lgerm2829
2019

English Linguistics : Corpus Linguistics

In view of the health context linked to the spread of the coronavirus, the methods of organisation and evaluation of the learning units could be adapted in different situations; these possible new methods have been - or will be - communicated by the teachers to the students.

5 credits	22.5 h	Q1

Teacher(s)	Gilquin Gaëtanelle ;
Language :	English
Place of the course	Louvain-la-Neuve
Main themes	The course introduces the main tenets of corpus linguistics and the methods and techniques used to work with large collections of spoken or written electronic data. It covers the following topics: corpus design: • data collection, archiving and markup. • corpus typology: spoken and written corpora; monolingual vs multilingual; native vs learner; diachronic vs synchronic. • major electronic corpora: British National Corpus, International Corpus of English, International Corpus of Learner English, MICASE, Louvain International Database of Spoken English Interlanguage, etc. • corpus annotation (POS-tagging, lemmatization, parsing, semantic tagging, prosodic annotation, error tagging). • automated analysis of lexis, grammar and discourse. Special attention is paid to the links between corpus linguistics and foreign language learning, contrastive and translation studies and natural language processing.
Aims	1 By the end of the course, students are expected to have a solid theoretical background in corpus linguistics and master the main techniques and tools used to analyse spoken and written computerized data. They will be able to read the scientific literature and conduct their own research in the field. - - - - *The contribution of this Teaching Unit to the development and command of the skills and learning outcomes of the programme(s) can be accessed at the end of this sheet, in the section entitled "Programmes/courses offering this Teaching Unit".*
Evaluation methods	**Due to the COVID-19 crisis, the information in this section is particularly likely to change.** During the term: one or several written assignments involving the analysis of corpus data and counting for 20% of the final grade. In case of second exam enrolment, the students who have not obtained at least 10/20 for the assignment(s) will have to do it/them again. In January or September: written exam counting for 80% of the final grade. A WORD OF CAUTION: students who have not handed in their written assignment(s) on time will fail this course overall (7/20 or less if the mean is lower).
Teaching methods	**Due to the COVID-19 crisis, the information in this section is particularly likely to change.** The course relies partly on required readings that the students are expected to do before class and that will lead to (in-class or online) discussions in which the students should be ready to participate fully. Some (in-class or online) hands-on sessions will be organized to familiarize students with text handling software tools.
Content	The course provides a theoretical and practical introduction to corpus linguistics. It presents the main concepts related to corpus linguistics, as well as some of its possible applications in different fields.
Inline resources	Moodle
Faculty or entity in charge	LMOD

Program title	Acronym	Credits	Prerequisite	Aims
Master [120] in Translation	TRAD2M	5		🔍
Master [60] in Modern Languages and Literatures : General	ROGE2M1	5		🔍
Master [120] in Modern Languages and Literatures : German, Dutch and English	GERM2M	5		🔍
Master [60] in Modern Languages and Literatures : German, Dutch and English	GERM2M1	5		🔍
Master [120] in Linguistics	LING2M	5		🔍
Master [120] in Modern Languages and Literatures : General	ROGE2M	5		🔍

Programmes containing this learning unit (UE)