



COURSE DATA

DATA SUBJECT

Code: 44225
Name: Applied corpus linguistics
Cycle: Master's Degree
ECTS Credits: 3
Academic year: 2025-26

STUDY (S)

Degree	Center	Acad. year	Period
2243 - Master's Degree in Advanced English Studies	Facultat de Filologia, Traducció i Comunicació	1	First quarter

SUBJECT-MATTER

Degree	Subject-matter	Character
2243 - Master's Degree in Advanced English Studies	Applied corpus linguistics	COMPULSORY

COORDINATION

GREGORI SIGNES CARMEN

FUSTER MARQUEZ MIGUEL

SUMMARY

This course is designed as an introduction to the various techniques employed in accessing and compiling a corpus. The course offers an overview of the recent advances and methods in corpus linguistics and their contribution to the current research in various fields, with specific focus on applied linguistics. Its primary objective is to understand the potential of applying the techniques used in corpus linguistics to the analysis of typologically different texts using mainly public online reference corpora of various kinds. Additionally, the course will instruct on a number of steps and parameters to build one's own corpora, and implementing morphological and semantic annotation. This will enhance the students' understanding of the quantitative and qualitative results obtained in their analysis and how to interpret these results.

PREVIOUS KNOWLEDGE

RELATIONSHIP TO OTHER SUBJECTS OF THE SAME DEGREE

There are no specified enrollment restrictions with other subjects of the curriculum.

OTHER REQUIREMENTS



No previous knowledge in linguistics is required. A solid C1 level of English is highly recommended.

COMPETENCES / LEARNING OUTCOMES

-

Ability to apply theoretical knowledge and skills to conduct critical, rigorous analysis using various tools, including printed and digital formats (ICT), within English Studies.

Ability to develop autonomy in learning, including resource and information search, and to access bibliographic and documentary sources in different areas of English Studies.

Ability to identify fundamental methodologies, theories, topics, and terminological, theoretical, formal, and ideological principles necessary for initiating linguistic or literary research in English Studies.

Ability to make informed judgments using established criteria and personal reflections. Achievement of proficiency in academic and scientific-technical English, both in written and oral forms.

Ability to organize, structure, and develop ideas in various academic documents, such as essays, technical reports, and advanced research projects masters theses within the field of English Studies.

Ability to publicly present from an advanced and more complex perspective than that of undergraduate studies experiences, ideas, or reports within the field of English Studies.

Advanced ability to understand from a more complex perspective than that of undergraduate studies the methodology required for solving problems specific to the field of English Studies.

Advanced ability to understand from a more complex perspective than that of undergraduate studies those concepts, principles, theories, or models related to different areas of English Studies.

Capacidad para interpretar y analizar textos representativos de la pluralidad lingüística y cultural de las sociedades anglófonas contemporáneas con el fin de reflexionar sobre su relevancia, no sólo en referencia a los contextos socio-lingüísticos, históricos, políticos y culturales en los que se inscriben, sino también en relación con la globalización de la cultura, dentro de los Estudios Ingleses.

Students should apply acquired knowledge to solve problems in unfamiliar contexts within their field of study, including multidisciplinary scenarios.

Students should demonstrate self-directed learning skills for continued academic growth.

Students should possess and understand foundational knowledge that enables original thinking and research in the field.

Use acquired knowledge as a basis for originality in the development or application of ideas, often in a research context.

DESCRIPTION OF CONTENTS



1. Introduction to corpus linguistics

Introductory session: presentation of didactic guide (program, methodology and evaluation).
General introduction to Corpus Linguistics and methodologies.
Presentation of some useful interfaces in this field.
Corpus data and corpus design.
Computer tools and documents for the following sessions.

2. Introduction to the use of free corpus programs

Presentation of free concordance programs.
Downloading of documents and conversion to plain text.
Basic techniques in corpus exploration: word lists, concordance, raster, collocations and groupings.

3. Exploring free online reference corpora

The value of reference corpora for your own corpus research.
Exploration of prestigious online public corpora.
Exploration of some of their basic techniques

4. Exploring and using Sketch Engine

Introduction to Sketch Engine for creating your own corpus and exploring built-in corpora.
Automatic annotation systems in Sketch Engine.
English reference corpus in Sketch Engine.
Sketch Engine tools: key terms, keywords, word sketches and engrams.

5. Quantification and interpretation of corpus results

Critical appraisal of the contribution of corpus linguistics in recent research.
Statistical analysis of quantitative data and its significance in correlation with corpus-based qualitative analysis.

Learning outcomes

Know the current research in Corpus Linguistics in both theoretical and applied fields in the English language.

Understand the typology, organization, topics, and use of different corpora for research in the field of Applied Linguistics.

Use various computer tools to obtain and interpret data meaningfully.

Know how to store results in appropriate databases based on solid selection principles.

Know methods for data presentation.



Be able to interpret linguistic data quantitatively using different statistical techniques and software tools.

Be able to relate quantitative and qualitative analysis to the theoretical foundations of current research methods.

Understand the parameters for compiling a relevant research corpus in areas such as English language teaching, translation, English for academic and professional purposes, media, or political discourse.

Promote the handling of data extracted from native speakers and/or learners and possess basic knowledge of corpus tagging and automated processing.

Understand and appreciate, from within the scope of the degree itself, inequalities based on sex and gender in society; integrate the different needs and preferences based on sex and gender in the design of solutions and problem solving.

WORKLOAD

PRESENCIAL ACTIVITIES

Activity	Hours
Theory	24,00
Total hours	24,00

NON PRESENCIAL ACTIVITIES

Activity	Hours
Attendance at other activities	0,00
Individual or group project	16,00
Independent study and work	20,00
Preparation of lessons	2,00
Preparation for assessment activities	10,00
Resolution of case studies	3,00
Total hours	51,00

TEACHING METHODOLOGY

TM1. Theoretical classes.

TM2. Group work.

TM3. Fixed readings (autonomous reading or with class discussion, selecting relevant information, relating concepts, studying, completing worksheets, writing outlines or summaries).

TM4. Design and development of assignments for oral or written presentations.

TM5. Theoretical and practical problem solving. Compilation of corpus for data analysis. Different genres and modes (written, oral, visual, digital and/or multimodal).

EVALUATION



BREAKDOWN OF THE EVALUATION FOR THE FIRST CALL:

Active participation in class: 20%

2. Evaluation of competencies by written test: 30%
3. Evaluation of competencies through written/oral projects in class: 50%

BREAKDOWN OF THE EVALUATION FOR THE SECOND CALL:

1. Assignments made through participation are compensated with a practical test: 20%
2. Evaluation of competencies through a written test: 30%
3. Evaluation of competencies through projects through a practical test. Students must demonstrate their knowledge of corpus methodologies, accounting for 50% of the assessment.

IMPORTANT: Intellectual honesty is a core value in academic communities and essential for the fair assessment of students' work. All assignments submitted for this course must be of **original authorship**. Submissions that involve **fraudulent collaboration** or the use of **generative artificial intelligence** (such as ChatGPT or similar tools) will not be accepted, unless their use is explicitly authorised by the teaching staff and forms part of the course content.

Exams and practical exercises must be free from spelling and grammatical errors. Each spelling, punctuation, or expression mistake will result in a deduction from the final grade, which may lead to a failing mark.

The general grading system will follow the regulations of the Universitat de València on assessment and grading, approved by the Governing Council on May 30, 2017 (ACGUV 108/2017).

REFERENCES

Core



- Baker, Paul (ed) (2009) *Contemporary corpus linguistics*. London-New York: Continuum.
- Cheng, Winnie (2012) *Exploring corpus linguistics: language in action*. London: Routledge.
- Flowerdew, Lynne (2012) *Corpora and Language Education*. Houndmills: Palgrave Macmillan.
- Kawaguchi, Yuji, Makoto Minegishi & Jacques Durand (2009) *Corpus analysis and variation in linguistics*, Amsterdam-Philadelphia: John Benjamins.
- Meunier, Fanny et al. (eds) (2011) *A taste for corpora: in honour of Sylviane Granger*. Amsterdam-Philadelphia: John Benjamins.
- O'Keefe, Anne, Michael McCarthy & Ronald Carter (2007) *From corpus to classroom: language use and language teaching*, Cambridge: C.U.P.
- Renouf, Antoinette & Andrew Kehoe (2009) *Language and Computers Studies in Practical Linguistics*, Vol. 69. *Corpus Linguistics: refinements and reassessments*, Amsterdam-New York: Rodopi.

Supplementary

- Baker, Paul & Tony McEnery (2015) (eds.), *Corpora and discourse studies: Integrating discourse and corpora* (Palgrave Advances in Language and Linguistics). Basingstoke: Palgrave Macmillan.
- Cantos Gómez, Pascual (2013) *Statistical method in language and linguistic research*. Oakville, CT: Equinox.
- Fuster-Márquez, M. & Almela, M. (2017) *Words, Corpus and Back to Words*. *Quaderns de Filologia: Estudis Linguistics* 22. Universitat de València.
- Fuster-Márquez, M. & Gregori-Signes, Carmen (2018) *Persuading Consumers: The Use of Conditional constructions in British Hotel Websites*. *Discourse and Communication* 12(6): 587-607.
- Garside, Roger, Geoffrey Leech & Tony McEnery (eds) (1997) *Corpus Annotation: Linguistic Information from Computer Text Corpora*, London-New York: Routledge.
- McEnery, Tony and Andre Wilson (2001) *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- McEnery, Tony, Richard Xiao & Yukio Tono (2006) *Corpus-based language studies*. London-New York: Routledge.
- Partington, Alan (1998) *Patterns of Meaning*. Philadelphia: John Benjamins Publishing Company.
- Tognini-Bonelli, Elena (ed.) (2001) *Corpus linguistics at work*. Amsterdam-Philadelphia: John Benjamins Publishing Company.
- Stubbs, Michael (2001) *Words and Phrases: Corpus Studies of Lexical Semantics*. Wiley.