Syntax Analysis

Dr. M. Moshiul Hoque

CSE, CUET

Phrasal Categories in Bangla

- Every sentence is composed of one or more phrases.

- **03 types of phrases:**
    - ❑Noun phrase (NP)
    - ❑Verb phrase (VP)
    - ❑Adjective phase (AP)

Noun Phrase (NP)

- The root element of a NP is a noun or pronoun.

  ➢ A single noun or pronoun can act as NP

  ➢ determiner, quantifier, adjective, Bivokti, may accompany with N.

- *pakhi (cvwL)*, *pakhiti (cvwLwU)*, *ekti pakhi (GKwU cvwL)*, *ekti sundor pakhi (GKwU my›'i cvwL)*

Verb Phrase (VP)

- The root of the verb or any of its from can be used as VP.
- A verb from with NPs and/or with an adjective phrase can make up an VP.

❑Underlined words

- *selera khelche* (*ছেলেরা খেলছে*),

- *selera ball khelche* (*ছেলেরা বল খেলছে*),

- *se bhalovabe lekhapora kore* (*সে ভালোভাবে লেখাপড়া করে*)

Verb Phrase (VP)

▪More than one word may need to be grouped into a single verb.

▪**amara _tazmohal darshan korlam_** (আমরা তাজমহল দর্শন করলাম)

▪VP usually used as the predicate of a sentence

▪it combines with a subject NP to form whole sentence.

Adjective Phrase (AP)

- Consists of one or more adjectives only.
- often used to modify nouns & thus appear as the elements of an NP.
  - ❑*se amake ekti <u>mojar</u> boi dilo* (*‡m Avgv‡K GKwU gRvi eB w`j*)
    - ▪ underlined word is an AP.

- AP that appears before the VP, modifies he verb.
  - ❑*se <u>khub druto</u> likhte pare* (*‡m Lye 'a"Z wjL‡Z cv‡i*)

Structure of the Noun Phrase (NP)

- **Recall**

  ❖the root element of a NP is a noun.

  ❖A single noun can act as a NP

  ❖determiner, quantifier, adjective, and Bivokti may accompany with it.

PS rule for Bangla NP

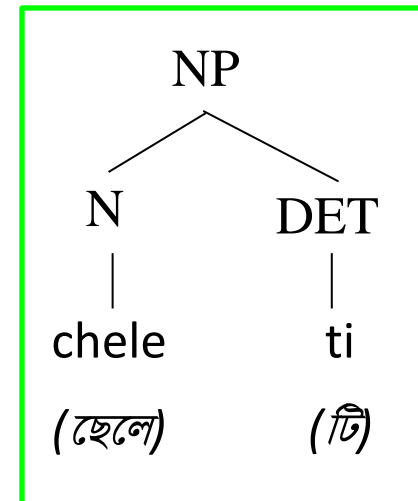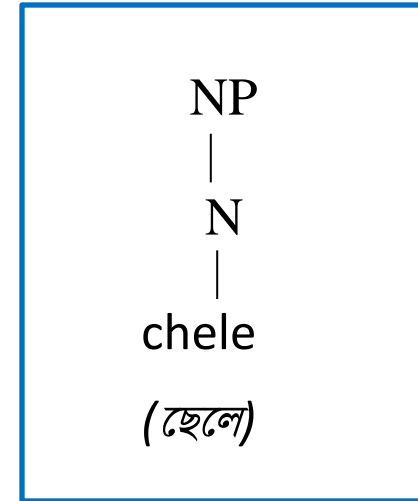| Rule No.	PS rule
01.	NP $\rightarrow$ N (DET)
02.	NP $\rightarrow$ (DEMO) (SPR) (AP) N
03.	NP $\rightarrow$ NP Biv (NP)
04.	NP $\rightarrow$ N PM
05.	SPR $\rightarrow$ QFR (PP)
06.	DEMO $\rightarrow$ (DD) DO)

Rule 1: NP → N (DET)

- one head noun or head noun with a determiner

- () indicates that the use of DET is optional.

- DET always placed into RHS of head noun.
  - ✓ *chele* (ছেলে)
  - ✓ *cheleti* (ছেলেটি)

NP
|
N
|
chele
(ছেলে)

NP
/ \
N   DET
| |
chele ti
(ছেলে) (টি)
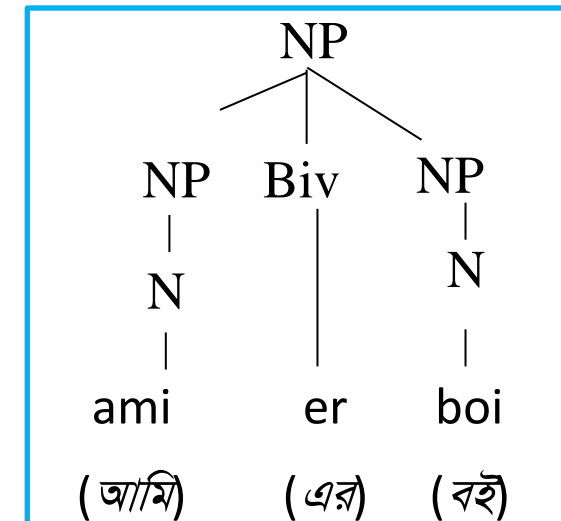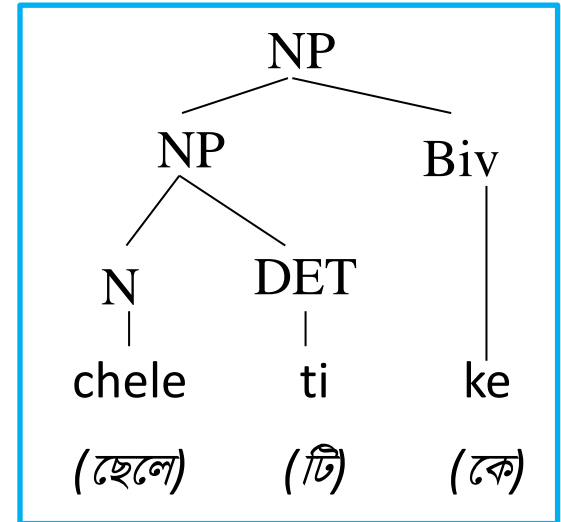
Rule 2: NP → (DEMO) (SPR) (AP) NP

- NP's can be made by noun in conjunction with demonstrator (DEMO) or specifier (SPR) or adjective phrase (AP) or combination any two or three optional part.

Rule 3: NP → NP Biv (NP)

- Bivokti may appear at the end of a NP or between two NPs.
- *ke ($\neq K$), dige ($w‘\neq M$), digoke ($w‘M\neq K$), der ($\neq‘i$), digere ($w‘\neq M\neq i$), re ($\neq i$), ra (iv), er (Gi), diger ($w‘\neq Mi$), kar (Kvi), ker ($\neq Ki$)* etc.
  - ✓ **amar boi ($Avgvi\ eB$)**
  - ✓ **cheletike ($\neq Q\neq jwU\neq K$)**

NP
├─ NP
│ ├─ N — chele (ছেলে)
│ └─ DET — ti (টি)
└─ Biv — ke (কে)

NP
├─ NP — N — ami (আমি)
├─ Biv — er (এর)
└─ NP — N — boi (বই)

Rule 4: **NP → N PM**

- There are a number of plural markers:
- *ra (iv), era (Giv), gulo (ˌ≠jv), der (≠'i), gon (Mb), mala (gvjv), borga (eM©),* etc.

Head Noun	Plural Marker	Plural Form
manush (মানুষ)	**era (এরা)**	**manushera (মানুষেরা)**
chele (ছেলে)	**ra (রা)**	**selera (ছেলেরা)**
chele (ছেলে)	**der (দের)**	**cheleder (ছেলেদের)**
boi (বই)	**gulo (গুলো)**	**boigulo (বই গুলো)**

Rule 4: **NP → N PM**

- **Mostly used**: *ra (ির), era (এর), gulo (গুলো), der (দের*
  - ☐ *ra (ির),* & *era (এর)* are used with nouns of the subject
  - ☐ *der (দের)* is used with non-subject nouns
  - ☐ *gulo (গুলো)* is used with any one.

- *ami cheleder moth dekhlam (আমি ছেলেদের মঠ দেখলাম)*
- *chelera khelchhe (ছেলেরা খেলছে)*
- *ami boigula porechhi (আমি বই গুলো পড়েছি)*
- *pakhi gulo urchhe (পাখি গুলো উড়ছে)*
  - ➤ **PM:** *cheleder (ছেলেদের), chelera (ছেলেরা), boigulo (বই গুলো)*

Rule 5: SPR → QFR (PP)

- Specifier (SPR) (wb‡`©kK) is composed
  - ❑ quantifier (QFR) (cwigvcK)
  - ❑ post-position (PP) (AbymM©)

- PP has been used as an optional.

- SPR indicates the number, quantity, definite or indefinite etc. of the head noun.
  - ✓ In any sentence, if the subject is N, then SPR can be placed into the left-hand-side of it

Quantifier (QFR)

- **Numeral words**: *ek (GK), dui (ᵧB), tin (wZb),* etc.

- **Collective quantifier**: *annek (A≠bK), bahu (eû), sob (me), koek (K†qK), proti (cÖwZ), prottek (cÖ≠ZK),* etc.

- ❑*bohu lok (eû †jvK),*
- ❑*anek paki (A≠bK cvwL)*

Post Position (PP)

- PP optional part of the specifier.
- Used to indicate either definite or indefinite nouns.
  - ➢ In an indefinite NP with PP, PP placed on the LHS of the head noun
  - ➢ In definite NP, PP has been placed on RHS of the head noun.
- ***ekjon boddrolok (GKRb), doshta boi (ˈkUv eB).***
  - ❏ *ekjon (GKRb) & doshta (ˈkUv)* indicates as a specifier.
  - ❏ *jon (Rb) & ta (Uv)* indicates the PP
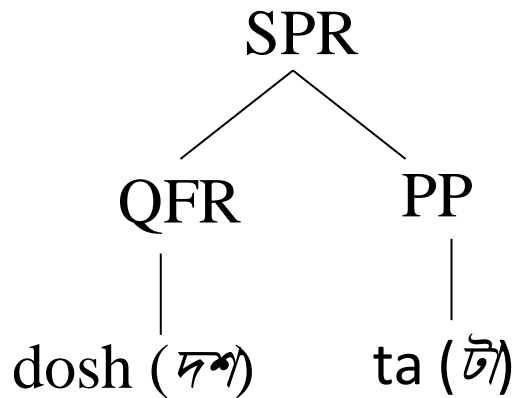- *ti (wU), ta (Uv), khana (Lvbv), khani (Lvwb)*, etc. are used as PP

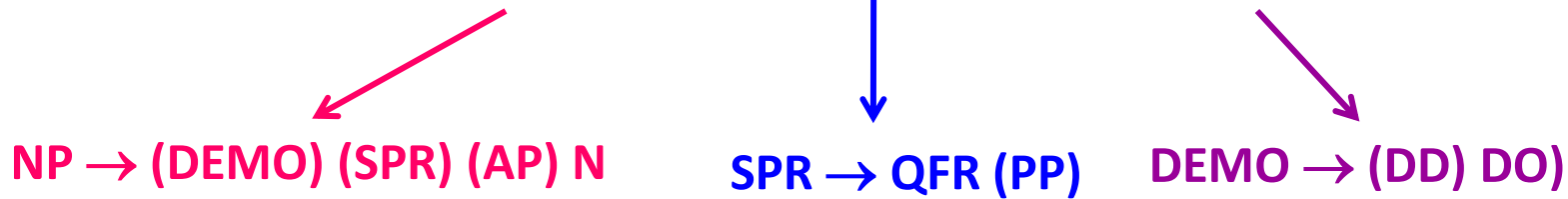# *ekjon (GKRb) & doshta (ˈkUv)*

SPR
- QFR — PP
  - ek (এক) — jon (জন)
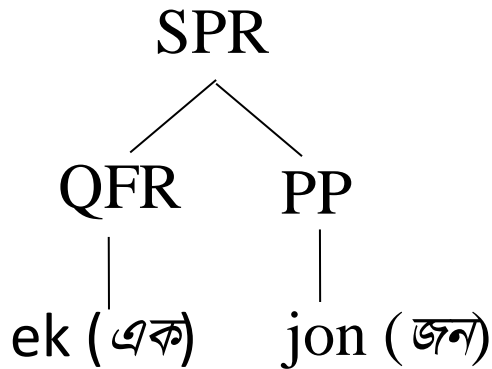
- There is some specifier where PP is absent.
- *anek lok (অনেক লোক)*
- *aanek (অনেক)* indicates the specifier that has contained only the quantifier part.

SPR
- QFR — PP
  - dosh (দশ) — ta (টা)

SPR
- QFR
  - anek
  - (অনেক)

Rule 2 + Rule 5 + Rule 6

NP → (DEMO) (SPR) (AP) N

SPR → QFR (PP)

DEMO → (DD) DO)

- **NPs:**
  - ✓ *ekti bhalo chele (GKwU fvj †Q‡j)*
  - ✓ *ekti chele (GKwU †Q‡j)*
  - ✓ *ooe sudor cheleti (IB my›`i †Q‡jwU)*
  - ✓ *ooe prothom sundor cheleti (IB cÖ_g my›`i †Q‡jwU)*

NP → SPR AP NP

NP
- SPR
  - QFR — ek (এক)
  - PP — ti (টি)
- AP
  - AD — bhalo (ভাল)
- NP
  - N — chele (ছেলে)

NP → DEMO AP NP

NP
- DEMO
  - DD — ooe (ওই)
- AP
  - AD — sundor (সুন্দর)
- NP
  - N — chele (ছেলে)
  - DET — ti (টি)

NP → SPR NP

NP
- SPR
  - QFR — ek (এক)
  - PP — ti (টি)
- NP
  - N — chele (ছেলে)

NP → DEMO NP

NP
- DEMO
  - DD — oee (ওই)
  - DO — Prothom (প্রথম)
- NP
  - N — chele (ছেলে)
  - DET — ti (টি)
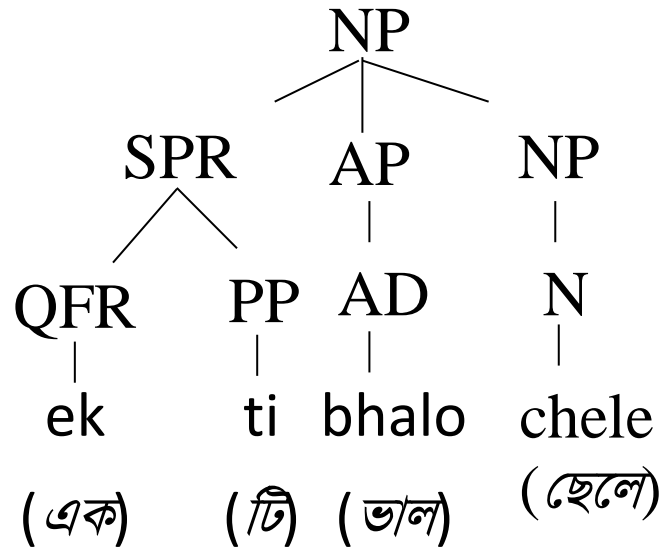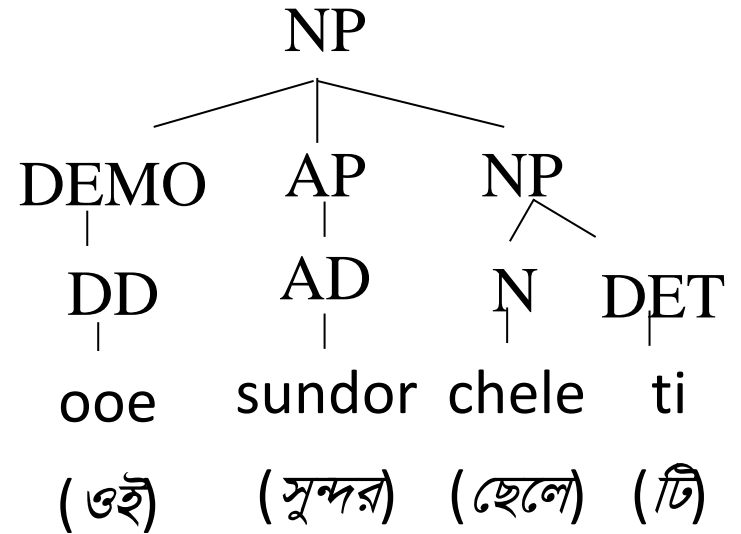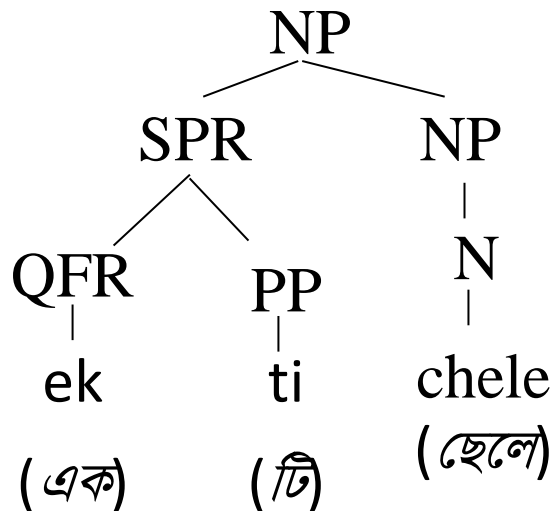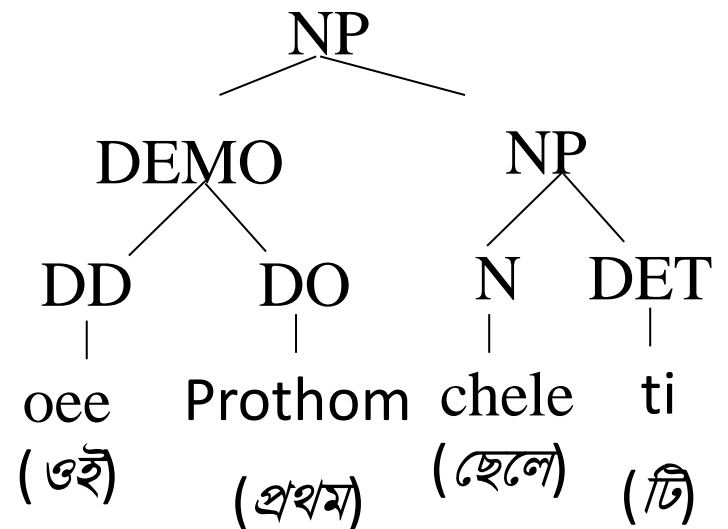
Rule 6: DEMO → (DD) (DO)

- Demonstrator (DEMO) combined
  - ❑demonstratic deictic (DD) (ms‡KZ evPK cÖ‵k©K)
  - ❑demonstratic ordinal (DO) (µgmsL¨v)
- ***DD****: a (G), oo (I), se (‡m), aae (GB), oee (IB) etc.*
- ***DD:*** *prothom (cÖ_g), ditria (wØZxq), tritia (Z...Zxq) etc*

- ▪ demonstrator words used with specifier or determiner.
  - ❖*ooe cheleti (IB †Q‡jwU)*
- ▪ DD are generally definite and cannot be used without a determiner

Structure of the Verb Phrase (VP)

- Every sentence in Bangla must have a VP
  - ➢The root of the verb or any of its form can be used as a VP
  - ➢A verb form (VF) with a NP
  - ➢with an AP
- The compulsory of a VP is the **VF &** optional part may contain a NP/AP

PS Rule for VP

(a) **VP → (NP) (AP) VF**

(b) **VF → VR AUX**

Verb Form (VF)

- **Verb** is so called because it is the most important word in a sentence.

  ❑it is the core or backbone of every sentence

  ❑It is the word used for stating something about a person or thing

- Segmented into two parts.

  ❑(i) Verb Root (VR)

  ❑(ii) Auxiliary (AUX)
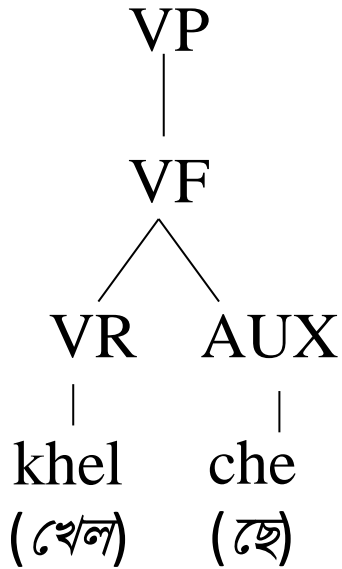
Verb Root (VR)

- Verb root cannot be divisible

- Appearance of the VR never changes due to the change of the number or gender of the subject
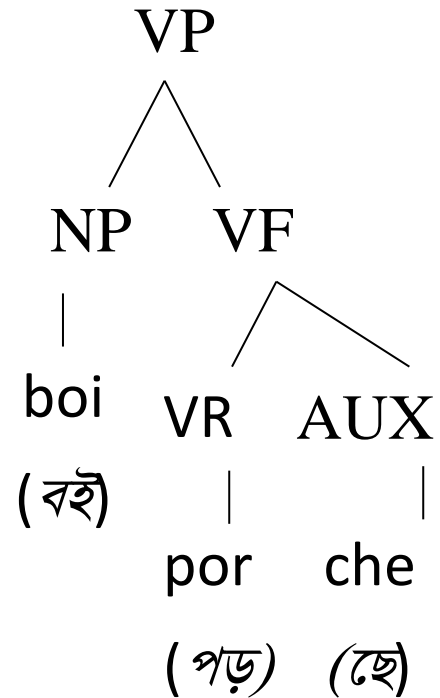
Auxiliary (AUX)

- **03 parts:**
  - ☐ Aspect
  - ☐ Tense
  - ☐ Person + Class (Concord)
- Concord depends on person and class.

- AUX depends on the person & class of the subject & also on the tense
  - *se khelche (ৰm ⊥Lj≠Q)*
  - *karim boi porche (Kwig eB co≠Q)*
  - *ami take bhalobabe chini (Avwg Zv≠K fv≠jvfv≠e wPwb)*
- Underlined words indicate the VP's.

se _khelche_ (সে খেলছে)

```
        VP
        |
        VF
       /  \
     VR    AUX
     |      |
    khel   che
   (খেল)   (ছে)
```

karim _boi porche_ (করিম বই পড়ছে)

```
          VP
         /  \
       NP    VF
       |     / \
      boi  VR   AUX
     (বই)  |     |
          por   che
         (পড়)  (ছে)
```

Another way to extends the VP rule:

karim <u>*boi porche*</u> (*Kwig* <u>*eB co≠Q*</u>)
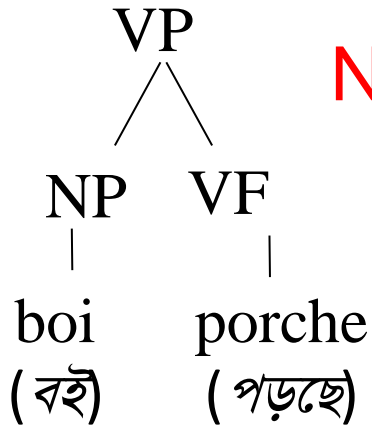
NP → AP VF

ami <u>*take bhalobabe chini*</u> (*Avwg* <u>*Zv≠K fv≠jvfv≠e wPwb*</u>)

NP → NP AP VF
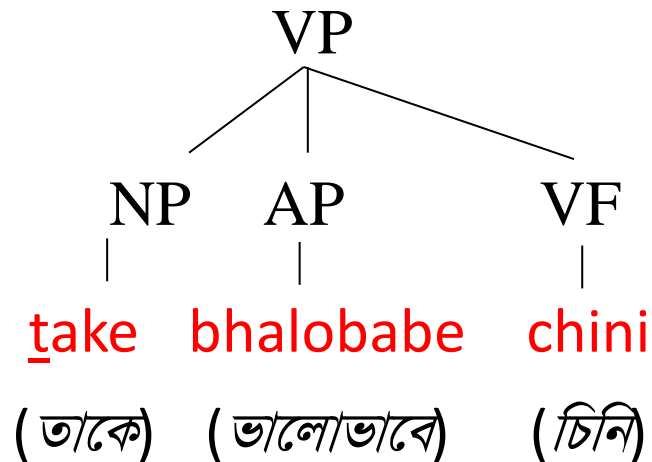
Structure of the Adjective Phrase (AP)

- Consists of one or more adjective only,
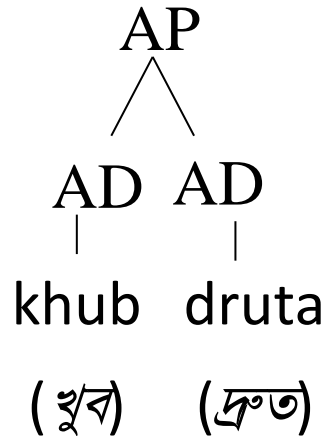   - APs are often used to modify nouns & this appears the element of an NP
    - AP that appears before the VP modifies the verb
    - Generally, AP is used just before the head noun or NP or VP.

$$\boxed{\textbf{\textcolor{blue}{AP} → \textcolor{blue}{AD / ADs}}}$$

rahim **<u>khub durta</u>** *likhe*

(iwng <u>Lye 'a"Z</u> wj≠L)

AP
├ AD ─ AD
│ khub · druta
(খুব) · (দ্রুত)

AP
│
AD
│
bhalo
(ভালো)

rahim **<u>bholo</u>** *chele (iwng fv≠jv †Q≠j)*

Person and Class in Bangla

03 types of classes
- ➤ Honorific
- ➤ Pejorative/ Negligible
- ➤ General/ Non-Honorific

03 types of person
- ❑ First Person
- ❑ Second Person
- ❑ Third Person

Person and Class in Bangla

❑ Classes are not applicable for all kinds of person

➢ First person has no classes

➢ Second person consists of 03 classes

  ❖ Second person Honorific

  ❖ Second person Non-Honorific

  ❖ Second person Pejorative

➢ Third person consists of 02 classes

  ❖ Third person Honorific

 ❖ Third person Non-Honorific

Classification of person and classes

Person	Singular	Plural	Class
First	ami (আমি)	amara (আমরা)	No Class
	apani (আপনি)	apanara (আপনারা)	Honorific
Second	tumi (তুমি)	tomara (তোমরা)	Non-Honorific
	tui (তুই)	tora (তোরা)	Pejorative
Third	tini (তিনি)	tara (তারা)	Honorific
	se (সে)	tara (তারা)	Non-Honorific

Tenses in Bangla

- 'tense' comes from the Latin word 'tempus' that is meant 'time'.
  - ➤ tenses show the time of action and its degree of completeness.
- tense is a form of the verb showing the time of the happening of an action.
- 03 kinds of tense in Bangla grammar.
- The inflection of Bangla verb can have several forms depending on the tense of the subject of the verb

Tenses in Bangla
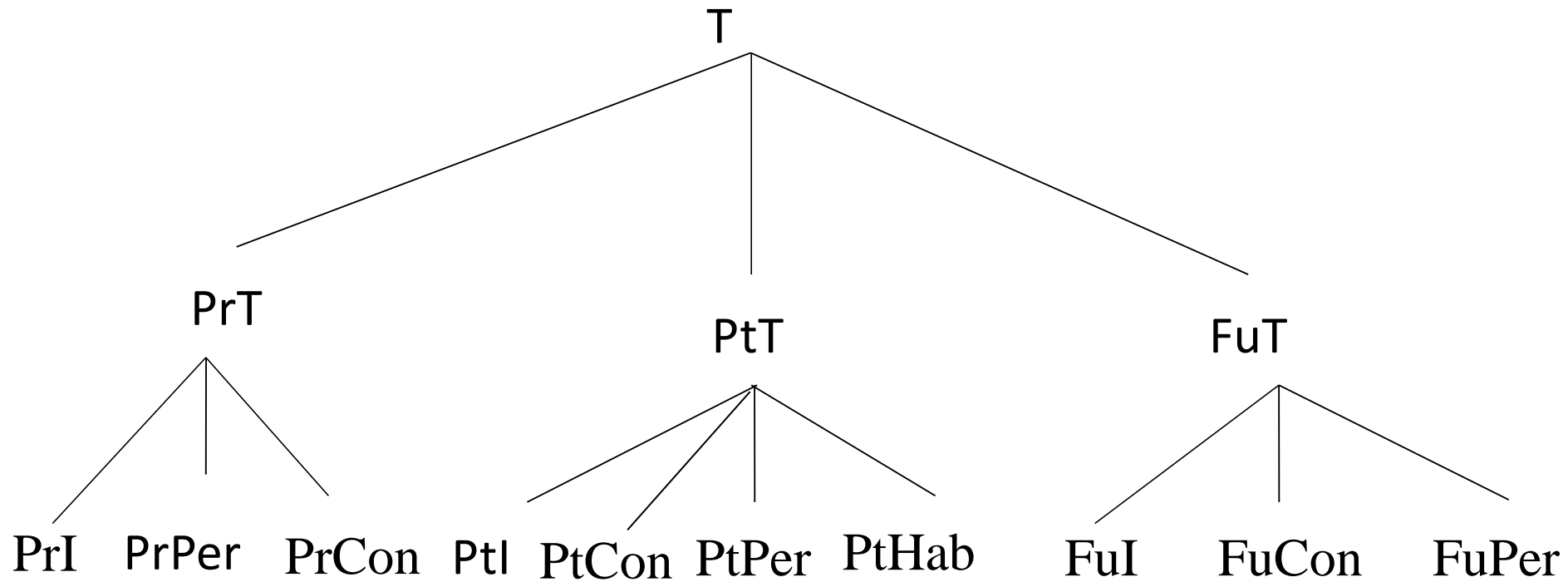
Conclusion

- Phrasal analysis plays key role in NL understanding phase during Parsing and hence MT
- 03 types of phrase analysis & establishes a set of rules to describe the each phrase in Bangla.
- Can be Included more phrasal categories
  - ➢ adverbial phrase,
  - ➢ inter-jectional phrase
- we find the more rules & describes the given language sentences more effectively