

Computational Linguistics (CL3.101)

Summer 2021, IIIT Hyderabad

31 May, Monday (Lecture 3)

Taught by Prof. Radhika Mamidi

Language Processing

The transition between the various stages of language processing is bidirectional, to enable re-analysis.

Machine Translation

A simple architecture of a machine translation system is as follows:

- Pre-processors: font converters, identifying idioms/collocations/phrasal verbs etc.
- POS tagger
- Morphological analysis
- Chunker
- Parser

These steps constitute the Source Language Analysis phase. This order, however, is not constant across languages; for instance, in Dravidian languages, POS tagging happens after morphological analysis.

- Transfer rules
- Preposition mapping rules
- Word sense disambiguation
- Bilingual dictionary lookup

These are the Source Language-Target Language Mapping phase. It is not an exhaustive list; there are more steps.

- Morphological generator
- Sentence generator
- Font converters

These final steps make up the Target Language Generation phase.