

Course: **Computational Linguistics** (exam code: 2000886)  
Professor: **Cristiano Chesi** ([c.chesi AT unisi.it](mailto:c.chesi@unisi.it), 0577-234762)  
Student hours: **Friday, 11:00** (e-mail before coming), room 327, CISCL, S. Niccolò  
Start date: **Thursday, 5 October**

Language and Mind:  
Linguistics and Cognitive Studies (LM39)  
2018-19

## COMPUTATIONAL LINGUISTICS SYLLABUS

### Program

The topic of the course is language and computation: on the one hand we will see how linguistic intuitions can be formalized so as to implement computer programs. On the other we will dig into the cognitive plausibility of these computational models.

We will also discuss how to store in an appropriate format linguistic information (corpora) and how to query linguistic databases.

This course will provide you with theoretical and technical background to evaluate the advantages (and the limits) of a quantitative and qualitative approach to the language study.

The main topics that will be analyzed are the following ones:

- annotated corpora and lexical information;
- morpho-syntactic parsers;
- syntactic and semantic tagging;
- machine translation systems;
- (time permitting) language in the brain;

No programming experience is required. A strong interest/curiosity in (Generative) Linguistics, Artificial Intelligence and Cognitive Psychology is indeed requested.

At the end of the course, the student (possibly in **group of maximum 3 people**) should present a **short project** (20 minutes max) describing problems and advantages of an invented (artificial) device communicating in natural language with human in a precise scenario (to be defined by the student/group).

### Evaluation

Class participation (20% of final grade) – please, interact during the lectures (read stuff before the lecture, ask questions, discuss ideas...)  
Project presentation (40% of final grade) – you should deliver a short presentation during the last lecture: this should be a little project presentation that is a tentative investigation of some of the topic we discussed in class (students are encouraged to prepare the project in groups of a maximum of 3 people)  
Oral exam (40% of final grade) on course topics (see References)

### References

- Lecture notes and e-learning: <http://elearning.unisi.it/moodle/course/view.php?id=2085>
- Jurafsky, D. & Martin, J. H. (2008) *Speech and Language Processing*. Prentice-Hall. (2nd edition)  
<http://www.cs.colorado.edu/~martin/slp.html>  
chapters: 1, 2, 3, 4, 5, 12, 13, (14), 15, (16, 17, 18), 19, (20) (optional chapters)
- Other readings will be presented during the lectures
- software (home-made software will be discussed and delivered during the classes):  
Antconc <http://www.laurenceanthony.net/software.html>  
TLearn <http://crl.ucsd.edu/innate/tlearn.html>

## Tentative program (there might be changes!)

ALWAYS double check the web page <http://elearning.unisi.it/moodle/course/view.php?id=2085> for announcements and materials!


**Lecture 1** - Thursday 5 October 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Course presentation, Intro to Natural Language Processing (NLP) and Linguistic Computation.


**Lecture 2** - Friday 6 October 2018, hour 9:00 - 11:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Intro to corpus linguistics


**Lecture 3** - Thursday, 18 October 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Formal Grammars and Chomsky's Hierarchy


**Lecture 4** - Friday, 19 October 2018, hour 9:00 - 11:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): CHILDES corpus


**Lecture 5** - Thursday, 25 October 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Lexicon, morphology and... spell checkers


**Lecture 6** - Friday, 26 October 2018, hour 9:00 - 11:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): Corpus creation and morpho-syntactic annotation


**Lecture 7** - Thursday, 15 November 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Syntactic Paring (Top-down Vs Bottom-up parsing, left corner, Earley algorithm)

**Lecture 8** - Friday, 16 November 2018, hour 9:00 - 11:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): Grammar writing and parsing


**Lecture 9** - Thursday, 29 November 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Advanced parsing: minimalist grammars and linguistic performance


**Lecture 10** - Friday, 30 November 2018, hour 9:00 - 11:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Knowledge representation, Bag-of-Words and disambiguation


**Lecture 11** - Thursday, 13 December 2018, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Sub-symbolic approach to language: neural networks


**Lecture 12** - Friday, 14 December 2018, hour 9:00 - 10:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): word sense disambiguation (using WORDNET)

**Lecture 13 (optional)** - Wednesday, 9 January 2019, hour 16:00 - 18:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Language in the Brain

**Lecture 14** - Thursday, 10 January 2019, hour 14:00 - 16:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): Simple Recurrent Networks (using T-Learn)


**Lecture 15** - Friday, 11 January 2019, hour 9:00 - 11:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Introduction to Machine Learning (and to Translation)

**Lecture 16 (optional)** - Wednesday, 23 January 2019, hour 16:00 - 18:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): machine learning software (using WEKA)

**Lecture 17** - Thursday, 24 January 2019, hour 14:00 - 16:00, Aula 20 (San Niccolò, Via Roma 56, Floor 0)

 Lab time (please bring your own laptop): course summary

**Lecture 18** - Friday, 25 January 2019, hour 11:00 - 13:00, Aula 15 (San Niccolò, Via Roma 56, Floor 0)

 Your project discussion