# Introduction to Computational Linguistics

Hyunjoo Cho

Fachschaft General & Computational Linguistics
**University of Tübingen**

WS 2024/25
Pre-course

## Definition

Scientific study of language from a computational perspective and build NLP tools

## Goals

1. Modelling and simulating human language to make it understandable for computers
2. Improving applications integrating linguistic data structures.

Adequate representation of the properties of human language into a formal system is needed.

1. Human language is prone to ambiguities and variations.
2. Computer are restricted to more formal systems and are vulnerable to any ambiguous or spontaneous changes.

## Natural vs. Formal Languages

## Natural Language

1. Languages that we speak.
2. Not designed by people.
3. Evolved naturally as they pass from generations to generations.

## Formal Language

1. Artificial languages.
2. Designed by people for specific purpose.
3. Main examples: programming languages.

## Related fields

1. Linguistics
2. Computer Science
3. Cognitive Science (Psycholinguistics)
4. Artificial Intelligence
5. Mathematics (Logic)
6. Philosophy
7. Neuroscience
8. NLP (often used as synonym)

## What is Natural Language Processing?

Mapping the given input (natural language) into useful representation & analyzing different aspects of the input

Examples:

1. POS tagging
2. Lemmatisation
3. Dependency parsing

## Main fields of application

Machine translation

1. Google translate
2. DeepL
3. Linguee

## Main fields of application

Text editors/spell checkers

1. Notepad
2. Grammarly
3. Wordtune

## Main fields of application

Chatbots

1. ChatGPT
2. Customer support systems

## Main fields of application

Speech recognition systems/text-to-speech synthesisers

1. Speech services
2. Google translate
3. 'Read aloud' options in browsers

## Machine translation

1. rules based approach
2. interlingual approach
3. dictionary-based approach
4. statistical approach
5. deep learning based approach (neural machine translation)

## Chatbots

1. heavily related to artificial intelligence, machine learning and natural language processing
2. require a large amount of conversational data to train
3. input/output database is usually fixed

## Career options

1. Natural Language Processing Engineer
2. Computational Linguist
3. Data Scientist
4. AI Engineer
5. Programmer

## Suggested literature and sources

For Introduction to CL


1. Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall, Upper Saddle River, NJ, 2nd edition edition, 2009
2. Ralph Grishman. Computational linguistics: an introduction. Cambridge University Press, 1986.
3. Turing, Alan (1950), "Computing Machinery and Intelligence", Mind, LIX (236).
4. John R. Searle. Minds, brains, and programs. Behavioral and Brain Sciences 3, 1980.
5. Dickinson, Markus, et al. Language and Computers. Wiley, 2012.

## Suggested literature and sources

For Logic


1. Magnus, P. D. FORALLX: An introduction to formal logic., 2017.
2. L.T.F. Gamut. Logic, Language, and Meaning, Volume 1: Introduction to Logic.

## Suggested literature and sources

For Java/DSA I

1. Savitch, Walter. Java: An Introduction to Problem Solving and Programming. Pearson, 2010.

Questions?