



# Natural Language Processing with Speech Recognition

*Author Heena Jand*

*Assistant professor in Computer Science department*

*Guru Nanak national College, Doraha, Ludhiana, Punjab*

*E-mail: [heenajand21@gmail.com](mailto:heenajand21@gmail.com)*

## **Abstract:**

Natural Language Processing (NLP) is the branch of science which is directly associated to human language which we called as natural language. Language meant for communicating about the world. By studying language, we can come to understand more about the world. This paper describes the classification of NLP, its steps, and application of NLP such as speech recognition. Speech recognition is the vast area of research. NLP deals with more knowledge-related aspects thus requiring the support of learning capabilities by the processors of text. NLP is used to analyze text, allowing machines to understand how humans speak. This human-computer interaction enables real-world applications like automatic text summarization, sentiment analysis, topic extraction, named entity recognition, parts-of-speech tagging, relationship extraction, stemming, etc. NLP is commonly used for text mining, machine translation, and automated question answering.

**Keywords:** NLP, NLU, NLG, speech recognition.

## **1. Introduction**

Natural Language Processing (NLP) is refers to Artificial Intelligence and Linguistics methods of communicating with an intelligent systems using words or statements written in natural language which is understandable by humans like English. Natural Language Processing (NLP) was formulated to built a software that understand the natural language so that user can have natural conversation with the system instead of using any programming language like java or C. Natural Language Processing (NLP) ease the user's work and to satisfy the wish of users to communicate with computer system in natural language. With NLP, a computer is able to listen to a natural language being spoken by a person, understand the meaning of it, and then if needed, respond to it by generating natural language to communicate back to the person. It deals with knowledge-related aspects thus requiring the support of learning capabilities by the processors of text.

NLP is considered as a hard problem in computer science. To understand the human language is to understand not only the words or phrases but the concepts and how they are linked with each other to create meaningful information. NLP algorithms are typically based on machine learning algorithms. Instead of hand-coding large sets of rules, NLP can rely on machine learning to automatically learn these rules by analyzing a set of examples (i.e. a large corpus, like a book, down to a collection of sentences), and making a statical inference. In general, the more data analyzed, the more accurate the model will be.

## **2. Components of NLP**



A language can be defined as a set of rules or set of symbols. Symbols are combined and used to broadcast information. Every language uses its grammar to make sentences similarly NLP considers the hierarchical structure of language: several words make a phrase, several phrases make a sentence and, ultimately, sentences convey ideas. By analyzing language for its meaning, NLP systems have long filled useful roles, such as correcting grammar, converting speech to text and automatically translating between languages.

NLP basically divided into two parts: *Natural Language Understanding* and *Natural Language Generation*.

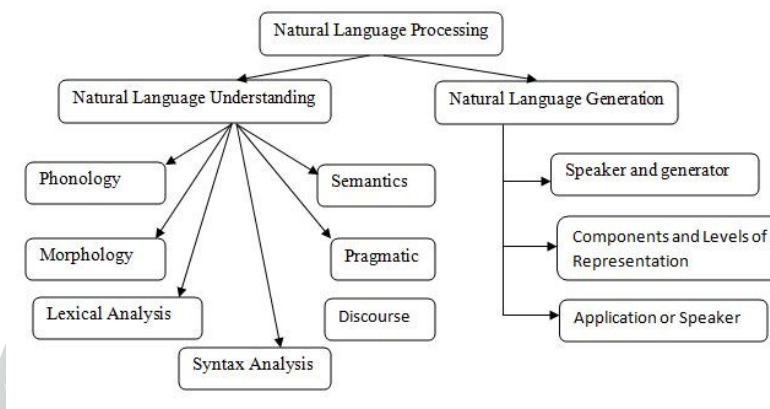


Fig 1: Classification of NLP

## 2.1 Natural Language Understanding

- **Phonology:** Phonology is the study of sound systematically. Here the word ‘phono’ means sound or voice and logy refers to word or speech. It uses semantic use of sound to encode meaning of human language.
- **Morphology:** It is a study of construction of words from primitive meaningful units.
- **Lexical Analysis:** It involves identifying and analyzing the structure of words. It interprets the meaning of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is divides the whole chunk of text into paragraphs, sentences and words.
- **Syntax:** It involves analysis of words in the sentences for grammar and arranging words in a manner that shows the relationship among the words.
- **Semantics:** It is concerned with the meaning of words and how to combine words into meaningful phrases and sentences.
- **Discourse:** The meaning of any sentence depends upon the meaning of the sentence just before it. So that discourse integration manages the connectivity of sentences.
- **Pragmatics:** It deals with using and understanding sentences in different situations and how the interpretation of the sentences is affected.

## 2.2 Natural Language Generation



- **Speaker and generator:** Speaker and generator are used to generate text in relevant situations.
- **Components and Levels of Representation:** The process of generating language involves following tasks:
  - Content selection:* Information should be selected and included in the set.
  - Textual Organization:* The information must be textually organized according the grammar, it must be ordered both sequentially and in terms of linguistic relations like modifications.
  - Linguistic Resources:* To support the information's realization, linguistic resources must be chosen. In the end these resources will come down to choices of particular words, idioms, syntactic constructs etc.
  - Realization:* The selected and organized resources must be realized as an actual text or voice output.
- **Application or Speaker:** Here speaker just initiate the process of generation and store history and structure of the content.

### 3. Application of NLP

NLP is applied in various fields like machine translation, spam filters, text categorization, algorithm trading, summarization, answering questions etc.

- |                               |                                 |
|-------------------------------|---------------------------------|
| • <i>Machine translation:</i> | • <i>Medicine</i>               |
| • <i>Spam filters</i>         | • <i>Dialogue system</i>        |
| • <i>Text categorization</i>  | • <i>Information extraction</i> |
| • <i>Summarization</i>        | • <i>Speech recognition</i>     |
| • <i>Answering questions</i>  | • <i>Voice recognition etc</i>  |
| • <i>Algorithmic trading</i>  |                                 |

### 4. Speech Recognition

Today, science is dealing directly to the processing of natural language like human language which makes it different from other activities in the fields. The research in this field has been started many years ago but at that time there was lack of resources but now a days, it expands its field for research in speech processing.

The main goal of speech recognition is to develop a system which take spoken language as an input and respond in an appropriate way depending on the meaning of the input. It is a process of taking words, phrases, sentences spoken by humans as an input and convert it into electronic signals then transformed these signals into coding patterns. It is a complex and dynamic task which finds the sequence of characters which forms a word. It is a complicated process because different people speaks differently so phoneme and the transition varies from person to person.

Here are some basics of speech recognition.

*Utterance:* It is the vocalization or speaking of a word or words that represents a single meaning to the computer. It can be a single word, few words or sentence/s.

*Speaker dependence:* This type of system is dependent on particular speaker only.



**Vocabularies:** Vocabularies (or dictionaries) are lists of words or utterances that can be recognized by the SR system.

**Accuract:** It is the ability of the system to recognize the word and examine its meaning accurately.

**Training:** A system is trained if it has a speaker in repeat standards or common phrases then comparison algorithms are adjusted to match with that particular speaker. In this training a recognizer usually improve its accuracy.

#### 4.1 Types of Speech Recognition

**Isolated words:** Isolated word recognizers usually require each utterance to have quiet (lack of an audio signal) on BOTH sides of the sample window. It doesn't mean that it accepts single words, but does require a single utterance at a time. Often, these systems have "Listen/Not-Listen" states, where they require the speaker to wait between utterances (usually doing processing during the pauses). Isolated Utterance might be a better name for this class.

**Connected Words:** Connect word systems are similar to Isolated words, but allow separate utterances to be 'run-together' with a minimal pause between them.

**Continuous Speech:** Continuous recognition is the next step. Recognizers with continuous speech capabilities are some of the most difficult to create because they must utilize special methods to determine utterance boundaries. Continuous speech recognizers allow users to speak almost naturally, while the computer determines the content. Basically, it's computer dictation.

**Spontaneous Speech:** There appears to be a variety of definitions for what spontaneous speech actually is. At a basic level, it can be thought of as speech that is natural sounding and not rehearsed. An ASR (automatic speech recognition) system with spontaneous speech ability should be able to handle a variety of natural speech features such as words being run together, "ums" and "ahs", and even slight stutters.

**Voice Verification/Identification:** Some ASR systems have the ability to identify specific users. This document doesn't cover verification or security systems.

#### 4.2 Challenges in Speech Recognition

- Performance degrades due to noise from outer sources.
- Lack of fault tolerant capacity.
- When resource is not ready and user start to speak which leads to lack of synchronization.
- Accuracy and reliability affected with unwanted inputs.

#### 4.3 Uses and Application

**Dictation:** Dictation is the most common use for ASR systems today. This includes medical transcriptions, legal and business dictation, as well as general word processing. In some cases special vocabularies are used to increase the accuracy of the system.

**Command and Control:** ASR systems that are designed to perform functions and actions on the system are defined as Command and Control systems. Utterances like "Open Netscape" and "Start a new xterm" will do just that.





**Telephony:** Some PBX/Voice Mail systems allow callers to speak commands instead of pressing buttons to send specific tones.

**Wearables:** Because inputs are limited for wearable devices, speaking is a natural possibility.

**Medical/Disabilities:** Many people have difficulty typing due to physical limitations such as repetitive strain injuries (RSI), muscular dystrophy, and many others. For example, people with difficulty hearing could use a system connected to their telephone to convert the caller's speech to text.

## 5. Conclusion

This paper attempts to review the Natural language processing, its levels and uses in speech recognition. The research in the field of NLP and speech recognition is a continue process from last 50 years. The speech recognition will consider the best when the error rate during the recognition can minimize. This paper defined various types of speech recognition used in various fields. In future there is need to remove the errors and make this field of artificial intelligence more and more useful.

## 6. References

- [1.] Green Jr, B. F., Wolf, A. K., Chomsky, C., & Laughery, K. (1961, May). Baseball: an automatic question-answerer. In Papers presented at the May 9-11, 1961, western joint IREAIIEE-ACM computer conference (pp. 219-224). ACM.
- [2.] Kamp, H., & Reyle, U. (1993). Tense and Aspect. In From Discourse to Logic (pp. 483-689). Springer Netherlands.
- [3.] F. Ivon, "Self-learning techniques for grapheme-to-phoneme conversion," in Onomastica Research Colloquium, (London, UK), December 1994.
- [4.] L. R. Bahl, P. F. Brown, P. V. de Souza, and R. L. Mercer, "A tree-based statistical language model for natural language speech recognition," in Acoustics, Speech and Signal Processing, *IEEE Transactions on*, vol. 37, Issue 7, (Yorktown Heights, NY, USA), pp. 1001–1008, July 1989.
- [5.] Makhoul, J. & Schwartz, R. (1995) *Proc. Natl. Acad. Sci. USA* 92, 9956-9963.
- [6.] Bates, M. (1995) *Proc. Natl. Acad. Sci. USA* 92, 9977-9982.

## 7. Bibliography

Name: Heena jand

Qualification: MSC (IT)

Assistant professor in computer science department.

Guru Nanak National College Doraha, Ludhiana Punjab.

Email- [heenajand21@gmail.com](mailto:heenajand21@gmail.com)

