- Introduction
- Why NLP
- Application 0f NLP
- Components of NLP
- Overview of linguistics .
- Grammar and Languages .

Natural Language Processing

- A sub-field of artificial intelligence
- AIM: to build intelligent computers that can interact with humans like human being

- how human being communicate with each other.
- computer should replicate the same thing

Natural languages

- It refers to the languages spoken by people
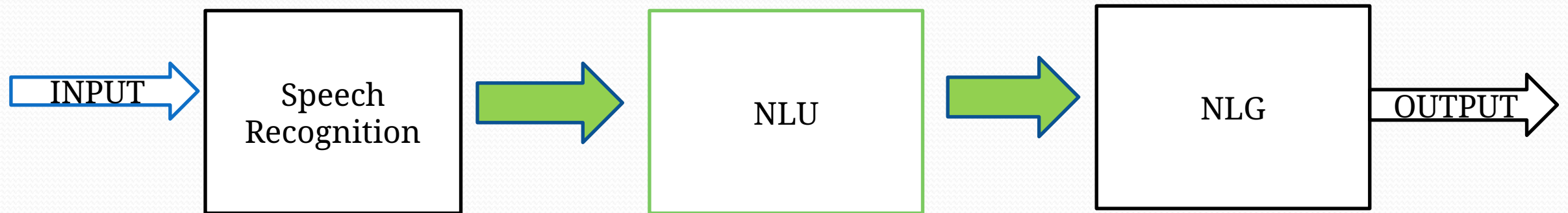- E.g. English , Japanese, as opposed to artificial languages like c++, java etc.

why NLP

- Huge amount of data .
- Internet = at least 20 billion pages
- ☐ Text Data - web sites, blog, tweets
- ☐ Audio Data - speech etc.

Application of NLP

I.   Speech Recognition
II. Sentiment analysis
III. Machine translation
IV. chat bots etc.

INPUT → Speech Recognition → NLU → NLG → OUTPUT

Components of NLP

1. Natural Language understanding.(NLU)
2. Natural languages generation.(NLG)

NLU:

What do the users say ? their intention? Meaning?

Ambiguity.

➢ challenges:

Lexical ambiguity

- actual it is tokenized
- two or more possible meanings for a single word

- E.g. : the tank was full of water .

Syntactic ambiguity:

- *syntactic ambiguity* is the presence of two or more possible <u>meanings</u> within a single <u>sentence</u> or sequence of <u>words</u>.

- Also called structural ambiguity

- E.g. : old men & women were taken to safe place .

Semantic ambiguity

- occurs whenever one or more words within a sentence can be understood as having two or more distinct meanings

- Eg : car hit the pole while it was moving .

Pragmatic ambiguity

- Arises when the statement is not specific, and the context does not provide the information needed to clarify the statement

- E.g.   The police are coming.

Natural language generation(NLG)

What should we say to user ?

- ☐ It should be intelligent and conversational
- ☐ Deals with structured data
- ☐ Text / sentence planning

Overview of linguistics

❖ The scientific study of language and its structure, including the study of grammar, syntax and phonetics(the study and classification of speech and sound).

❖ Means familiarity with the basics of grammar is certainty important

❖ We must understood how words & sentences are combined to produce meaningful word string.

❖ In NL , the sentence is basic language element .

Sentence :

- Made up of words which express a complete thought .
- To express a complete thought you must have Subject & Predicate

- ✓ The subject is what sentence is about
- ✓ predicate says something about the subject.

- The way a sentence is used determine its mood , declarative , imperative ,interrogative, or exclamatory.
- A word function in a sentence as a part of speech.

- Pos are nouns, pronouns, verbs, adjectives, adverbs, prepositions, conjunction & interjections.

- A language can be consider set of strings of finite or infinite length.

- Well formed sentences are constructed using set of rules called grammar .

- A grammar G is a formal specification of sentence structures that are allowable in language .