

Applied Text Mining in Python

Basic Natural Language Processing

What is Natural Language?

- **Language used for everyday communication by humans**
 - English
 - 中文
 - हिन्दी
 - русский язык
 - español

c u 2nite



compared to the artificial computer languages

What is Natural Language Processing?

- Any computation, manipulation of natural language
- Natural languages evolve
 - new words get added
 - old words lose popularity
 - meanings of words change
 - language rules themselves may change

selfie

thou

learn

position of verbs in
sentences!

NLP Tasks: A Broad Spectrum

- Counting words, counting frequency of words
- Finding sentence boundaries
- Part of speech tagging
- Parsing the sentence structure
- Identifying semantic roles
- Identifying entities in a sentence
- Finding which pronoun refers to which entity

and much more ...