# Paninian Grammar

Processing Indian Languages

# Nature of Indian Languages

- English Language has been modelled using CFG as the grammar formalism.

- But CFG may not be a viable choice for Indian languages. The reasons are the properties of Indian languages:

1. Asian languages are **SOV (Subject-Object-Verb) ordered** and inflectionally rich.

2. **Word-order free** - Some languages, like Sanskrit, have the flexibility to allow word groups representing subject, object and verb to occur in any order. Position of object and subject can be changed.

# Free Word-Order

- ***Maan bachche ko khaanaa detii hai***

- Mother child to food give

- Mother gives food to the child.

- .

- ***Bachche ko Maan khanaa detii hai***

- Child to mother food give

- Mother gives food to the child

# Need for Paninian Grammar

- Due to differences between English and Indian languages, there is a need to find a new framework for handling Indian languages.

- Verb groups are formed differently in Indo-Aryan and Dravidian languages.

- **Sanskrit** is different from other Indian languages because it has **5 tenses and 3 numbers** – The translation of "He goes" and "He is going" is same.

- **Hindi** is unique – it has **no neuter gender**.

- All nouns are categorized as masculine or feminine, verb form must have a gender agreement with the subject.
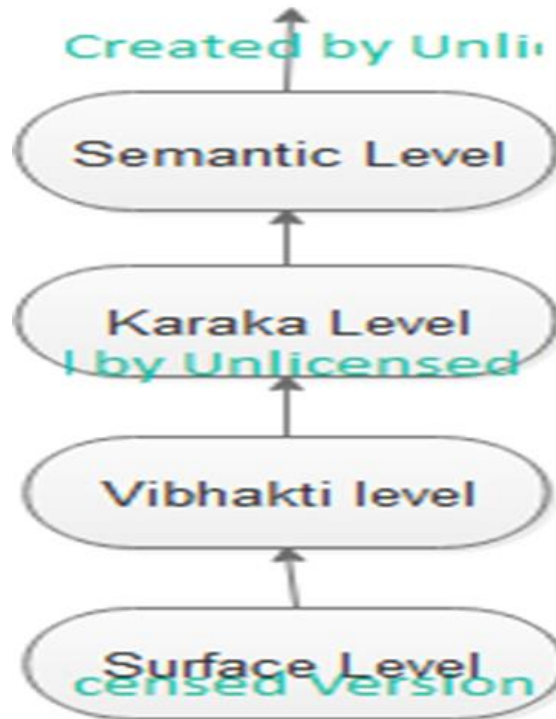
# What is Paninian Grammar?

- Paninian Grammar was written by Panini in 500 BC in Sanskrit.

- This framework can be used for other Indian languages and Asian languages as well.

- Paninian Grammar seeks to provide complete, maximally concise and theoretically consistent analysis of Sanskrit Grammatical Structure.

# Components of Paninian Grammar

| | |
|---|---|
| **Astadhyayi** | System of 400 grammatical rules |
| **Swasutras** | Inventory of phonological segments |
| **Dhatupatha** | List of about 2000 verbal roots consisting of subclassification |
| **Ganapatha** | Inventory of classes of lexical items associated with rules |

# Syntactico-Semantico Framework

- Paninian Grammar framework is said to be **Syntactico-Semantico** – because one can go from surface layer to deep semantics by passing through intermediate layers.

Created by Unli

Semantic Level

Karaka Level
by Unlicensed

Vibhakti level

Surface Level
censed Version

# Levels of Paninian Grammar - explained

- **Surface level** – is the uttered or the written sentence. It is the bottommost level in the hierarchy.

- **Vibakthi level** -  level at which there are local word groups based on case endings, preposition or postposition markers. Vibakthi for verb gives information about Tense Aspect and Modality(TAM). Therefore it is called TAM label. TAM labels are purely syntactic, determined from verb form and auxiliaries.

- **Karaka level** – Between the Vibakthi and the semantic level is the karaka level. It includes karaka relations and information about karaka, karma , Sampradhana and Vyaapara – elements of Karaka Theory.

- **Semantic level** – is the topmost level. It refers to what speaker thinks in his mind.

# Semantic Model of Paninian Grammar

Every verbal root (dhaatu) denotes an action consisting of:

1. an activity (a vyaapaara), and

2. a result (a phala)

*Result* is the state which when reached, the action is complete. *Activity* consists of actions carried out by different participants or *karakas*[3] involved in the action.

# Action

- An action is a complex which could be divided into sub-actions.

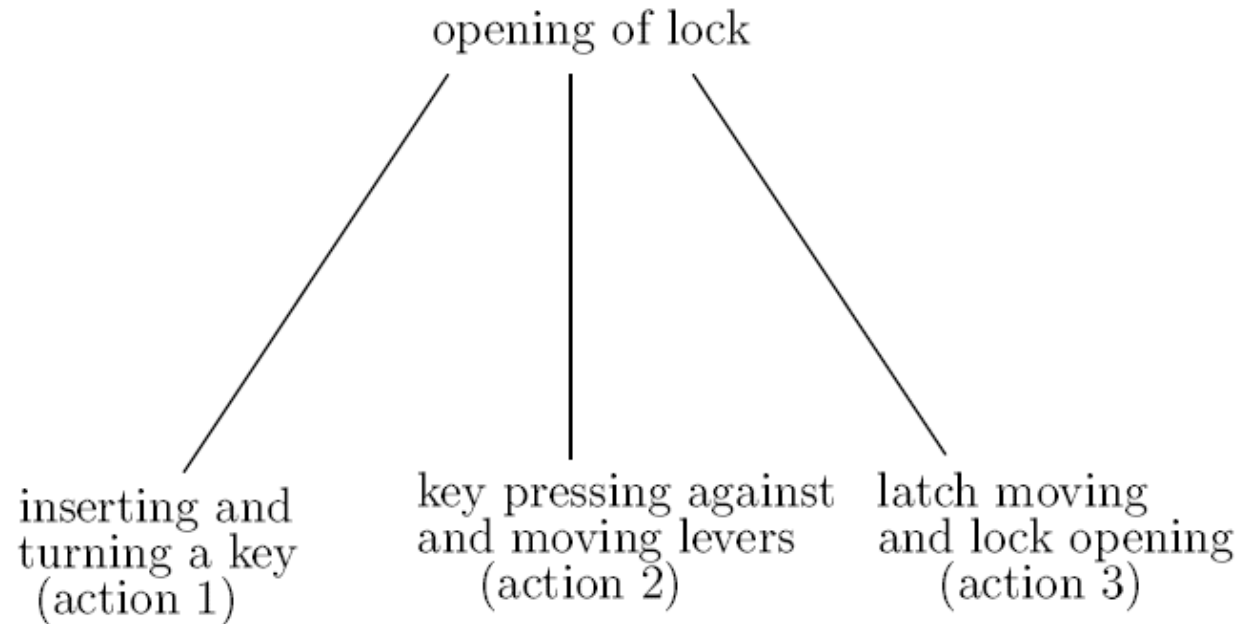- **Each** of the **sub-actions** has the **semantic relationships** with associated objects.

opening of lock

inserting and
turning a key
(action 1)

key pressing against
and moving levers
(action 2)

latch moving
and lock opening
(action 3)

Figure 5.1: Structure of an action

# Vivaksha and Swatantra

- Vivaksha refers to the speaker's viewpoint or attitude towards the activity,

- A sentence has
  - Statement of the objective activity
  - Information regarding speaker's viewpoint.

Out of all the participants in an action, there is one who is 'swatantra' or most independent, and is called $karta^4$ karaka. Thus, in the following sentences:

```
A.1 The boy opened the lock.
A.2 The key opened the lock.
A.3 The lock opened.
```

the karta is, respectively, the boy, the key, and the lock.

# Karta, Karma , Karana and Sampradhana

- Karta is the ashraya (locus) of the activity.

- Karma is the object.

- Karana karaka is the instrument. With the vyapara of the karana, the phala(result) is immediately achieved.

- Consider "**Rama killed Ravana with an arrow**"

- Rama = Karta , Ravana = Karma, Karana = arrow

- Sampradhana is the beneficiary.

- Consider "**the boy gave a rose to the girl**"

- The girl is the sampradhana.

# Issues in Paninian Grammar

- **Utsarga-apvada:** It is a multi-layered implementation – an approach to implement Paninian Grammar.
  - Rules are arranged in multiple layers such that each layer consists of rules which are in exception to rules in the higher layer.

- The two problems challenging are:
  - Computational implementation of Paninian Grammar
  - Adaptation of Paninian grammar to Indian and other languages.

- Another main problem is that mapping between Vibakthi and the semantic relation is not one to one.
  - Two different Vibakthi can represent same relation or the same Vibakthi can represent different relations in different contexts.

# Thank You