



Teaching Guide

Teaching Guide				
Identifying Data				2019/20
Subject (*)	Corpus Linguistics and Computer Science in English		Code	613505109
Study programme	Mestrado Universitario en Estudos Ingleses Avanzados e as súas Aplicacións (2019)			
Descriptors				
Cycle	Period	Year	Type	Credits
Official Master's Degree	2nd four-month period	First	Optional	3
Language	English			
Teaching method	Face-to-face			
Prerequisites				
Department	Letras			
Coordinador	Crespo Garcia, Maria Begoña	E-mail	begona.crespo.garcia@udc.es	
Lecturers	Crespo Garcia, Maria Begoña	E-mail	begona.crespo.garcia@udc.es	
Web	https://www.udc.es/grupos/muste/			
General description	The aim of this course is to introduce the field of corpus linguistics and corpus-based methodology for the linguistic study of English. General issues about the design, collection and analysis of corpus material will be discussed. In practical sessions students will learn how to search and analyse text corpora using software tools. Students will develop the ability to describe and discuss the extent to which language usage varies and can be tracked with reference to genuine examples attested in corpora. With this course students will acquire skills in: ? the principles and methodology of English corpus linguistics; ? the practical use of linguistic corpora and related software; ? the discussion of specialized literature on corpus linguistics; ? the presentation of research results.			

Study programme competences

Code	Study programme competences
A1	E01 ? Familiarity with the main research models in linguistic research.
A2	E02 ? Familiarity with the main resources, tools and methodologies in linguistic research.
A14	E14 ? Familiarity with and application of techniques and methods of quantitative linguistic analysis
B1	CB6 ? Students should have the knowledge and understanding necessary to provide a basis or opportunity for originality in the development and/or application of ideas, often in a research context.
B3	CB8 - Students must be able to integrate knowledge and to deal with the complexity of judgement formulation starting with information, which might be incomplete or limited, and which includes reflections on social and ethical responsibilities linked to the application of their knowledge and judgement.
B6	G01 ?The capacity to delve into those concepts, principles, theories or models related with the different fields of English Studies is a necessary skill, as is the ability to solve specific problems in a particular field of study via appropriate methodology.
B8	G03 ? An efficient use of new information technology and communication in English Studies is a necessary skill.
B10	G05 ? Skills related to research and the handling of new knowledge and information in the context of English Studies are to be acquired by students
B11	G06 ? Students should be able to develop a critical sense in order to assess the relevance of both existing research in the fields of English Studies, and their own research.
B12	G07 ?Linguistic competence (C2 level) in oral and written English must be developed and consolidated.
B14	G09 ? Students are expected to be able to carry out research projects of an academic nature in the different fields of English Studies
B15	G10 ? The ability to present and defend a research project using adequate terminology and resources appropriate to the field of study is a skill which should be acquired.

Learning outcomes

Learning outcomes	Study programme competences
-------------------	-----------------------------



	AR1	BR1
	AR2	BR3
	AR14	BR6
		BR8
		BR10
		BR11
		BR12
		BR14
		BR15

Contents	
Topic	Sub-topic
1. Introduction (description vs. theory; corpus vs. computational linguistics; brief history of corpus linguistics).	
2. What is a corpus? (defining a corpus, types of corpora, corpus resources).	
3. Corpus design and compilation (size, representativeness, etc.).	
4. Corpus annotation (tagging, parsing, other types of annotation).	
5. Data retrieval (software tools, concordances, wildcards, keywords, word lists, etc.).	
6. Data analysis (quantitative and qualitative analyses, normalized frequencies, frequency distribution, statistical significance).	
7. Applications of corpora for the linguistic analysis of English.	
Practical demonstrations with a selection of different corpora.	

Planning				
Methodologies / tests	Competencies	Ordinary class hours	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A14 B1 B3	10	5	15
ICT practicals	B8 B10 B11	10	18	28
Oral presentation	B6 B8 B10 B12	1	9	10
Practical test:	A2 B12 B14 B15	1	18	19
Introductory activities	A1	1	0	1
Personalized attention		2	0	2

(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.

Methodologies	
Methodologies	Description
Guest lecture / keynote speech	Interactive lecturing on Corpus Linguistic Issues.
ICT practicals	Hands-on sessions with software tools and corpora for text analysis.
Oral presentation	Oral presentation of a case study. Students will have to provide a hand-out and a PowerPoint presentation.
Practical test:	on the role of corpus linguistics as a framework/methodology for linguistic investigation.
Introductory activities	Ice-breaking activities.

Personalized attention



Methodologies	Description
ICT practicals	Personal tutoring for specific support.

Assessment			
Methodologies	Competencies	Description	Qualification
ICT practicals	B8 B10 B11	Hands-on sessions with software tools and corpora for text analysis.	30
Oral presentation	B6 B8 B10 B12	Oral presentation of a case study. Students will have to provide a hand-out and a PowerPoint presentation.	30
Practical test:	A2 B12 B14 B15	A test on the role of corpus linguistics as a framework/methodology for linguistic investigation.	40

Assessment comments
Second opportunity: Students will have to repeat only those parts (exercises and case study) in which they did not reach a pass in the first opportunity. In the second opportunity the case study will be submitted as a written essay (between 2,000-3,000 words, excluding references). Students officially exempt from class attendance: Students will have to hand in all course exercises (60%) and submit an essay on a case study (between 2,000-3,000 words, excluding references) (40%). All exercises and tests will be done in English. Overall correctness in language and in the use of formal conventions is indispensable. Therefore, grammatical and orthographic mistakes will be penalised. Those who are officially registered as part-time students, and have been granted permission not to attend classes, as stipulated in the regulations of this University, will be assessed in either of the opportunities according to the same criteria specified for the second opportunity. Part-time students who have been granted a "dispensa académica" must contact the teacher at the beginning of the semester in order to design the necessary changes and adjustments relating to the different activities to be carried out along the academic year. LINGUISTIC ACCURACY Both the theoretical and practical examination must meet the minimum requirements of linguistic correctness (spelling and grammar, punctuation, syntax concordance, lexical precision, or formal register). The level in English language expected of students is equivalent to a C1 level. Deficiencies in this area will be penalized in qualifying. Essays and written work may be tested through Turnitin in order to detect wrong citations, plagiarism or any other type of fraud. Should this occur, the rules for plagiarism will be applied. (Normas de avaliación, revisión e reclamación das cualificacións dos estudos de grao emestrado universitario). Turnitin recognises papers previously turned in by other people (or the student him/herself) at this university or other universities, as well as other material found on Internet.

Sources of information



Basic	Aarts, Jan, Pieter de Haan & Nelleke Oostdijk (eds.). 1993. English language corpora: Design, analysis and exploitation. Amsterdam: Rodopi. Baker, Paul. 2010. Sociolinguistics and corpus linguistics. Edinburgh: Edinburgh University Press. Baker, Paul, Andrew Hardie & Tony McEnery. 2006. A glossary of corpus linguistics. Edinburgh: Edinburgh University Press. Biber, Douglas, Susan Conrad & Randi Reppen. 1998. Corpus linguistics. Investigating language structure and use. Cambridge: Cambridge University Press. Biber, Douglas & Randi Reppen (eds.). 2015. The Cambridge handbook of English corpus linguistics. Cambridge: Cambridge University Press. Cantos, Pascual. 2011. Statistical methods in language and linguistic research. London: Equinox. Hoffmann, Sebastian, Stefan Evert, Nick Smith, David Lee & Ylva Berglund Prytz. 2008. Corpus linguistics with BNCweb - a practical guide. Frankfurt am Main: Peter Lang. Hunston, Susan. 2002. Corpora in applied linguistics. Cambridge: Cambridge University Press. Kennedy, Graeme. 1998. An introduction to corpus linguistics. London: Longman. Kilgarriff, Adam & Gregory Grefenstette. 2003. Introduction to the special issue on the web as corpus. Computational Linguistics 29(3). 333-347. Lavid, Julia. 2005. Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI. Madrid: Cátedra. Lindquist, Hans. 2009. Corpus linguistics and the description of English. Edinburgh: Edinburgh University Press. Lüdeling, Anke & Merja Kytö (eds.). 2008. Corpus linguistics. An international handbook. Volume I. Berlin/New York: Walter de Gruyter. McEnery, Tony & Andrew Wilson. 1996. Corpus linguistics. Edinburgh: Edinburgh University Press. McEnery, Tony, Richard Xiao & Yukio Tono. 2006. Corpus-based language studies. An advanced resource book. London: Routledge. Meyer, Charles. 2002. English corpus linguistics: An introduction. Cambridge: Cambridge University Press. Mitkov, Ruslan (ed.). 2003. The Oxford handbook of computational linguistics. Oxford: Oxford University Press. Mukherjee, Joybrato. 2004. The state of the art in corpus linguistics: Three book-length perspectives. English Language and Linguistics 8(1). 103-119. Oakes, Michael. 1998. Statistics for corpus linguistics. Edinburgh: Edinburgh University Press. Sinclair, John. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press. Stubbs, Michael. 1996. Text and corpus analysis. Oxford: Blackwell.
Complementary	

Recommendations
Subjects that it is recommended to have taken before
Subjects that are recommended to be taken simultaneously
Cognition and Cognitive Processes in English/613505011
Contrastive Linguistics in English /613505111
Subjects that continue the syllabus
Other comments

(*)The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.