# UNIVERSIDADE DA CORUÑA

| **Teaching Guide** | | | | |
|---|---|---|---|---|
| **Identifying Data** | | | | **2018/19** |
| **Subject (*)** | Corpus Linguistics and Computer Science in English | | **Code** | 613505010 |
| **Study programme** | Mestrado Universitario en Estudos Ingleses Avanzados e as súas Aplicacións (2013) | | | |
| **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** | Moskowich-Spiegel Fandiño, Isabel | | **E-mail** | isabel.moskowich-spiegel.fandino@udc.es |
| **Lecturers** | Moskowich-Spiegel Fandiño, Isabel | | **E-mail** | isabel.moskowich-spiegel.fandino@udc.es |
| **Web** | www.imaes.eu/?page_id=31 | | | |
| **General description** | | | | |

| **Study programme competences / results** | |
|---|---|
| **Code** | **Study programme competences / results** |
| 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 |
| 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. |
| B7 | G02 ? Students must be capable of applying the knowledge acquired in the multidisciplinary and multifaceted area of English Studies |
| B8 | G03 ? An efficient use of new information technology and communication in English Studies is a necessary skill. |
| B9 | G04 ? Students must be able to publicly present their ideas, reports or experiences, as well as give informed opinions based on criteria, external norms or personal reflection. All of this implies having sufficient command of both oral and written academic and scientific language |
| 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. |
| B13 | G08 ? Students should become progressively autonomous in the learning process, and in the search for appropriate resources and information, via the use of bibliographic and documentary sources related to English Studies. |
| 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 / results** | | |
| To know and use the principles and methodology of English corpus linguisti | AR2 | BR8 BR10 BR13 | |
| To acquire the practical use of linguistic corpora and related software | | BR8 | |

| | | |
|---|---|---|
| To be able to discuss specialized literature on corpus linguisti | AR1 | BR6 |
| | AR14 | BR7 |
| | | BR11 |
| | | BR12 |
| | | BR14 |
| To be able to reach conclusions that can be generalised basinf on particular linguistic data and present the research results. | | BR9 |
| | | BR11 |
| | | BR12 |
| | | BR14 |
| | | BR15 |

| Contents | |
|---|---|
| **Topic** | **Sub-topic** |
| 1. Introduction (description vs. theory; corpus vs. computational linguistics; brief history of corpus linguistics). | descrición vs. teoría<br>lingüística de corpus vs. lingüística computacional<br>breve historia da lingüística de corpus |
| 2. What is a corpus? (defining a corpus; types of corpora; corpus resources). | definición de corpus<br>tipos de corpus<br>recursos. |
| 3. Corpus design and compilation (size and representativeness, etc.). | Mostreo<br>tamaño<br>representatividade<br>equilibrio |
| 4. Corpus annotation (tagging, parsing, other types of annotation). | Codificación, etiquetado e anotación.<br>etiquetado morfolóxico<br>etiquetado sintáctico<br>outros tipos de anotación |
| 5. Data retrieval (software tools, concordances, wildcards, keywords, word lists, etc.). | ferramentas de software<br>concordancias<br>comodíns<br>palabras chave<br>listas de palabras (frequencias) |
| 6. Data analysis (quantitative and qualitative analyses; normalized frequencies; frequency distribution; statistical significance). | análises cuantitativas e cualitativas<br>frecuencias normalizadas<br>distribución de frecuencias<br>significatividade estatística |
| 7. Applications of corpora for the linguistic analysis of English. | |
| 8. Hands-on demonstrations with a selection of corpora. | |

| Planning | | | | |
|---|---|---|---|---|
| **Methodologies / tests** | **Competencies / Results** | **Teaching hours (in-person & virtual)** | **Student?s personal work hours** | **Total hours** |
| Introductory activities | A1 A2 B11 | 10 | 20 | 30 |
| Document analysis | B6 B8 | 0 | 10 | 10 |
| ICT practicals | A14 B13 | 4 | 6 | 10 |
| Case study | B7 B9 B10 B12 B14 | 0 | 11 | 11 |
| Personalized attention | | 14 | 0 | 14 |
| (*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students. | | | | |

| Methodologies |
|---|
| |

| Methodologies | Description |
|---|---|
| Introductory activities | - Lectures on the role of corpus linguistics as a framework/methodology for linguistic investigation. |
| Document analysis | - Analisys of particular cases for a written piece of work. |
| ICT practicals | - Hands-on sessions with software tools and corpora for text analysis (available at computer clusters and via internet). |
| Case study | - Seminars and oral presentations of the student?s written piece of work |

| Personalized attention | |
|---|---|
| **Methodologies** | **Description** |
| ICT practicals | Students will be using the Moodle platdorm as well as other virtual means to contact the teacher. Besides this, they will receive personalised attention during the teacher's office hours. |

| Assessment | | | |
|---|---|---|---|
| **Methodologies** | **Competencies / Results** | **Description** | **Qualification** |
| Document analysis | B6 B8 | O estudantado terá que facer as lecturas obrigatoriase, nalgún caso, escribir unha recensión ademais da exercicios de diverso tipo e actividades que en liña pola plataforma Moodle ou por outras vías electrónicas . | 55 |
| Case study | B7 B9 B10 B12 B14 | Realización dun traballo con datos para a súa análise cuantitativa e cualirtativa sobre un aspecto concreto da lingua inglesa. O estudantado terá que facer a presentación oral do traballo e entregar a presentación. | 30 |
| ICT practicals | A14 B13 | Participación activa nas aulas e nos debates virtuais | 15 |

| Assessment comments |
|---|

First opportunity:

15%    Active participation in the

sessions. Please note that attendance is compulsory.

55%    Exercises and programmed

readings.

30%    Oral presentation of a case study. Students

will have to provide a hand-out and a PowerPoint presentation.

Exercises

as well as the hand-out and the PowerPoint presentation for the case study will

be submitted via the teaching platform.

Pieces of work including plagiarized material will be marked with 0. To detect plagiarism, the application Turnitin could be used. This application

recognises papers previously turned in by other people (or the student him/herself) at this university or other universities.

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.

**Sources of information**

| | |
|---|---|
| **Basic** | Aarts, J., P. de Haan & N. Oostdijk (eds.). 1993. English Language Corpora: Design, Analysis and Exploitation. Amsterdam: Rodopi.Baker, P. 2010. Sociolinguistics and Corpus Linguistics. Edinburgh: Edinburgh University Press.Baker, P., A. Hardie & T. McEnery. 2006. A Glossary of Corpus Linguistics. Edinburgh: Edinburgh University Press.Biber, D., S. Conrad & R. Reppen. 1998. Corpus Linguistics. Investigating Language Structure and Use. Cambridge: C.U.P.Cantos, P. 2011. Statistical Methods in Language and Linguistic Research. London: Equinox.Hoffmann, S., S. Evert, N. Smith, D. Lee & Y. 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, G. 1998. An Introduction to Corpus Linguistics. London: Longman.Kilgariff, A. & G. Grefenstette. 2003. Introduction to the Special Issue on the Web as Corpus. Computational Linguistics 29/3: 333-347.Lavid, J. 2005. Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI. Madrid: Cátedra.Lindquist, H. 2009. Corpus Linguistics and the Description of English. Edinburgh: Edinburgh University Press.Lüdeling, A. & M. Kytö (eds.). 2008. Corpus Linguistics. An International Handbook. Volume I. Berlin/New York: Walter de Gruyter.McEnery, T. & A. Wilson. 1996. Corpus Linguistics. Edinburgh: Edinburgh University Press.McEnery, T., R. Xiao & Y. Tono. 2006. Corpus-Based Language Studies. An Advanced Resource Book. London: Routledge.Meyer, Ch. 2002. English Corpus Linguistics: An Introduction. Cambridge: Cambridge University Press.Mitkov, R. (ed.). 2003. The Oxford Handbook of Computational Linguistics. Oxford: Oxford University Press. Mukherjee, J. 2004. ?The State of the Art in Corpus Linguistics: Three book-length perspectives.? English Language and Linguistics 8/1: 103-119.Oakes, M. 1998. Statistics for Corpus Linguistics. Edinburgh: Edinburgh University Press.Sinclair, J. 1991. Corpus, Concordance, Collocation. Oxford: Oxford University Press.Stubbs, M. 1996. Text and Corpus Analysis. Oxford: Blackwell. |
| **Complementary** | - Aarts, J., P. de Haan &amp; N. Oostdijk (eds.). 1993. English Language Corpora: Design, Analysis and Exploitation. Amsterdam: Rodopi. - Baker, P. 2010. Sociolinguistics and Corpus Linguistics. Edinburgh: Edinburgh University Press. - Baker, P., A. Hardie &amp; T. McEnery. 2006. A Glossary of Corpus Linguistics. Edinburgh: Edinburgh University Press. - Biber, D., S. Conrad &amp; R. Reppen. 1998. Corpus Linguistics. Investigating Language Structure and Use. Cambridge: C.U.P. - Cantos, P. 2011. Statistical Methods in Language and Linguistic Research. London: Equinox. - Hoffmann, S., S. Evert, N. Smith, D. Lee &amp; Y. 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, G. 1998. An Introduction to Corpus Linguistics. London: Longman. - Kilgariff, A. &amp; G. Grefenstette. 2003. Introduction to the Special Issue on the Web as Corpus. Computational Linguistics 29/3: 333-347. - Lavid, J. 2005. Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI. Madrid: Cátedra. - Lindquist, H. 2009. Corpus Linguistics and the Description of English. Edinburgh: Edinburgh University Press. - Lüdeling, A. &amp; M. Kytö (eds.). 2008. Corpus Linguistics. An International Handbook. Volume I. Berlin/New York: Walter de Gruyter. - McEnery, T. &amp; A. Wilson. 1996. Corpus Linguistics. Edinburgh: Edinburgh University Press. - McEnery, T., R. Xiao &amp; Y. Tono. 2006. Corpus-Based Language Studies. An Advanced Resource Book. London: Routledge. - Meyer, Ch. 2002. English Corpus Linguistics: An Introduction. Cambridge: Cambridge University Press. - Mitkov, R. (ed.). 2003. The Oxford Handbook of Computational Linguistics. Oxford: Oxford University Press. - Mukherjee, J. 2004. "The State of the Art in Corpus Linguistics: Three book-length perspectives." English Language and Linguistics 8/1: 103-119. - Oakes, M. 1998. Statistics for Corpus Linguistics. Edinburgh: Edinburgh University Press. - Sinclair, J. 1991. Corpus, Concordance, Collocation. Oxford: Oxford University Press. - Stubbs, M. 1996. Text and Corpus Analysis. Oxford: Blackwell. |

| Recommendations |
|---|
| **Subjects that it is recommended to have taken before** |
| |
| **Subjects that are recommended to be taken simultaneously** |
| |
| **Subjects that continue the syllabus** |
| |
| **Other comments** |

Students must have a C1 level in English in order to be able to follow the course. 

Students are
expected to complete all assignments and readings suggested by the lecturer(s),
and to come to the sessions prepared to discuss them. Students must visit the
eLearning platform for the course regularly. Students are expected to check
their university email on a regular basis: announcements and last-minute
changes will be notified via email.
Academic misconduct (including cheating, plagiarism, self-plagiarism, collusion
or fabrication of results) will not be tolerated and will be penalised.