

Faculty of Arts and Letters



CLIG2190 Introduction to the quantitative analysis of linguistic data

[30h] 4 credits

This course is taught in the 1st semester

Teacher(s): Cédric Fairon
Language: French
Level: Second cycle

Aims

At the end of the course, the student will be able to:

- build a corpus of written or oral data on which to carry out a linguistic study (at a lexical, phonetic, semantic, or discourse level);
- compare and use different corpus processing softwares;
- do a quantitative study of a corpus;
- do a qualitative study of a corpus.

Main themes

The objective of this course is to introduce the students to the theory and methodology of corpus linguistics. Corpus linguistics is an applied field of linguistics, which aims at providing empirical foundations to theoretical research. It has continuously gained more importance in the last 10 years with the development of computer capacities and the availability of electronic textual data (cd-rom, Internet, etc.).

Content and teaching methods

The course is organised in three parts (mixing theory and practice) as follows:

1. constitution of oral and written corpora for linguistic analysis :
 - a. typology (finite corpus vs. dynamic corpus; mono vs. multilingual corpora, etc.);
 - b. corpora of written data: existing resources; manual constitution and automatic retrieval of texts; encodings and annotations; shared publication formats (TEI) and meta-data;
 - c. oral corpora: existing resources; recording and digitalisation; transcription and annotation; sociolinguistic profiles.
2. principles and methods of quantitative analysis: lexical statistics, (concordances, collocation extraction, graphical representation tools); introduction to statistical software.
3. principles and method of qualitative analysis: analysis rules; coding criteria, (parameters and categories).

The course is interactive: students are asked to prepare each course by either readings (a portfolio is provided on the Icampus web site) or applied exercises (corpus development, searches in corpora, software experimentation, etc.). Each session starts with question/ answers sessions or correction and discussion of exercises.

Other information (prerequisite, evaluation (assessment methods), course materials recommended readings, ...)

Prerequisite: one course of general introduction to linguistics.

Evaluation: Students are evaluated on their regular participation. Students have to carry out a personal project for the final evaluation.

Other credits in programs

LING21MA Première année du master en linguistique, à finalité (4 credits)
 approfondie en linguistique générale et appliquée

LING21MS Première année du master en linguistique, à finalité spécialisée (4 credits)
 en ingénierie linguistique