* Overview of linguistics :-

- "Linguistics" is the scientific study of language.
- It involves the analysis of language form, language meaning and language in context.
- Linguists analyse the human language by observing an interplay between sound and meaning.

→ We can also consider four major sub-disciplines :-

1. Historical linguistics.
2. Syntax and morphology.
3. Semantics and pragmatics.
4. Phonology.

* Grammers and Language :-

- Linguistics have attemped to define grammers since the inception (उत्पत्ति) of natural languages like English, Sanskrit etc.

- Noam Chomsky gave a mathematical model of grammer in 1956, which is effective for writing computer languages.

* Grammer:-

A grammer G can be formally written as a 4-tuple (N, T, S, P) where:-

- N or V_N is a set of variables or non-terminals symbols.
- T is a set of Terminal symbols.
- S is a special variable called the Start symbol, $S \in N$.
- P is Production rules for Terminals and Non-Terminals.

→ four types of grammers:-

- Type 0:- Most general, with no restrictions.
- Type 1:- Context Sensitive. The Right hand side of the Production must contain atleast as symbols.

__/__/____

- Type 2 :- Context-free. The left side of Production must contain exactly one symbol.

- Type 3 :- Regular Expressions.

Example :-

Grammer G1 -

(ξ S, A, B ϒ, ξ a, b ϒ, S, ξ S → AB, A → a, B → b ϒ)

Here :-

- S, A, B are Non terminal Symbols.
- a, b are Terminal Symbols.
- S is the Start Symbol.
- Productions, P : S → AB, A → a, B → b

* Languages :- NLP is the capability of computer software to understand the natural language.

→ There are variety of languages in the world.
Each language has its own structure.
↓
Grammer
↓
Certain Set of rules.

__/__/____

Certain Set of Rules.
↓ determine.

what is allowable.
what is not allowable.

for example:- In English, S V O
= = = =
↓ ↓ ↓
Subject Verb Object

I eat Apple.

other languages:- S O V
O S V.

⇒ The set of all strings that can be derived from a grammer is said to be the language generated from that grammer.

A Language generated by a grammar G_1 is a subset formally defined by:-

$$L(G_1) = S \Rightarrow G_1 \omega$$

$$L(G) = \{ W / W \in \Sigma^*, S \Rightarrow G_1 W \}.$$