

COURSE OUTLINE

(1) GENERAL

SCHOOL	HUMANITIES		
ACADEMIC UNIT	DEPARTMENT OF MEDITERRANEAN STUDIES		
LEVEL OF STUDIES	UNDERGRADUATE		
COURSE CODE	ΓΥΕ-25	SEMESTER	7
COURSE TITLE	CORPUS PROCESSING		
INDEPENDENT TEACHING ACTIVITIES <i>if credits are awarded for separate components of the course, e.g. lectures, laboratory exercises, etc. If the credits are awarded for the whole of the course, give the weekly teaching hours and the total credits</i>	WEEKLY TEACHING HOURS	CREDITS	
	3	5	
<i>Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).</i>			
COURSE TYPE <i>general background, special background, specialised general knowledge, skills development</i>	Special background		
PREREQUISITE COURSES:	No		
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	Greek. In the case of ERASMUS students, Greek & English		
IS THE COURSE OFFERED TO ERASMUS STUDENTS	Yes		
COURSE WEBSITE (URL)	http://dms.aegean.gr/en/undergraduate-studies/program-of-studies-2019-20/		

(2) LEARNING OUTCOMES

<p>Learning outcomes <i>The course learning outcomes, specific knowledge, skills and competences of an appropriate level, which the students will acquire with the successful completion of the course are described.</i></p> <p><i>Consult Appendix A</i></p> <ul style="list-style-type: none"> • <i>Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area</i> • <i>Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B</i> • <i>Guidelines for writing Learning Outcomes</i> <p>After the successful complete of the course, students should be able to:</p> <ul style="list-style-type: none"> - perceive what a Corpus is - understand the need for Corpora - distinguish corpora types

- apply normalization when needed
- know about Linux and why its use by linguists for corpus processing purposes
- know Linux environment
- understand what programming is
- implement basic tools such as word lists
- deal with issues of corpus processing for Humanities and Social Sciences purposes
- deal with problems in a formal, algorithmic way

General Competences

Taking into consideration the general competences that the degree-holder must acquire (as these appear in the Diploma Supplement and appear below), at which of the following does the course aim?

<i>Search for, analysis and synthesis of data and information, with the use of the necessary technology</i>	<i>Project planning and management</i>
<i>Adapting to new situations</i>	<i>Respect for difference and multiculturalism</i>
<i>Decision-making</i>	<i>Respect for the natural environment</i>
<i>Working independently</i>	<i>Showing social, professional and ethical responsibility and sensitivity to gender issues</i>
<i>Team work</i>	<i>Criticism and self-criticism</i>
<i>Working in an international environment</i>	<i>Production of free, creative and inductive thinking</i>
<i>Working in an interdisciplinary environment</i>	<i>.....</i>
<i>Production of new research ideas</i>	<i>Others...</i>
	<i>.....</i>

- Search for, analysis and synthesis of data and information, with the use of the necessary technology
- Adapting to new situations
- Decision-making
- Working independently
- Team work
- Working in an interdisciplinary environment
- Production of new research ideas

(3) SYLLABUS

- What is a Corpus – definitions
- Corpora types
- Corpus annotation
- Frequency normalisation
- Corpus-based studies
- Corpus-based applications to Humanities and Social Sciences
- Corpus design and construction
- Greek corpora on the web
- Linux
- The Terminal – basic Linux commands
- Linux for Humanities and Social scientists
- Programming for the extraction of word-frequency lists
- program checking/correcting
- Saving the results for future reference

(4) TEACHING and LEARNING METHODS - EVALUATION

<p>DELIVERY <i>Face-to-face, Distance learning, etc.</i></p>	Face-to-face	
<p>USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY <i>Use of ICT in teaching, laboratory education, communication with students</i></p>	Use of ICT in teaching, laboratory education, communication with students	
<p>TEACHING METHODS <i>The manner and methods of teaching are described in detail. Lectures, seminars, laboratory practice, fieldwork, study and analysis of bibliography, tutorials, placements, clinical practice, art workshop, interactive teaching, educational visits, project, essay writing, artistic creativity, etc.</i></p> <p><i>The student's study hours for each learning activity are given as well as the hours of non-directed study according to the principles of the ECTS</i></p>	<p>Activity</p>	<p>Semester workload</p>
	Lectures	13 hrs (0.52 ECTS)
	Teaching in Laboratory of informatics	26 hrs (1.04 ECTS)
	Personal study	43 hrs (1.72 ECTS)
	Assignments	40 hrs (1.6 ECTS)
	Final examination	3 hrs (0.12 ECTS)
	Course total	125 hrs (5 ECTS)
<p>STUDENT PERFORMANCE EVALUATION <i>Description of the evaluation procedure</i></p> <p><i>Language of evaluation, methods of evaluation, summative or conclusive, multiple choice questionnaires, short-answer questions, open-ended questions, problem solving, written work, essay/report, oral examination, public presentation, laboratory work, clinical examination of patient, art interpretation, other</i></p> <p><i>Specifically-defined evaluation criteria are given, and if and where they are accessible to students.</i></p>	<p>Language of evaluation: Greek. In case of ERASMUS students: Greek & English</p> <p>Method of evaluation:</p> <ul style="list-style-type: none"> - assignment - end of semester exam <p>Method of evaluation is given in class</p>	

(5) ATTACHED BIBLIOGRAPHY

Greek language

Γούτσος, Δ., Σηφιανού, Μ. & Θεοδωρακοπούλου, Α. (2005) *Η ελληνική ως ξένη γλώσσα*. Αθήνα: Πατάκης.

Γούτσος, Διονύσης (2015) *Εισαγωγή στη Γλωσσολογία Σωμάτων Κειμένων*. Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, Αποθετήριο "Κάλλιπος", www.kallipos.gr.

Μαρκόπουλος, Γ. Α. (2006) *Ζητήματα Υπολογιστικής Γλωσσολογίας*. Γ. Α. Μαρκόπουλος.

Μικρός, Γ. Κ. (2009) *Η Ποσοτική Ανάλυση της Κοινωνιογλωσσολογικής Ποικιλίας – Θεωρητικές και Μεθοδολογικές Προσεγγίσεις*. Αθήνα: Μεταίχμιο.

Μικρός, Γ. (2016) *Υπολογιστική Υφολογία*. Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα - Αποθετήριο "Κάλλιπος"

Σαριδάκης, Ι. (2010) *Σώματα Κειμένων και μετάφραση – Θεωρία και Εφαρμογές*. Αθήνα: Παπαζήση.

Παναρέτου, Ε. (2009) *Νομικός Λόγος – Γλώσσα & Δομή των Νόμων*. Αθήνα: Παπαζήση.

Τάντος, Α. (2016) *Υπολογιστική Γλωσσολογία. Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα* - Αποθετήριο "Κάλλιπος".

Φραντζή, Κ. (2012) *Εισαγωγή στην Επεξεργασία Σωμάτων Κειμένων*. Αθήνα: Ίων.

Foreign language

Baker, P. (2009) *Contemporary Corpus Linguistics* (Contemporary Studies in Linguistics). Continuum International Publishing Group Ltd.

Biber, D., Conrad, S. & Randi Reppen, R. (1998) *Corpus linguistics: investigating language structure and use*. Cambridge University Press.

Cheng, Winnie (2011) *Exploring Corpus Linguistics: Language in Action* (Routledge Introductions to Applied Linguistics). Routledge.

Dougherty, D. (1990) *sed & awk. Sebastopol, CA: O'Reilly & Associates, Inc.*

Friedl, J. E.F. (2006) *Mastering regular Expressions*. Sebastopol, CA: O'Reilly Media, Inc.

Hyland, Ken (2013) *Corpus Applications in Applied Linguistics*. Bloomsbury Academic.

Kennedy, G. (1998) *An Introduction to Corpus Linguistics*. Longman.

McEnery, A., Xiao, R. & Tono, Y. 2006. *Corpus-based Language Studies: an advanced resource book*. Routledge, 2006.

McEnery, A. (2011) *Corpus Linguistics: Method, Theory and Practice* (Cambridge Textbooks in Linguistics). Cambridge University Press.

Oakes, M. P. (1998) *Statistics for corpus linguistics*. Edinburgh University Press.

O’Keeffe, A. (2012) *The Routledge handbook of corpus linguistics* (Routledge Handbooks in Applied Linguistics). Routledge.

Stubbs M. (1996) *Text and Corpus Analysis: Computer Assisted Studies of Language and Institutions* (Language in Society). John Wiley & Sons.

Watt, A. (2005) *Beginning Regular Expressions*. Indianapolis, IN: Wiley Publishing, Inc.

Welsh, M., Dalheimer, M. K. & L. Kaufman (2011) *Ο Οδηγός του LINUX*. Αθήνα: Κλειδάριθμος.

Wolfgang, T, (2007) *Corpus Linguistics: A Short Introduction*. Continuum.

Related academic journals:

International Journal of Corpus Linguistics (IJCL)

Corpora

Corpus Linguistics and Linguistic Theory (CLLT)

Corpus Pragmatics

Language Resources and Evaluation