

Lec12: Trees

HUL 242

25/2

Phrase Structure rules

- Generalizations about structure represented by phrase structure rules
- Rules generate trees
- A set of PS rules for some simple Hindi sentences aka 'Toy grammar'

Phrase Structure Rules

Re-write rules

$XP \rightarrow X Y Z$

label for constituent is written as \rightarrow elements making up constituent

Noun Phrases

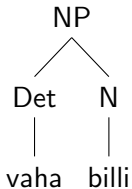
In the simplest form, therefore $NP \rightarrow N$

NP
|
N
|
billi

Noun Phrase

Optional determiner

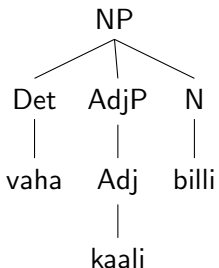
vaha billi



$NP \rightarrow (Det) N$

Noun Phrase

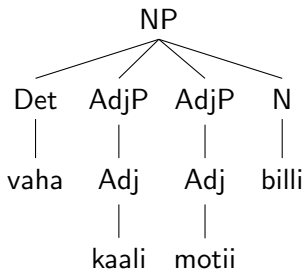
vaha kaali billi



$NP \rightarrow (Det) (AdjP) N$

Noun Phrase

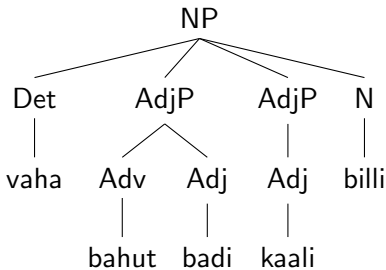
vaha kaali motii billi



$NP \rightarrow (Det) (AdjP^+) N$

Adj Phrase

vaha bahut badi kaali billi



$NP \rightarrow (Det) (AdjP^+)$ N

$AdjP \rightarrow (Adv)$ Adj

Note, we treat *bahut* as an adverb of degree, similar to *very*, which can modify other adjectives *very colourful*

Syntactic Sisterhood

Modifiers are attached within the phrase they modify

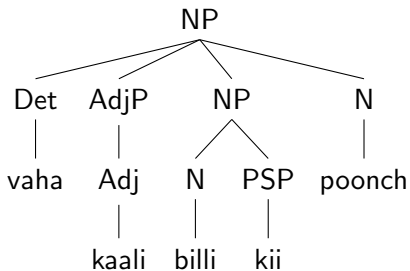
Principle of Modification

If an XP modifies some head Y, then XP must be sister to Y

The AdjP *kaali* modifies the noun *billi*, and it's a sister of N *billi*

Noun Modifiers

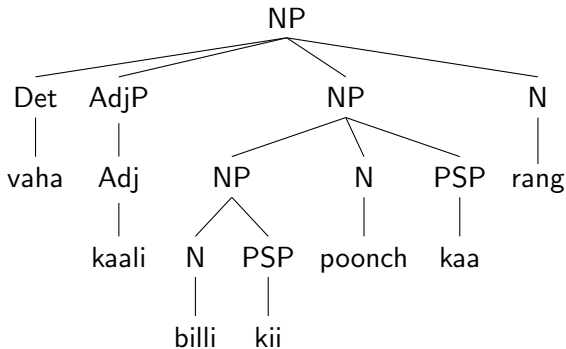
vaha kaali billi ki poonch



NP → (Det) (AdjP+) (NP) N (PSP)

Noun Modifiers

vaha kaali billi ki poonch ka rang



NP → (Det) (AdjP+) (NP+) N (PSP)

Heads of Phrases

Note that the head of the NP phrase is an N
although an optional (NP) exists, $*NP \rightarrow N N$ is not a valid rule

$NP \rightarrow N NP$ is possible

This relationship is asymmetric: the (NP) modifies NP, but not the other way round

NP not PP

- Postpositions in Hindi like *ke/kii/kaa* or *se, ne* and *ko* don't form postpositional phrases (PP)
- This is because these elements are distinguished from postpositions like *ke saath* or *ke upar*, which form PPs
- *ke/kii/kaa* or *se, ne* and *ko* indicate **case** [grammatical roles of nouns] e.g. possessive, instrumental, ergative (agentive) and accusative
- Sometimes, *raam ko, sitaa se* are also known as 'case phrases'
- In English, nouns don't have overt grammatical case markers (except pronouns such as *they* (nom), *them* (acc), *their* (poss))
- For subjects and objects English, grammatical roles are indicated via structural positions i.e. before/after the verb
- All other noun phrases in English require prepositions

Verb Phrases

billi soyi

VP

|

V

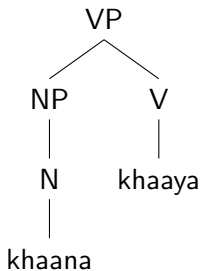
|

soyi

VP \rightarrow V

Verb Phrases

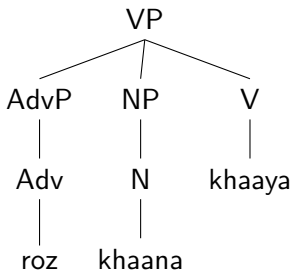
billi ne khaana khaaya



$VP \rightarrow (NP) V$

Verb Phrases

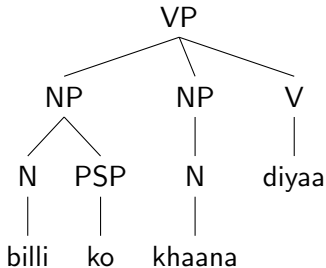
billi ne roz khaana khaaya



VP \rightarrow (AdvP) (NP) V

Verb Phrases

mohan ne billi ko khaana diyaa

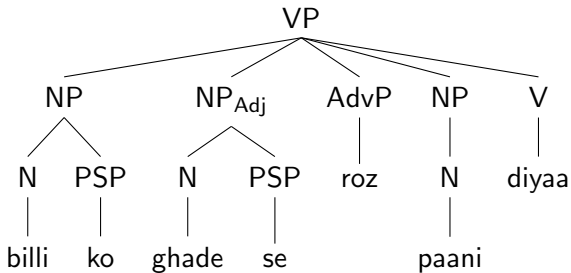


VP \rightarrow (NP) (AdvP) (NP) V

No Kleene plus because we want exactly 2 NPs (not more than 2)

Verb Phrases

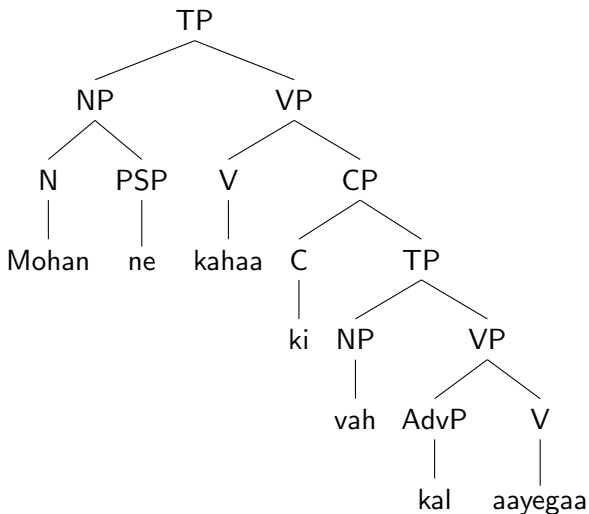
mohan ne billi ko ghade se roz paani diyaa



VP \rightarrow (NP) (NP_{Adj}+) (AdvP) (NP) V

Complementizer Phrase

Mohan ne kahaa [ki vaha kal aayegaa] (CP \rightarrow (C) S)

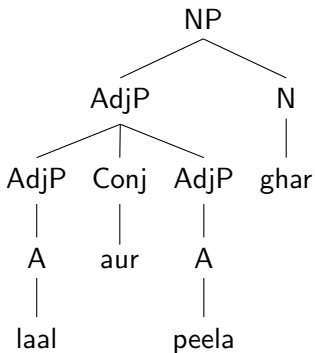


Complementizer Phrase

- The C in a CP is optionally present
- Compare ‘Mohan ne kaha ki vaha kal aayega’ with ‘Mohan ne kaha “main kal aaungaa”’
- S can be embedded inside a CP

Conjunct phrases

laal aur peela ghar



$XP \rightarrow XP \text{ conj } XP$

The Tensed Phrase

$S \rightarrow NP VP$

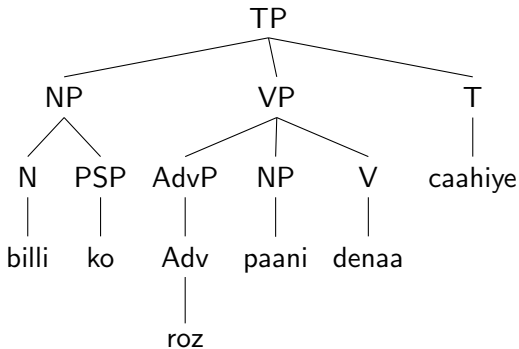
S is sometimes written as TP (Tensed Phrase) or IP (Inflectional Phrase)

Functional categories like auxiliaries and complementizers are also heads

But the T in TP is not always overtly realized

Tensed Phrase

billi ko roz paani denaa caahiye



TP \rightarrow NP VP (T)

Toy grammar

CP \rightarrow TP (C)

TP \rightarrow NP VP (T)

NP \rightarrow (Det) (AdjP+) N (PSP) (NP+)

VP \rightarrow (NP) (NP_{Adj}+) (AdvP) (NP) (AdvP) V (CP)

AdjP \rightarrow (AdvP) Adj

XP \rightarrow XP conj XP

Bottom up approach to drawing trees

mohan ne choti billi ko maara

- Mark part of speech tags
- Identify which items are modifiers and which others are heads
- 'grow' the tree according to the set of PS rules

Structural ambiguity

maine ladke ko maidaan mein bhaagte hue dekha

English

The hunter [killed the elephant] [in his pajamas]

The hunter [killed [the elephant [in his pajamas]]]

PP attachment

Tree structure disambiguates PP attachment in the tree

Major challenge for automatic parsing

PS rules across languages

- Languages with VSO order can have TP \rightarrow VP NP
- Others may not appear to have structure at all e.g. *Nimictumimaka* 'I'll give you money'
- Words that govern syntactic form are similar to those governing word form

Word orders

mohan ne billi ko maara
maara mohan ne billi ko
billi ko mohan ne maara
billi ko maara mohan ne

Is the TP \rightarrow NP VP rule valid?

Change in word order indicates emphasis or contrast

Posit a neutral word order for such languages (sim. Sanskrit, Latin)

References

These slides rely heavily on Andrew Carnie's *A Generative Introduction to Syntax*, Ch 3