

Computational Linguistics

Parameswari Krishnamurthy

IIIT Hyderabad

parameshkrishnaa@gmail.com

January 3, 2025

Linguistics

Linguistics is a field of science seeks to answer the following questions:

- what exactly do we know when we know a language?
- how is this knowledge acquired? and
- how is such knowledge used?

Linguistics

Linguistics is a field of science seeks to answer the following questions:

- what exactly do we know when we know a language?
- how is this knowledge acquired? and
- how is such knowledge used?

Further it also investigates,

- which characteristics are unique to a language?
- which are universal (found in all the world's languages)?
- how individual languages differ, how and why they change, "die" or "are born"?

Linguistics

Linguistics is a field of science seeks to answer the following questions:

- what exactly do we know when we know a language?
- how is this knowledge acquired? and
- how is such knowledge used?

Further it also investigates,

- which characteristics are unique to a language?
- which are universal (found in all the world's languages)?
- how individual languages differ, how and why they change, "die" or "are born"?

Like all sciences, linguistics is divided into theoretical and applied.

Theoretical Linguistics

- phonetics (articulation and perception of sounds)

Theoretical Linguistics

- phonetics (articulation and perception of sounds)
- phonology (sound system of a particular language)

Theoretical Linguistics

- phonetics (articulation and perception of sounds)
- phonology (sound system of a particular language)
- morphology (inflection and word formation)

Theoretical Linguistics

- phonetics (articulation and perception of sounds)
- phonology (sound system of a particular language)
- morphology (inflection and word formation)
- syntax (phrases and sentences)

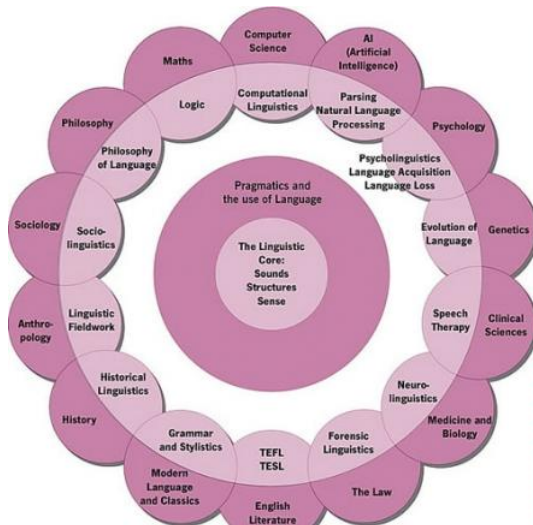
Theoretical Linguistics

- phonetics (articulation and perception of sounds)
- phonology (sound system of a particular language)
- morphology (inflection and word formation)
- syntax (phrases and sentences)
- semantics (the meaning of linguistic expressions)

Theoretical Linguistics

- phonetics (articulation and perception of sounds)
- phonology (sound system of a particular language)
- morphology (inflection and word formation)
- syntax (phrases and sentences)
- semantics (the meaning of linguistic expressions)
- pragmatics (language use)

Applied Linguistics



Computational Linguistics

Common names:

- Computational Linguistics (CL)
- Natural Language Processing (NLP)
- Language Engineering
- Human Language Technology (HLT)

Computational Linguistics

Why Computational Linguistics (CL) rather than Natural Language Processing (NLP)?

Computational Linguistics

- Computers dealing with language
- Modeling what people do
- Rule-Based Modelling
- Major role by Linguists

Natural Language Processing

- Applications on the computer side
- Majorly by statistical and machine learning based Modelling
- Major role by Computer Engineers

Computational Linguistics

- computational linguistics (broad sense): interdisciplinary research field (between linguistics and computer science) which develops concrete algorithms for natural language processing (machine translation, machine speech recognition ...)
- computational linguistics (narrow sense): discipline in modern linguistics which develops, implements and investigates computational models of human language.
- Computational linguistics is the study of computer systems for understanding and generating natural language. (Ralph Grishman, Computational Linguistics: An Introduction, Cambridge University Press 1986.)

- It deals with formal theories about the linguistic knowledge that a human needs for generating and understanding language.
- Computational linguists develop formal models simulating aspects of the human language faculty and implement them as computer programmes.

Computational linguists are interested in providing computational models of various kinds of linguistic phenomena. These models may be knowledge-based (hand-crafted) or data-driven (statistical or empirical).

- Computational linguistics is the scientific study of language from a computational perspective.
- field predates artificial intelligence
- 1950s in US to translate Russian scientific texts into English.
- tries to model human mind
- synthesizing and analysing of language.
- issue:
 - Linguistic Knowledge
 - World Knowledge
 - Domain Knowledge
 - Context Knowledge
 - Cultural Knowledge

History

- 1950: Machine Translation
- 1960: Chomsky
- 1964: ALPAC Report
- 1966: ELIZA
- 1980: knowledge-based CL
- 1990: statistical / machine learning approaches in CL

Computational Linguistics

