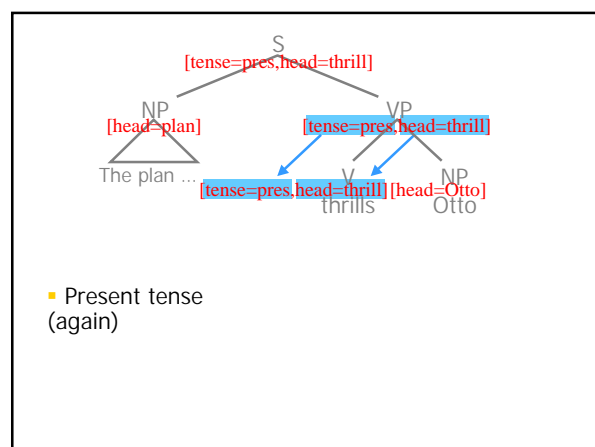
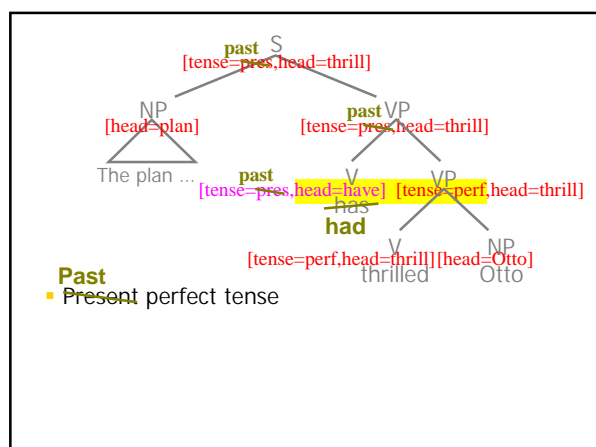
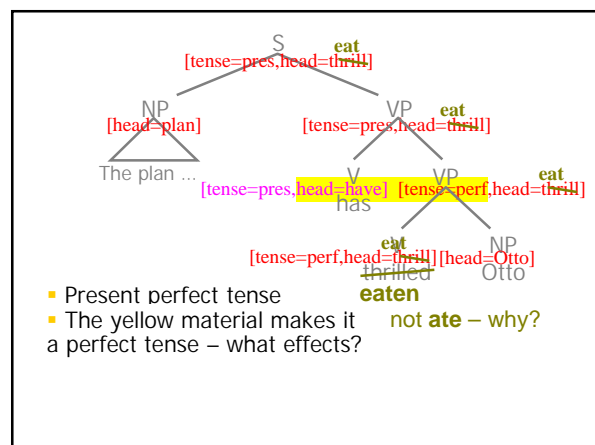
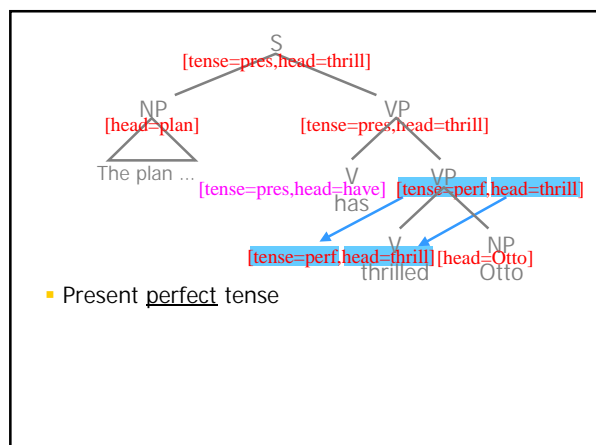
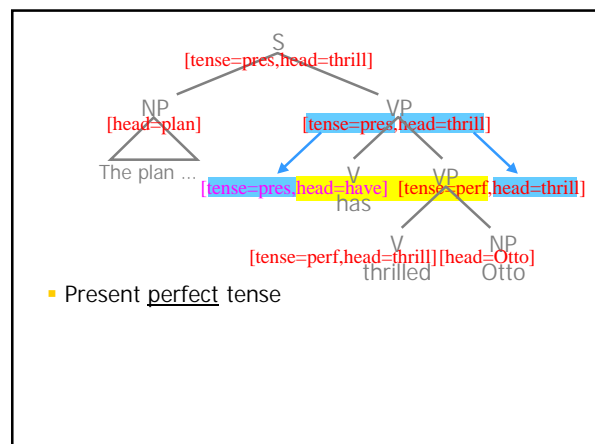
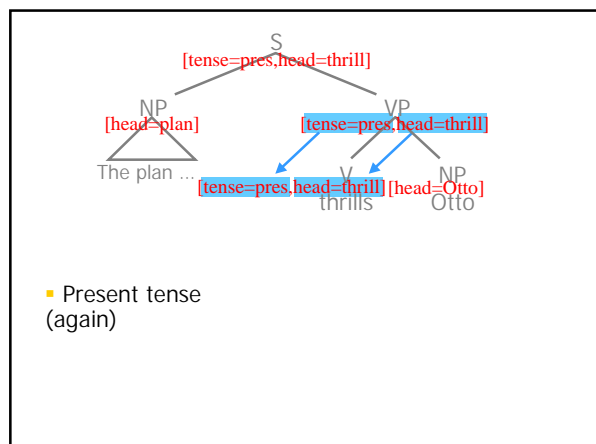
Past

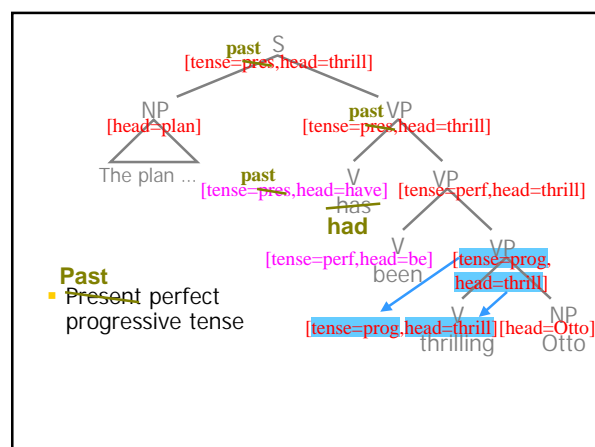
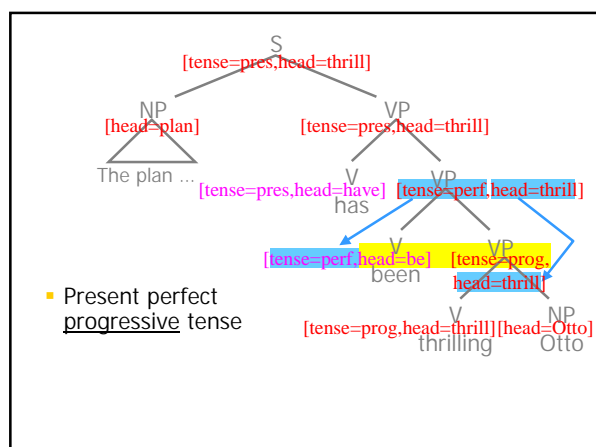
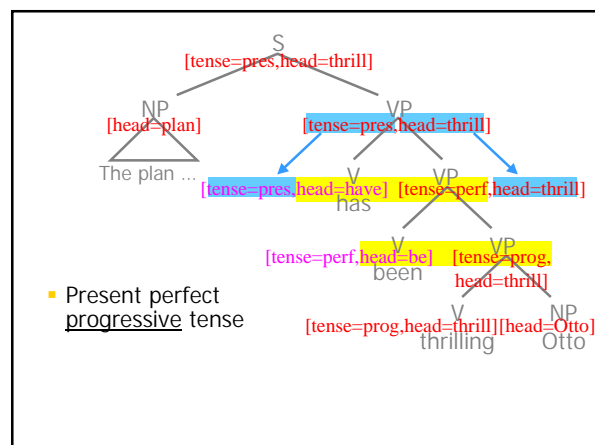
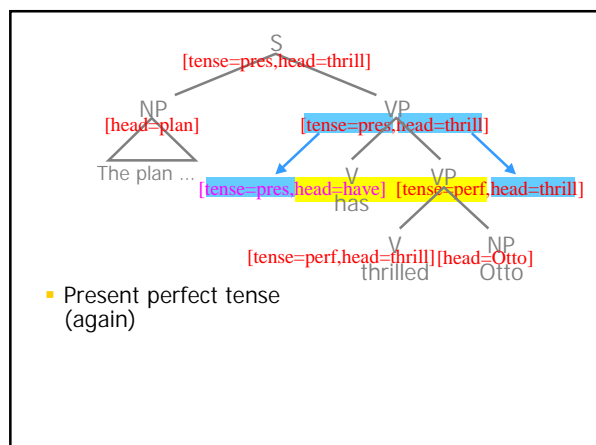
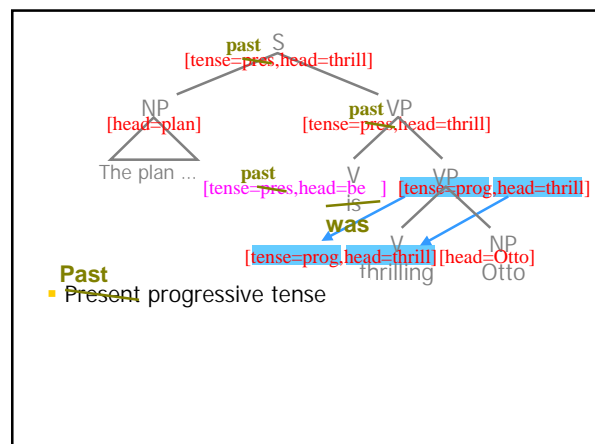
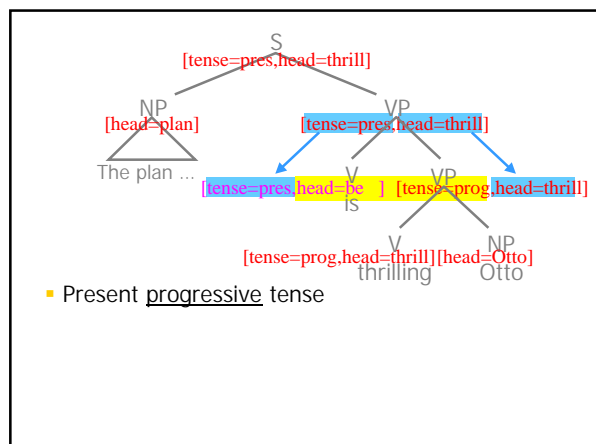
NP[$h=\text{plan}$] \rightarrow The plan ...

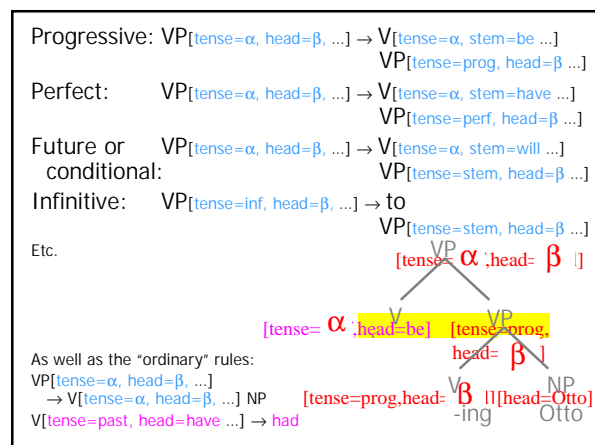
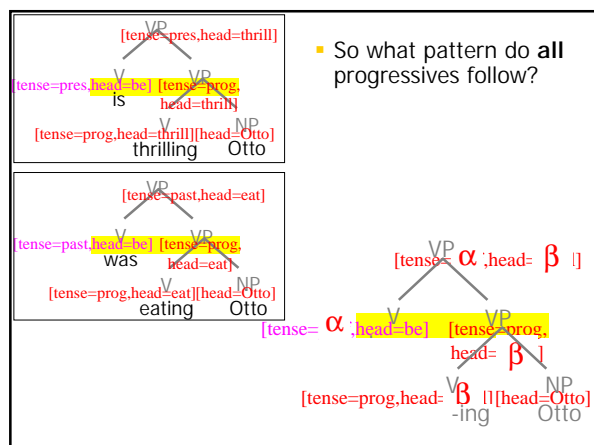
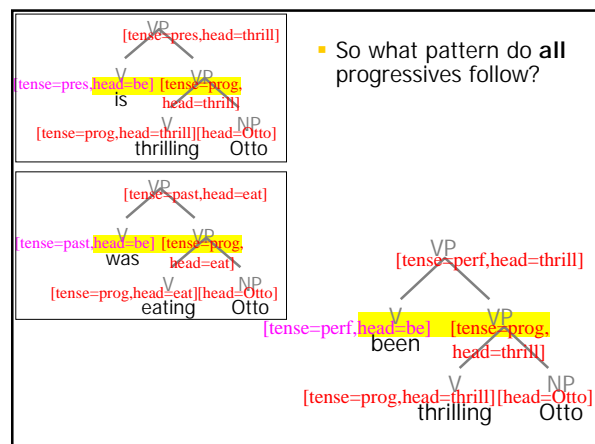
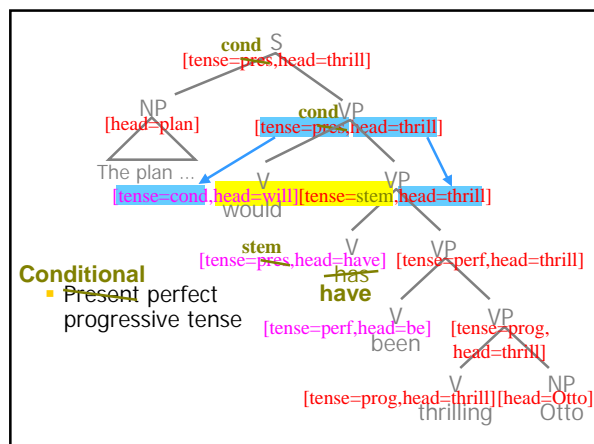
VP[$tense=\text{pres}, head=\text{thrill}$] \rightarrow thrills

VP[$tense=\text{past}, head=\text{thrill}$] \rightarrow thrilled

NP[$h=\text{Otto}$] \rightarrow Otto







Gaps ("deep" grammar!)

- Pretend "kiss" is a pure transitive verb.
- Is "the president kissed" grammatical?
 - If so, what type of phrase is it?

the sandwich that the president kissed e

I wonder what Sally said the president kissed e

What else has Sally consumed the pickle with e

Sally consumed e with the pickle

Gaps

- Object gaps:**
 - the sandwich that the president kissed e
 - I wonder what Sally said the president kissed e
 - What else has Sally consumed the pickle with e
 - Sally consumed e with the pickle

[how could you tell the difference?]

- Subject gaps:**
 - the sandwich that e kissed the president
 - I wonder what Sally said e kissed the president
 - What else has

Gaps

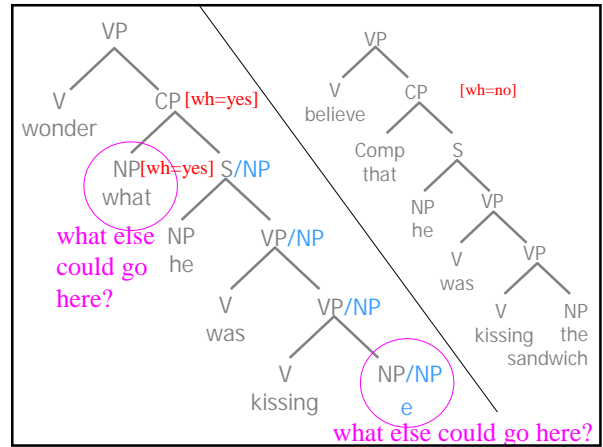
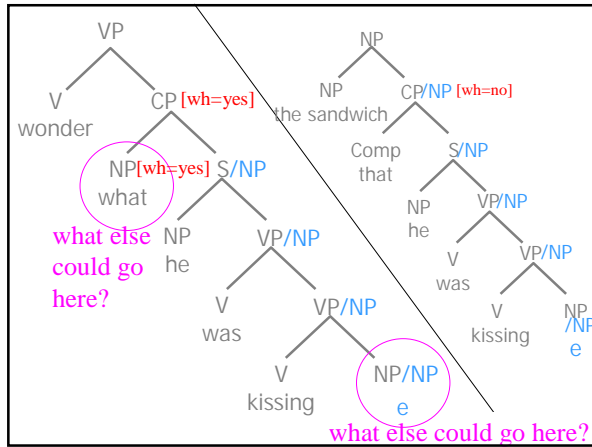
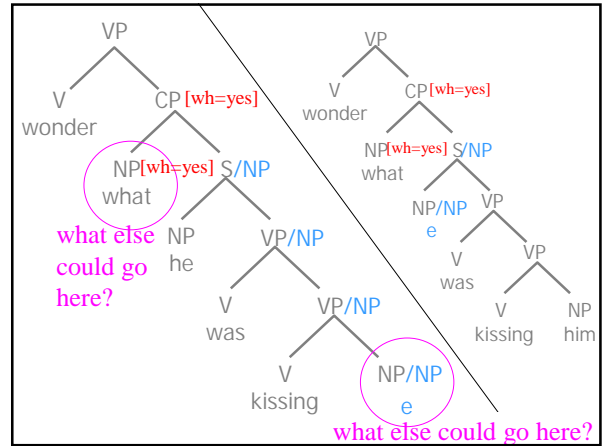
- All gaps are really the same – a missing NP:
 - the sandwich that
 - I wonder what
 - What else has
- the president kissed e
 Sally said the president kissed e
 Sally consumed the pickle with e
 e kissed the president
 Sally said e kissed the president

Phrases with missing NP:

$X[\text{missing=NP}]$
 or just X/NP for short

600.465 - Intro to NLP - J. Finger

37



To indicate what fills a gap, people sometimes "coindex" the gap and its filler.

- Each phrase has a unique index such as "i".
- In some theories, coindexation is used to help extract a meaning from the tree.
- In other theories, it is just an aid to help you follow the example.

the money_i I spend e_i on the happiness_j I hope to buy e_j
 which violin_i is this sonata_j easy to play e_j on e_i

He has gone

- Lots of features (tense, number, person, gap, vowels, commas, etc.)
- Sorry, that's just how language is ...
- You know too much to write it down easily!