



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	Yáñez Bouza, Nuria			
Profesorado	Yáñez Bouza, Nuria			
Correo-e	nuria.y.b@uvigo.es			
Web	http://fatic.uvigo.es/			

Descrición xeral GALEGO.
Este curso ofrece unha introdución ao eido da lingüística de corpus e á metodoloxía baseada en traballos de corpora para o estudo da lingua inglesa. Trataranse aspectos de deseño, compilación e análise de materiais de corpus. Nas sesións prácticas as/os estudantes aprenderán a facer búsquedas con ferramentas software. As/os estudantes desenvolverán a habilidade de describir e debater a variación lingüística e como se pode rastrexar o cambio e variación lingüísticos por medio de exemplos en corpora.

As/os estudantes deste curso adquirirán competencias en:

- ☐ os principios e metodoloxías da lingüística de corpus da lingua inglesa;
- ☐ o uso práctico de corpora e ferramentas software;
- ☐ debater lecturas no eido da lingüística de corpus;
- ☐ presentar os resultados da súa investigación.

ENGLISH.

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

Código	
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
B2	Capacidade para aplicar os coñecementos adquiridos na contorna multidisciplinar e multifacética dos Estudos Ingleses.
B3	Capacidade para utilizar de forma eficiente as novas tecnoloxías da información e da comunicación no ámbito dos Estudos Ingleses.

B4	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.
B5	Habilidade para investigar e manexar novos coñecementos e información no contexto dos Estudos Ingleses.
B6	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.
B7	Habilidades de consolidación e desenvolvemento da competencia lingüística (nivel C2) no uso falado e escrito da lingua inglesa.
B8	Autonomía progresiva na aprendizaxe, procura propias de recursos e información, accedendo para iso a fontes bibliográficas e documentais sobre os distintos ámbitos que conforman os Estudos Ingleses.
B9	Capacidade para realizar traballos de investigación de carácter académico nos distintos ámbitos dos Estudos Ingleses.
C2	Coñecemento dos principais modelos de investigación lingüística.
C3	Coñecemento dos principais recursos e ferramentas de investigación lingüística.
C11	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.

Resultados de aprendizaxe

Resultados previstos na materia	Resultados de Formación e Aprendizaxe
Coñecemento dos principios e das metodoloxías do eido da lingüística de corpus. As/os estudantes desenvolverán a habilidade de describir e debater a variación lingüística e como se pode rastrexar o cambio e variación lingüísticos por medio de exemplos en corpora.	B1 B5 B8 C2 C3 C11
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.	
Adquisición das competencias intelectuais para identificar os métodos axeitados para a compilación de datos, e para unha investigación baseada en principios éticos.	B2 B5 B6 C2
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.	
Adquirir uns coñecementos básicos sobre dereito e xestión	
Habilidade para deseñar, compilar e analizar datos de corpora, incluíndo o uso de ferramentas de software para a búsqueda, extracción e filtro de datos. Habilidade para interpretar a información en forma de gráficas, táboas, diagramas, etc.. Habilidade para aplicar métodos estatísticos aos datos obtidos.	B3 B9 C3 C11
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.	
Desenvolvemento das habilidades no contexto académico. Habilidade de deseñar un proxecto (case study), cun argumento sólido e lóxico, facendo uso de fontes bibliográficas. Capacidade de comunicar e presentar o proxecto nun estándar académico axeitado a nivel MA.	B1 B4 B6 B7 B9 C11
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.	
Práctica das competencias e estratexias relevantes para un programa de mestrado, por exemplo a eficacia do uso do tempo, o cumprimento coas datas de entrega de traballos, a aprendizaxe autónoma, as lecturas fóra da aula, o uso de bases de datos bibliográficas, etc. Aprendizaxe das convencións académicas en aspectos como a notación de termos lingüísticos e as notacións bibliográficas.	B7 B8 B9 C3
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.	
Familiarización coas estratexias axeitadas para unha comunicación óptima da presentación oral, incluíndo a presentación do handout e a Power Point.	B7 B8 C3
Students will become familiar with strategies for a good delivery of an oral presentation, including the presentation of handouts.	

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	3	5
Sesión maxistral	8	0	8
Resolución de problemas e/ou exercicios	3	38	41
Traballos e proxectos	1	20	21
*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.
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.
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	Resultados de Formación e Aprendizaxe	
Sesión maxistral	Participación e implicación do alumnado nas sesións presenciais: atención, comentarios e preguntas ao fío das explicacións e resolución de actividades encomendadas.	15	B1 B6 B7	C2 C3

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 poden ser auto-evaluados polo alumnado mediante a plataforma virtual.	55	B1 B2 B5 B6 B7 B8 B9	C2 C3 C11
Traballos e proxectos	Proba final cun compoñente práctico e teórico. O alumno deberá facer unha exposición oral de un caso práctico, na que se valorarán os contidos e a presentación en sí, que deberá estar en soporte ppt. O alumnado deberá entregar un 'hand-out' ao profesorado durante a presentación.	30	B1 B3 B4 B7 B9	C2 C3 C11

Outros comentarios sobre a Avaliación

GALEGO

Primeira edición de actas:

- 15% participación activa nas sesións. Deber lembrarse que **a asistencia é obrigatoria**.
- 55% exercicios, traballos, lecturas. As datas de entrega serán acordadas entre a profesora e o alumnado. Xeralmente, os traballos serán entregados despois das unidades temáticas.
- 30% presentación oral dun "case study", incluíndo un "handout" e presentación Power Point. As datas da presentación serán acordadas entre a profesora e o alumnado. En xeral, a presentación terá lugar nun prazo dun mes despois de rematar o curso.

Segunda edición de actas (xullo):

- As/os estudantes que suspendan na primeira edición de actas soamente terán que repetir as partes que teñan suspensas. En relación ao "case study", na segunda oportunidade deberán entregar un traballo escrito (2000-3000 palabras, excluindo as referencias e os apéndice) e deberán facer a presentación oral. As datas de entrega serán consensuadas coa profesora dentro dos prazos establecidos pola Xunta de FFT, e será a mesma para todo o alumnado.

ENGLISH

First opportunity

- 15% active participation in the sessions. Please note that attendance is compulsory.
- 55% exercises, assignments, and programmed readings. In general, assignments will be submitted after each thematic unit.
- 30% oral presentation of a case study. Students will have to provide a handout and a powerpoint presentation. In general, the deadline will be within a month after the end of the course.

Second opportunity (July) :

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 2000-3000 words excluding references). The submission deadline will be agreed on between the lecturer and the students, within the period established by the Xunta da FFT, and it will be the same for all registered students.

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.
- Kilgarriff, 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.

Recomendacións

Materias que se recomenda cursar simultaneamente

Variación e Cambio na Lingua Inglesa/V01M121V01201

Materias que se recomenda ter cursado previamente

Métodos e Recursos de Investigación Lingüística e a súa Aplicación á Lingua Inglesa/V01M121V01101

Outros comentarios

GALEGO.

Este curso será impartido en lingua inglesa.

A asistencia presencial neste curso é obrigatoria.

A profesora entregará handouts para as sesións presenciais antes da sesión.

Espérase que o alumnado realice tódolos traballos, exercicios e lecturas, e que asista a clase preparado para debatilos. Esperáase tamén que o alumnado consulte a plataforma virtual FAITIC regularmente, e que consulten o correo da universidade, pois as novidades e anuncios serán comunicados fundamentalmente por correo.

Non se tolerará ningún tipo de plaxio, e de cometerse éste será penalizado severamente cun suspenso na totalidade da materia. Alegar descoñecemento non eximirá a/o estudante da súa responsabilidade.

ENGLISH.

The course will be delivered in English.

Attendance is compulsory.

Handouts will be provided to accompany lectures and seminar work.

Students are expected to complete all assignments and readings suggested by the lecturer, 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 with a zero mark or with a fail for the entire course, depending on the task involved. Pledging lack of awareness about plagiarism will not excuse students from their responsibility.
