Natural Language Processing (NLP): An Overview

Outline

- Objectives
- History of NLP
- Generic NLP System
- Levels of NLP
- Knowledge in Language Processing
- Ambiguity in Natural Language
- Stages in NLP
- Challenges in NLP
- Applications of NLP

Objective

➢ Understand the history of Natural Language Processing   (NLP)
➢ Explore the components of a generic NLP system
➢ Examine the levels and stages of NLP
➢ Discuss knowledge in language processing
➢ Analyze ambiguity in natural language
➢ Identify challenges and applications of NLP

History of NLP

➢ 1950s: Emergence of Machine Translation (MT) for translating languages like Russian to English

➢ 1960s: Introduction of AI concepts revitalized NLP; development of LUNAR system for lunar rock analysis

➢ 1980s: Focus on computational grammar and reasoning for meaning

➢ 1990s: Growth in grammars, tools, and probabilistic models

➢ 2000s-Present: Advances in Machine Learning and Deep Neural Networks for text classification, QA, and sentiment analysis
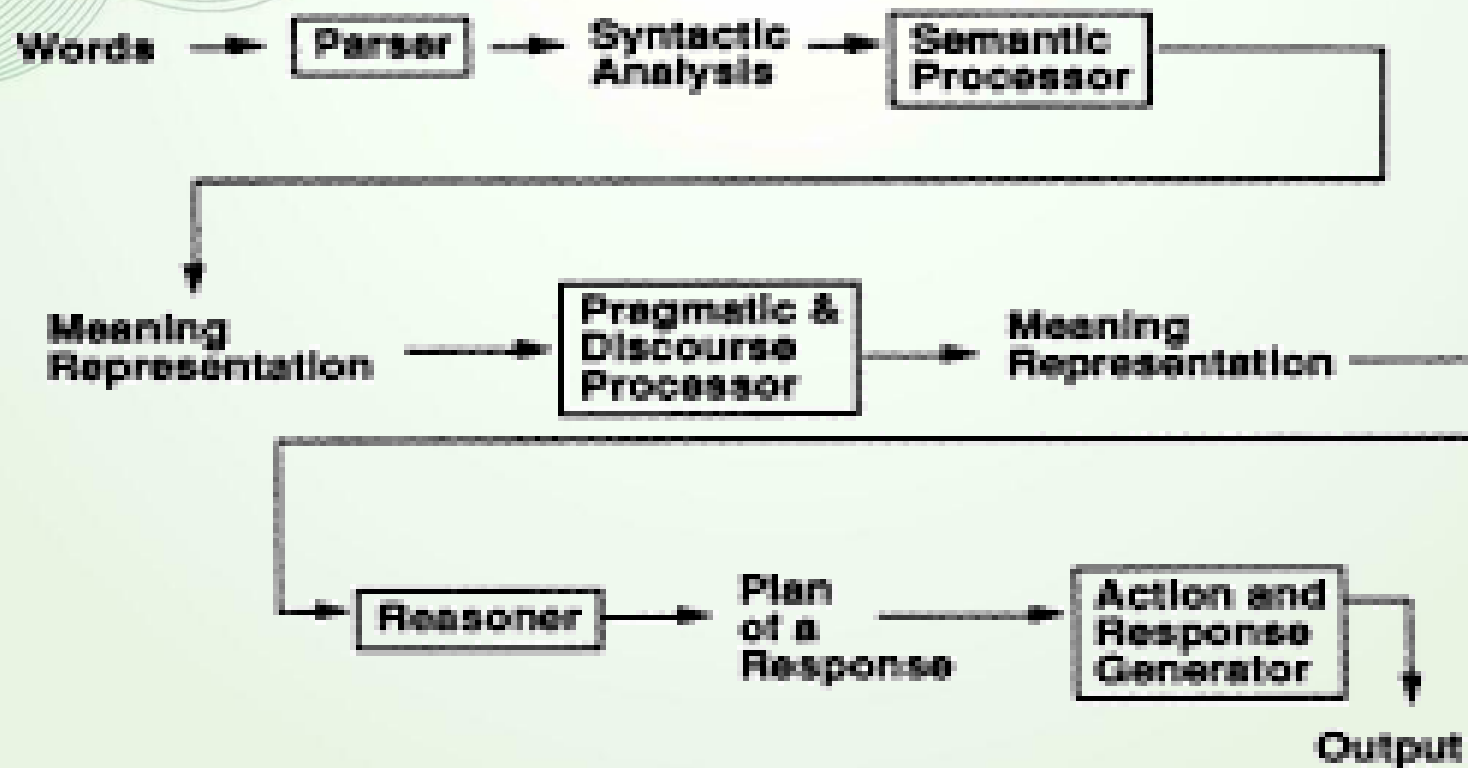
Generic NLP System

➢ Inputs: Text paragraphs, typed commands, or speech

➢ Preprocessing: Dictionary lookup, morphological analysis,lexical substitutions, part-of-speech tagging

➢ Outputs: Database answers, commands, spoken responses, or semantic analysis.

GENERIC NLP SYSTEM
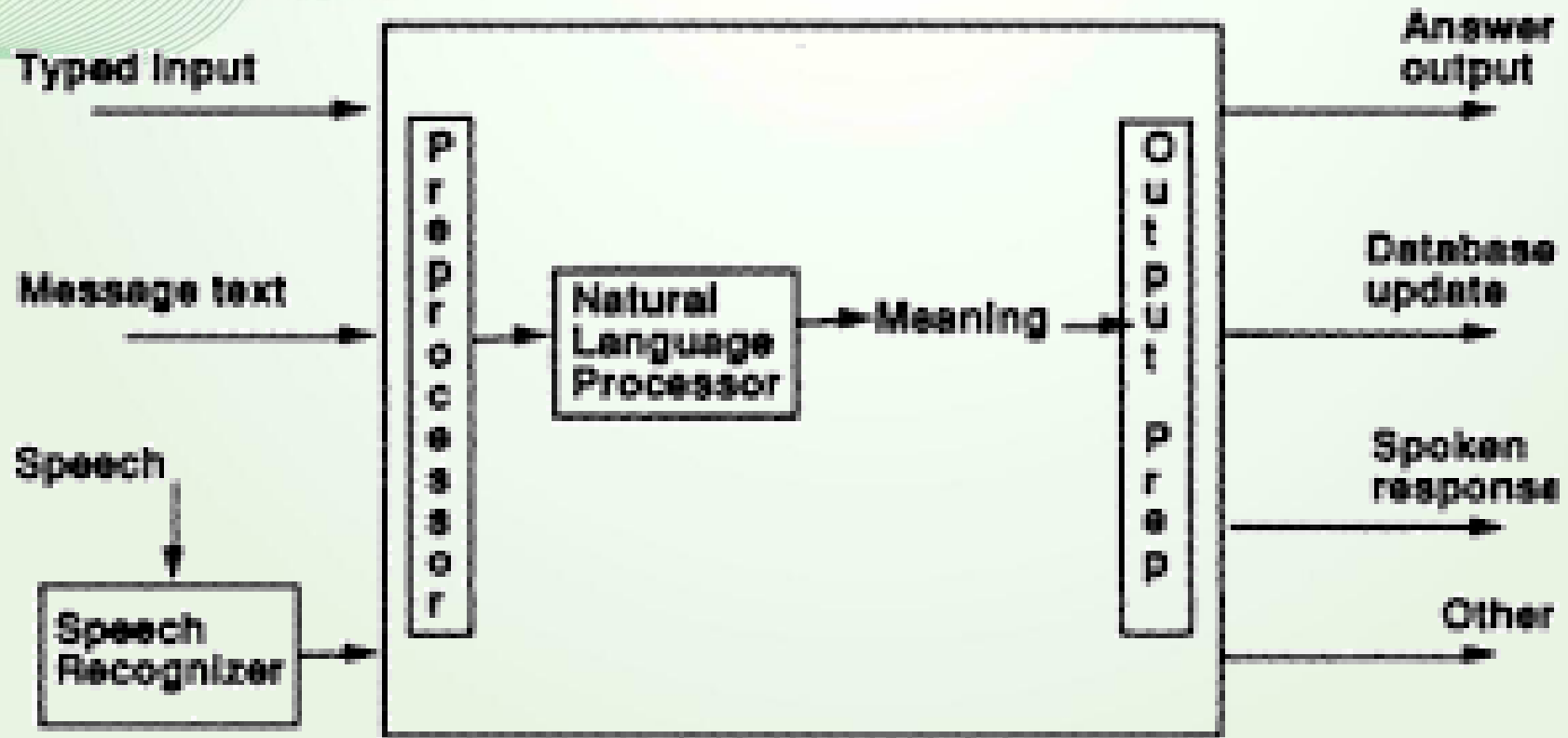
Generic NLP system

Pipeline view of the components of a generic NLP system

Components of NLP

Natural Language Understanding (NLU)
- Analyzes spoken/typed sentences to determine meaning
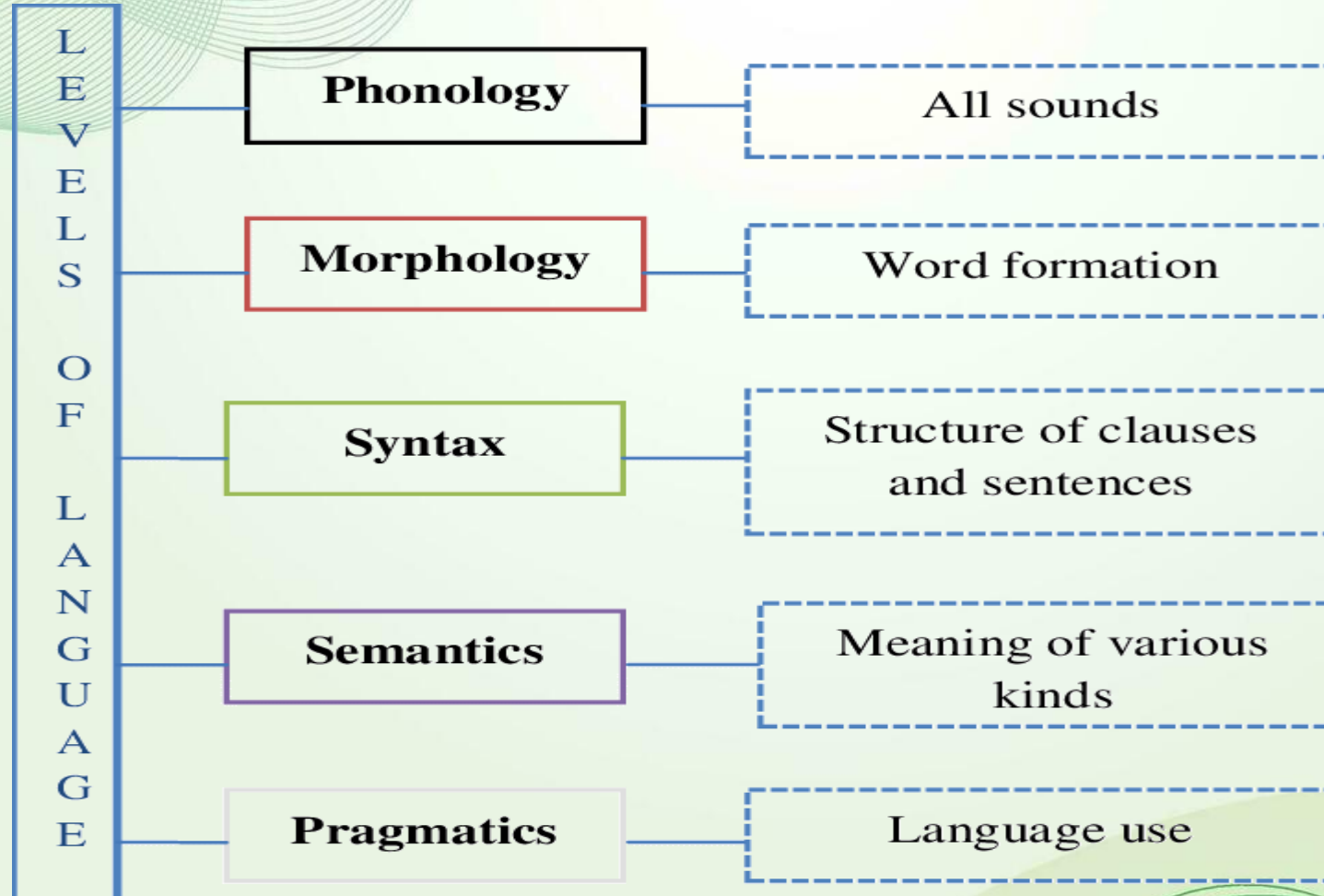- Involves morphological, syntactic, semantic, and discourse analysis Natural Language Generation (NLG
- Converts formal representations into natural language
- Involves deep planning and syntactic generation

Levels of NLP

1.Phonology: Interpretation of speech sounds
2.Morphology: Word construction from morphemes (e.g., truth+ful+ness)
3.Syntax: Sentence structure and word roles (e.g., "The dog ate my homework")
4.Semantics: Meaning of words and sentences
5.Pragmatics: Contextual use of sentences
6.Discourse: Inter-sentential relationships
7.Reasoning: Answering queries using database reasoning

LEVELS OF NLP

LEVELS OF LANGUAGE		
	Phonology	All sounds
	Morphology	Word formation
	Syntax	Structure of clauses and sentences
	Semantics	Meaning of various kinds
	Pragmatics	Language use

Knowledge in Language Processing

➤ Phonetic/Phonological: Sound patterns and syllable formation

➤ Morphological: Word formation from morphemes

➤ Syntactic: Sentence and phrase structure

➤ Semantic: Context-independent meaning

➤ Pragmatic: Contextual meaning in specific situations

➤ Discourse: Inter-sentential links and pronoun resolution

➤ World: General knowledge shared by speakers

Ambiguity in Natural Language

➤ Lexical Ambiguity: Multiple meanings of a word (e.g., "silver" as adjective, noun, or verb)

➤ Syntactic Ambiguity: Multiple parse structures (e.g., "He lifted the beetle with red cap")

➤ Semantic Ambiguity: Multiple sentence interpretations (e.g., "Seema loves mother and Sriya does too")

➤ Anaphoric Ambiguity: Ambiguous pronoun references (e.g., "It" in "The horse ran up the hill. It was very steep.")

➤ Pragmatic Ambiguity: Context-dependent meanings (e.g., "I love you too"0

Stages in NLP

1. Lexical Analysis: Dividing text into words, sentences, and paragraphs
2. Syntactic Analysis: Grammar and word order analysis
3. Semantic Analysis: Assigning meanings to syntactic structures
4. Discourse Integration: Context from preceding and succeeding sentences
5. Pragmatic Analysis: Interpreting intended meaning in context

5 Steps in Natural Language Processing

1. Lexical Analysis

2. Syntactic Analysis

3. Semantic Analysis

4. Discourse Integration

5. Pragmatic Analysis

Challenges in NLP

➢ Language Differences: Variations in vocabulary, phrasing, and inflection

➢ Training Data: Quality and quantity of data impact performance

➢ Development Time: Extensive time required for training

➢ Phrasing Ambiguities: Difficulty in parsing ambiguous statements

➢ Misspellings: Challenges in detecting and correcting errors

➢ Words with Multiple Meanings: Context-based disambiguation

➢ Phrases with Multiple Intentions: Recognizing varied user intents

➢ False Positives: Recognizing system limitations

➢ Conversation Flow: Maintaining interactive dialogue

Applications of NLP

➢ Machine Translation: Converting text/speech between languages (e.g., Google Translator)

➢ Information Retrieval: Searching relevant documents (e.g., Google, Yahoo)

➢ Text Categorization: Sorting documents into categories (e.g., spam filtering)

➢ Information Extraction: Extracting structured data from text

➢ Grammar Correction: Spelling and grammar checks in word processors

➢ Sentiment Analysis: Analyzing emotional tone in text

➢ Question-Answering Systems: Providing direct answers to queries

➢ Spam Detection: Identifying unwanted emails

➢ Chatbots: Customer service automation.