



COURSE UNIT (MODULE) DESCRIPTION

Course unit (module) title	Code
Corpus Linguistics and Phraseology/ Tekstynų lingvistika ir frazeologija	

Lecturer(s)	Department(s) where the course unit (module) is delivered
Coordinator: Assoc. Prof. Dr Rita Juknevičienė Other(s): Prof. Dr Jonė Grigaliūnienė	Faculty of Philology Department of English Philology

Study cycle	Type of the course unit (module)
2nd	Compulsory (MA Linguistics track) Optional (MA Literature, Culture, Linguistics; Media Discourse tracks)

Mode of delivery	Course unit delivery period	Language(s) of instruction
Face-to-face	Spring Term	English

Requirements for students	
Prerequisites: Courses in linguistics, lexicology, corpus linguistics	Additional requirements (if any): English (C1)

Number of ECTS credits allocated	Total student workload	Contact hours	Self-study hours
10	260	48	212

Purpose of the course unit (module): programme competences to be developed		
The purpose of the course unit is to deepen students' understanding of corpus-driven and corpus-based linguistic research with a particular focus on linguistic studies dealing with fixed multi-word expressions and formulaic language in English. The students will develop the following competences:		
I. generic competences: <ul style="list-style-type: none"> • cognitive; • social; • personal; 		
II. subject-specific competences: <ul style="list-style-type: none"> • knowledge of contemporary linguistic theories and methods and their application to the study of the English language; • research competence. 		
Learning outcomes of the course unit (module)	Teaching and learning methods	Assessment methods
Learning outcomes from the generic competences: <ul style="list-style-type: none"> - ability to find, analyse, synthesise and evaluate data needed for the studies; 	Interactive lectures, discussions; seminars; individual study; hands-on experience with corpora; mini-research projects	Continuous assessment (50%): a mini-research project for the seminars (25%) and a midterm test (25%).

<ul style="list-style-type: none"> - ability to apply the acquired knowledge in practice, identify problems and propose solutions; - ability to generate ideas and choose appropriate forms of expression; - ability seek new knowledge and skills and apply them for tasks in a new environment; - ability to recognise different cultural identities and attitudes, explore others' ideas despite cultural differences, express one's ideas in ways that are understandable to others; - ability to take responsibility for the quality of one's own and the team's activities; - ability to take initiative and carry out one's activities, rationally assess the strengths and weaknesses of one's work, reflect and be responsible for one's decisions and actions. <p>Learning outcomes from the subject-specific competences:</p> <ul style="list-style-type: none"> - ability to effectively communicate in English in writing and orally; analyse and produce academic texts in English following relevant conventions and genre requirements; - ability to make effective use of modern information and communication technologies, databases and resources (corpora, corpus processing tools, dictionaries); - ability to raise a relevant research question in linguistics and design a research study in the field of corpus-based or corpus-driven phraseology; - ability to select the theoretical framework(s), critically evaluate relevant research methods or approaches and empirical material; - ability to conduct a research study by adopting corpus-based and corpus driven approaches to data collection and processing, interpreting findings of the study, drawing conclusions and evaluating the results in the context of other studies; - ability to present the results of individual research projects in writing and orally to a variety of audiences and demonstrate the ability to participate in discussions. 		<p>Written examination (50%): a written assignment, e.g. a critical literature review on a given topic (ca 700 words).</p>
--	--	--

Content: breakdown of the topics	Contact hours	Self-study work: time and assignments
---	----------------------	--

	Lectures	Tutorials	Seminars	Exercises	Laboratory work	Internship/work placement	Contact hours	Self-study hours	Assignments
1. Introduction to the course unit: aims, structure, assessment.			1				1		
2. Introduction to Corpus Linguistics: past and present. Key terms and issues in Corpus Linguistics. Corpus-based approach vs corpus-driven approach to language. Best known corpora.	2						2	10	Selected chapters from Tognini-Bonelli (2004)
3. The status of Corpus Linguistics. Early Corpus Linguistics and the Chomskyan Revolution. Ongoing debate: Widdowson, Stubbs, Sinclair etc.	2						2	10	Selected chapters from Tognini-Bonelli (2004); Stubbs (2007); Widdowson (2000)
4. Why use corpora? How can corpora help in language learning, teaching and research?	2						2	10	Selected chapters from Tognini-Bonelli (2004)
5. Corpus creation. Corpora design criteria: representativeness, balance and sampling, size. Different types of corpora: sample, monitor, spoken, developmental, bilingual, multilingual, learner. WEB as a corpus.	4						4	16	Selected chapters from Gatto (2014); Sinclair (2004)
6. Corpora in Language research. Understanding of meaning in Corpus Linguistics. Extended Unit of Meaning. Lexical Grammar.	2						2	10	Selected chapters from Sinclair (1991); Sinclair (1996)
7. Making statistical claims: Raw frequency and normalized frequency. Tests of statistical significance. Tests for significant collocations.	2						2	10	Selected chapters from Brezina (2018)
8. Using corpora in the classroom. Learner corpora, learner corpora research.	2						2	10	Selected chapters from McEnery et al. (2006); O'Keeffe et al. (2007); Granger (1998)
9. Corpus-based and corpus-driven phraseological research: overview of the field and discussion of recent studies.			3				3	12	Reading: Gries (2008)
10. Types of formulaic expressions: collocations. Statistical and phraseological definitions of collocations. Association measures (MI, t-score, Log Likelihood).			6				6	26	Reading: a chapter from Juknevičienė (2017); tasks on data retrieval and analysis
11. Types of formulaic expressions: lexical bundles/n-grams/recurrent sequences. Retrieval and analysis. Structural and functional classification.			6				6	26	Biber et al. (2014); tasks on data retrieval and analysis
12. Research tools and corpus analysis software (AntConc, #LancsBox, SketchEngine, Wordsmith Tools).			6				6	26	Tasks on data retrieval and analysis
13. Applications of corpus-driven research in translation, lexicography and language teaching.			6				6	26	Tasks and presentations of individual projects
12. Course summary and assessment.			4				4	20	
Total	16		32				48	212	

Assessment strategy	Weight,%	Deadline	Assessment criteria
<i>Continuous assessment</i>			
A mid-term test on theory	25	mid-term	correctness of the answers to test questions;

		test	language accuracy of the answers; more specific requirements will be presented during the course
A presentation of an individual research project	25	end of term	content (clarity and appropriateness of the information presented; conformity to presentation requirements to be specified during the seminars)
Examination			
A written assignment	50	examination date	A critical literature review on a given topic (ca 700 words) assessed for the content, text structure and language use. More specific requirements will be presented during the course.

Author	Year of publication	Title	Issue of a periodical or vol. of a publication	Publishing place and house or web link
Compulsory reading				
Biber, D., S. Conrad & V. Cortes	2004	<i>If you look at...: Lexical bundles in university teaching and textbooks</i>	25/3: 371-405	<i>Applied Linguistics</i> (Oxford Journals Online)
Brezina, V.	2018	<i>Statistics in Corpus Linguistics</i>		Cambridge University Press
Gatto, M.	2014	<i>Web as Corpus. Theory and Practice</i>		London/New Delhi/ New York/ Sydney: Bloomsbury
Granger, S. & F. Meunier (eds.)	2008	<i>Phraseology: An interdisciplinary perspective</i>		Amsterdam: John Benjamins
Gries, S. Th	2008	Phraseology and linguistic theory: A brief survey.	3-26	Amsterdam: J. Benjamins Publishing Company. (In S. Granger & F. Meunier (Eds) <i>Phraseology. An Interdisciplinary Perspective</i>)
Juknevičienė, R.	2017	<i>English Phraseology and Corpora: An introduction to corpus-based and corpus-driven phraseology</i>		Vilnius: Vilniaus universiteto leidykla
McEnery, M. , R. Xiao & Y. Tono	2006	<i>Corpus-Based Language Studies</i>		London/New York: Routledge
Sinclair, J.	1991	<i>Corpus, Concordance, Collocation</i>		Oxford: Oxford University Press
Sinclair, J. (ed.)	1996	<i>Looking Up</i>		London/Glasgow: Harper Collins Publishers
Sinclair, J.	1996	The Search for Units of Meaning.	IX (1): 76-106	<i>TEXTUS</i>
Sinclair, J. M	2004	<i>Trust the Text: Language, Corpus and Discourse</i>		London: Routledge
Stubbs, M.	2007	Notes on the History of Corpus Linguistics and Empirical Semantics	317-329	Joensuu: Joensuun Yliopisto. (In M. Nenones & S. Niemi (Eds) <i>Collocations and Idioms</i>)
Tognini-Bonelli, E.	2004	<i>Corpus Linguistics at Work</i>		Amsterdam: John Benjamins
Wray, A.	2008	Formulaic Language: Pushing the Boundaries	57/2: 182-197	<i>ELT Journal</i> (Oxford Journals Online)
Optional reading				
Aijmer, K. (ed).	2009	<i>Corpora and Language Teaching</i>		Amsterdam/Philadelphia: John Benjamins
Conklin, K. & N. Schmitt	2008	Formulaic sequences: are they processed more quickly than nonformulaic language by native and nonnative speakers?	29/1: 72-89	<i>Applied Linguistics</i> (Oxford Journals Online)
Cortes, V.	2015	Situating lexical bundles in the	197-216	Amsterdam: J. Benjamins

		formulaic language spectrum		Publishing Company. (In V. Cortes & E. Csomay (Eds) <i>Corpus-based Research in Applied Linguistics: Studies in Honor of Doug Biber</i>)
Granger, S. (Ed.)	1998	<i>Learner English on Computer</i>		London: Longman
Hyland, K	2008	As can be seen: Lexical bundles and disciplinary variation	27: 4-10	<i>English for Specific Purposes</i> (Elsevier)
Meunier, F. & S. Granger (Eds.)	2008	<i>Phraseology in Foreign Language Learning and Teaching</i>		Amsterdam: John Benjamins Publishing Company
McEnery, T. & A. Wilson	2001	<i>Corpus Linguistics</i>		Edinburgh University Press
Nesselhauf, N.	2005	<i>Collocations in a Learner Corpus</i>		Amsterdam: John Benjamins
O’Keeffe, A., McCarthy, M. & R. Carter	2007	<i>From Corpus to Classroom</i>		Cambridge: Cambridge University Press
Philip, G.	2011	<i>Colouring Meaning. Collocation and Connotation in Figurative Language</i>		Amsterdam/Philadelphia: John Benjamins.
Sinclair, J	2004	<i>Trust the Text</i>		London: Routledge
Stubbs, M.	2001	Texts, Corpora and Problems of Interpretation: A Response to Widdowson	22(2):149- 172	<i>Applied Linguistics</i>
Vilkaite, L.	2016	Formulaic language is not all the same: comparing the frequency of idiomatic phrases, collocations, lexical bundles, and phrasal verbs	8: 28-54	<i>Taikomoji kalbotyra</i>
Widdowson, H. G.	2000	On the Limitations of Linguistics Applied	21(1):3-25	<i>Applied Linguistics</i>
Corpus tools				
Anthony, L.	2018	AntConc [Computer Software]	v. 3.5.7	Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/software
Brezina, V., T. McEnery, & S. Wattam	2015	#LancsBox [Computer Software], presented in ‘A new perspective on collocation networks. Collocations in context’	20(2): 139– 173	International Journal of Corpus Linguistics Available from http://corpora.lancs.ac.uk/lancsbox/
Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, V. Suchomel	2014	The Sketch Engine: ten years on	1: 7-36	Lexicography Available from: http://www.sketchengine.eu
Scott, M.	2010	WordSmith Tools (v. 5).	v. 5	Oxford: Oxford University Press

Updated 2023-11-16