



DATOS IDENTIFICATIVOS

Lingüística Informática e de Corpus Aplicadas á Lingua Inglesa

Materia	Lingüística Informática e de Corpus Aplicadas á Lingua Inglesa			
Código	V01M121V01202			
Titulación	Máster Universitario en Estudios Ingleses Avanzados e as súas Aplicacións			
Descritores	Creditos ECTS	Sinale	Curso	Cuadrimestre
	3	OP	1	2c
Lingua de impartición	Inglés			
Departamento				
Coordinador/a	González Álvarez, María Dolores			
Profesorado	González Álvarez, María Dolores Yáñez Bouza, Nuria			
Correo-e	dglez@uvigo.es			
Web	http://fatic.uvigo.es/			
Descrición xeral	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.			

Competencias de titulación

Código	
A1	Capacidade para utilizar unha terminoloxía lingüística rigorosa así como unha notación adecuada.
A3	Coñecemento dos principais recursos e ferramentas de investigación lingüística.
A11	Capacidade de aplicar métodos de investigación cuantitativa e cualitativa, de executar varios procedementos de recolleita e análise de datos e de utilizar software específico para a investigación en lingüística inglesa.
B1	Capacidade de afondar naqueles conceptos, principios, teorías ou modelos relacionados cos distintos campos dos Estudos Ingleses, así como de coñecer a metodoloxía necesaria para a resolución de problemas propios de devandita área de estudo
B4	Capacidade para elaborar adecuadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, artigos científicos e formular hipóteses razoables.
B5	Capacidade para presentar en público experiencias, ideas ou informes, así como emitir xuízos en función de criterios, de normas externas ou de reflexións persoais para o que será necesario alcanzar un dominio suficiente da linguaxe académica e científica tanto na súa vertente escrita como oral.
B7	Habilidade para traballar e continuar a aprendizaxe de forma autónoma e autodirixida.
B8	Habilidades para investigar e manexar novos coñecementos e información no contexto dos Estudos Ingleses.
B9	Capacidade para adquirir un espírito crítico que leve aos estudantes a considerar a pertinencia das investigacións existentes nas áreas de estudo que conforman os Estudos Ingleses, así como das súas propias.
B11	Capacidade para traballar en equipo, cooperativo e compartido, fomentando a tolerancia e a aprendizaxe colaborativo.
B15	Desenvolvemento da capacidade comunicativa no ámbito académico.
B16	Capacidade crítica e de síntese do estudante mediante a lectura e revisión de artigos especializados.

B17	Autonomía progresiva na aprendizaxe, procuraas propias de recursos e información, accedendo para iso a fontes bibliográficas e documentais sobre os distintos ámbitos que conforman os Estudos Ingleses.
B18	Capacidade para obter información sobre a investigación máis recente que se está levando a cabo dentro do ámbito dos Estudos Ingleses, tanto en lingüística como en literatura e cultura.
B19	Capacidade para realizar traballos de investigación de carácter académico nos distintos ámbitos dos Estudos Ingleses.
B20	Capacidade para presentar e defender un traballo de investigación utilizando a terminoloxía e os recursos adecuados e apropiados dentro do campo obxecto de estudo.

Competencias de materia

Resultados previstos na materia	Tipoloxía	Resultados de Formación e Aprendizaxe
Students will gain good background knowledge of the principles and methodology of English corpus linguistics. 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.	saber saber facer	A11 B1 B4 B8 B19
Students will acquire intellectual skills in identifying adequate methods of data collection, and they will learn to carry out linguistic research in an appropriate and ethical way.	saber saber facer	B1 B4 B20
Students will gain the ability to design, collect and analyse corpus material, including the use of software tools to search, retrieve and sift data. Students will develop skills in interpreting information presented in the form of diagrams, tables and graphs; they will be able to apply and interpret tests of statistical significance.	saber saber facer	A3 A11
Students will improve their skills in academic writing. On successful completion of the course students will have shown the ability to design a research project and develop a written argument of depth and complexity, using primary sources and critical literature, with a standard of scholarly presentation of the material produced appropriate to MA study.	saber facer	B4 B16 B19 B20
Students will improve their skills in critical analysis such as the discussion of specialized literature on corpus linguistics.	saber saber facer	B9 B16 B18
Students will practice skills and strategies relevant to a successful completion of an MA degree programme, such as time management, group work, working to deadlines, self-directed study and research, etc. Students will be able to use appropriate notation and referencing conventions in linguistic research. They will also acquire skills in the use of library and database resources for independent research.	saber saber facer	A1 A3 B7 B9 B11 B17
Students will become familiar with strategies for a good delivery of an oral presentation, including the presentation of handouts.	saber saber facer	B5 B15

Contidos

Tema	
1. Introduction	Description vs. theory (armchair vs. butterflies). Corpus vs. computational linguistics. Brief history of corpus linguistics.
2. What is a corpus?	Defining a corpus. Types of corpora. Main corpus resources for linguistic research.
3. Corpus design and compilation	Sampling matters: e.g. size, representativeness, collecting and computerizing data, storage and copyright.
4. Corpus annotation	Tagging, parsing and other types of annotation.
5. Data retrieval	Software tools (e.g. Wordsmith). Precision and Recall. Concordances and queries. Keywords, word lists, frequency lists, etc.
6. Data analysis	Quantitative and qualitative analyses. Normalised frequencies and frequency distribution. Statistical significance.
7. Applications of corpora for the linguistic analysis of English	Selection of case studies on lexicon, grammar, discourse markers, register variation, historical change, sociolinguistics, etc.
8. Hands-on demonstrations with a selection of corpora	ARCHER, the Helsinki Corpus, etc.

Planificación

	Horas na aula	Horas fóra da aula	Horas totais
Actividades introdutorias	2	6	8
Sesión maxistral	5.5	6	11.5
Prácticas en aulas de informática	4	4	8
Resolución de problemas e/ou exercicios	2	15	17

Traballos e proxectos	0.5	30	30.5
*Os datos que aparecen na táboa de planificación son de carácter orientador, considerando a heteroxeneidade do alumnado.			

Metodoloxía docente	
	Descrición
Actividades introdutorias	Actividades encamiñadas a tomar contacto e reunir información sobre o alumnado, así como a presentar a materia.
Sesión maxistral	Exposición por parte das profesoras dos contidos sobre a materia obxecto de estudo, bases teóricas e/ou directrices dun traballo, exercicio ou proxecto a desenvolver polo estudante.
Prácticas en aulas de informática	Exposición por parte da profesora e posta en práctica por parte do alumnado dos coñecementos de corpora, software tools, queries and concordances, collocations, frequency lists, etc.
Resolución de problemas e/ou exercicios	Actividade na que se formulan problema e/ou exercicios relacionados coa materia. O alumnado debe desenvolver as solucións axeitadas ou correctas mediante a aplicación de procedementos de transformación da información dispoñible e a interpretación dos resultados. Adóitase empregar como complemento da lección maxistral.

Atención personalizada	
Metodoloxías	Descrición
Resolución de problemas e/ou exercicios	Para un óptimo seguimento e aproveitamento da materia, o alumnado contará con atención personalizada en modalidade presencial (na aula, no despaho durante horas de tutorías) e en modalidade virtual (a través das ferramentas que ofrece a plataforma de teledocencia e o correo electrónico). Esta atención personalizada estará encamiñada a atender as necesidades e consultas do alumnado relacionadas co estudo e/ou teams vinculados coa materia, proporcionándolle orientación, apoio e motivación no proceso de aprendizaxe.
Prácticas en aulas de informática	Para un óptimo seguimento e aproveitamento da materia, o alumnado contará con atención personalizada en modalidade presencial (na aula, no despaho durante horas de tutorías) e en modalidade virtual (a través das ferramentas que ofrece a plataforma de teledocencia e o correo electrónico). Esta atención personalizada estará encamiñada a atender as necesidades e consultas do alumnado relacionadas co estudo e/ou teams vinculados coa materia, proporcionándolle orientación, apoio e motivación no proceso de aprendizaxe.
Probas	Descrición
Traballos e proxectos	Para un óptimo seguimento e aproveitamento da materia, o alumnado contará con atención personalizada en modalidade presencial (na aula, no despaho durante horas de tutorías) e en modalidade virtual (a través das ferramentas que ofrece a plataforma de teledocencia e o correo electrónico). Esta atención personalizada estará encamiñada a atender as necesidades e consultas do alumnado relacionadas co estudo e/ou teams vinculados coa materia, proporcionándolle orientación, apoio e motivación no proceso de aprendizaxe.

Avaliación		
	Descrición	Cualificación
Sesión maxistral	Participación e implicación do alumnado nas sesións presenciais: atención, comentarios e preguntas ao fío das explicacións e resolución de actividades encomendadas.	5
Prácticas en aulas de informática	Participación e implicación do alumnado nas sesións de informática: atención, comentarios e preguntas ao fío das explicacións e resolución de actividades encomendadas.	5
Resolución de problemas e/ou exercicios	Resolución de exercicios, problemas ou pequenos traballos que o alumnado realizará en clase ou de forma autónoma fóra da aula e que serán puntuados pola profesora. Algúns destes exercicios serán auto-avaluados polo alumnado mediante a plataforma virtual.	30
Traballos e proxectos	Proba final cun compoñente práctico e teórico. O alumno deberá entregar un traballo escrito individual de 2,000-3,000 palabras (excluíndo bibliografía e apéndices) onde porá en práctica os coñecementos adquiridos no curso. Antes da entrega do traballo escrito, o alumnado fará unha exposición oral individual de dito traballo, no que se valorará os contidos e o aspecto estético. O alumnado recibirá feedback pra poder mellorar a versión final escrita.	60

Outros comentarios sobre a Avaliación

Attendance is **compulsory**.

Academic misconduct (including cheating, plagiarism, self-plagiarism, collusion or fabrication of results) will not be tolerated and will be **penalised** with a zero mark or with a fail for the entire course, depending on the task involved.

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.

Bibliografía. Fontes de información

Recommended bibliography

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.

Recomendaciones
