COURSE UNIT (MODULE) DESCRIPTION

Course unit (module) title	Code
Introduction to Corpus Linguistics	

Lecturer(s)	Department(s) where the course unit (module) is delivered
Coordinator: dr. Jolanta Šinkūnienė **Other(s):**	Department of English Philology Faculty of Philology

Study cycle	Type of the course unit (module)
BA programme, semester 4	Optional

Mode of delivery	Period when the course unit (module) is delivered	Language(s) of instruction
Seminars	Spring term	English

Requirements for students	
Prerequisites: English (C1)	Additional requirements (if any):

Course (module) volume in credits	Total student's workload	Contact hours	Self-study hours
5 ECTS	150	48	102

Purpose of the course unit (module): programme competences to be developed

The main objectives of the course are to bring the students up to date with the latest developments in corpus-based language studies, and to show the practical applicability of the corpus in professional contexts. One of the fields of English for Specific Purposes – the field of tourism - will be addressed as a case study to illustrate the benefits of corpus linguistics in specialised language acquisition. Other objectives of the course are to set the scene for a corpus-based language studies by addressing some of the key issues of corpus linguistics, to give a hands-on experience with the corpora and its annotation: to show that the use of corpora can be an integral part of courses in grammar, spoken and written proficiency, to draw the students' attention to complex patterns and phraseology rather than regularities and to provide additional support that language learning is a complex process involving hypothesis formation and testing. The use of corpora allows the students to become researchers: the students are encouraged to explore research questions using corpus data, tools and techniques. The course unit develops the following competences:

Generic competences:

1. Responsibility: the ability to set goals and make plans, and take responsibility for them
- will be able to set goals, choose and use resources necessary for the completion of a task, plan their time and follow deadlines
- will be able to take responsibility for their work / study results and learn from mistakes

2. Problem solving: the ability to solve problems by relying on analytical, critical, and creative thinking
- will be able to identify problems and challenges in their own and related fields
- will be able to identify problems by finding, analysing, and critically assessing relevant information, generate

new ideas, choose the most optimal solutions

3. Openness to change: the ability to understand the necessity of change and the intention to constantly improve oneself
- will be familiar not only with the changes taking place in their field of interest, but also their causes, challenges, opportunities
- will be open to new ideas, strive to change, and be creative and innovative
- will be able to evaluate the quality of their actions and achievements and will strive to acquire the competencies necessary for future change

Subject-specific competences:

4. Essential knowledge and skills in linguistics: perception of language as a phenomenon and perception of linguistics as a scientific discipline
- will know, understand and be able to define linguistics as a scientific discipline and to properly use and interpret the basic concepts and terms of linguistics
- will acquire knowledge of the main branches and methods of linguistics

5. Understanding and analysis of the English language system at various levels: phonetics, morphology, syntax, semantics, etc.
- will be able to describe, analyse and interpret English-language phenomena at various levels (phonetic, morphological, syntactic, semantic, pragmatic, language development, etc.) and compare them with Lithuanian or another language using appropriate terminology and methods
- will be able to effectively use and analyse English for specific purposes (science, politics, business, law, etc.)

6. Competence of a researcher of language and literature: the ability to independently conduct linguistic, literary or interdisciplinary research applying the acquired philological knowledge and skills in practice
- will be able to identify a linguistic, literary, cultural or interdisciplinary problem and choose appropriate empirical material, research methods and bibliographic sources
- will be able to collect, analyse, systematise and critically evaluate material from various sources and adequately apply philological research methods to investigate the selected research problem
- will be able to describe the results of the research in a scientific text, provide conclusions or recommendations, and present and defend them publicly

7. Ability to apply philological knowledge and skills in practice within and outside the University
- will be able to apply the acquired knowledge and skills of English philology during communication and cooperation in an intercultural (and multilingual) environment, creating, translating and editing various texts, teaching English as a foreign language, etc.
- will be able to use modern information technologies, data resources and research resources to conduct linguistic and literary analysis of English texts and present the results of analysis and/or interpretation to the public in order to contribute to the practical applicability of English philology studies (corpora, dictionaries, term banks and glossaries, thematic websites, literature maps, etc.)

Learning outcomes of the course unit (module)	Teaching and learning methods	Assessment methods
On successful completion of this course unit, the students should have: • an understanding of the foundations of Corpus Linguistics, key terminology; • an understanding of the construction of different corpora; • a broad sense of application of different corpora; The students should be able to: • extract, present and discuss data from the corpora; • use corpus data to answer various questions about vocabulary, syntax, etc. • carry out a research project using corpora.	• Seminars, group discussions, individual work (reading and analysis of research articles). • Research tasks, assignments. • Hands on experience of corpora. • Research project work.	Final mark will be based on the following assignments: • **home assignments & participation in class discussion** (20 %): • **research in progress presentations** (20 %) • **term paper**: (60 %): • structure + content (40%) • language use + academic conventions (20%) **The focus is on empirical data** (empirical or textual

| | | | | | | | analysis) data collection procedure, methods of analysis, introduction of term paper structure, clarity and precision. |

| Content: breakdown of the topics | Contact hours | | | | | | Self-study work: time and assignments | |
	Lectures	Tutorials	Seminars	Exercises	Laboratory work Internship/work	Contact hours	Self-study hours	Assignments
Introduction to the course unit: aims, structure, assessment scheme.			1			**1**		
Introduction to Corpus Linguistics: past and present. Key terms. Best known corpora. Different types of corpora. Why use corpora? How can corpora help in language learning, teaching, and research?			2			**2**	**4**	Aarts (2002), Gries (2010).
Corpus creation. Corpora design criteria: spoken, written language, representativeness, size. Statistical tests.			2			**2**	**4**	McEnery & Wilson (2001: 76-98)
The use of corpora in the English for Specific Purposes field. Tourism industry as an example.			2			**2**	**4**	Heller et al (2014: 425-458)
Corpora in language research: Meaning in Corpus Linguistics. Extended unit of meaning.			4			**4**	**8**	Sinclair (2004: 24-28); Sinclair (2000)
Historical corpora. Annotation and TEI/XML tagging using historical corpora as the basis.			4			**4**	**4**	Yáñez-Bouza (2024: 31-62). Denison et al. (2024: 109-128).
The use of corpora in language studies. Contrastive analysis. Examples of studies using parallel corpora.			4			**4**	**10**	McEnery & Wilson (2001: 103-118). Analysis of sample research papers using parallel corpora from Le Bruin et al. (2022); Van Olmen & Šinkūnienė (2021)
The use of corpora in language studies. Comparable specialised corpora in greater detail. Examples of studies using comparable specialized corpora. Preliminary discussion of paper topics.			4			**4**	**10**	McEnery & Wilson (2001: 119-130). Analysis of sample research papers using comparable corpora from Dirdal et al. (2022); Paquot et al. (2022); Šinkūnienė (2018); Van Olmen & Šinkūnienė (2021)
Research paper. Structure of a research paper. Writing a literature review. Referencing.			2			**2**	**8**	Swales (2004)

Topic									Self-study
Analysis of sample introductions.			2				2	10	Analysis of sample research paper introductions. Work on term paper.
Reporting data, methods, findings. Conclusions, implications.			2				2	12	Swales (2004: 219-239) Work on term paper.
Presentation of mini-proposals, peer-review of introductions, data & methods sections, individual feedback.			12				12	12	Peer-review of introductions. Work on term paper.
Oral presentations of research in progress			6				6	16	Preparation for the presentations. Work on term paper.
Summary and course evaluation			1				1		
Total			**48**				**48**	**102**	

Assessment strategy	Weight, %	Deadline	Assessment criteria
Cumulative assessment			
Attendance			The attendance of seminars is compulsory. Students who have missed more than 35% of the seminars with no justifiable reason will not complete the course.
Continuous assessment	20%	throughout the term	Active participation in class discussions, performance of practical tasks.
Presentations	20%	in May	Content, presentation skills, language accuracy and fluency.
Term paper	60%	June 10, 2026	Structure + content (40%) Language use + academic conventions (20%)

Author	Year of publication	Title	Issue of a periodical or volume of a publication	Publishing place and house or web link
Compulsary reading				
Aarts, J.	2002	Does corpus linguistics exist? Some old and new issues	In L. E. Breivik and A. Hasselgren (eds.). *From the COLT's mouth... and others': Language corpora studies in honour of Anna-Brita Stenström*, 1–19.	Amsterdam: Rodopi.
Gries, St.	2010	What is Corpus Linguistics?	*Language and Linguistics Compass* 3 (2009): 1–17, 10	
Denison, D., Yáñez-Bouza, N. & Oudesluijs, T.	2024	Editing The Mary Hamilton Papers (c.1740 – c.1850).	In Javier Calle-Martín & Jesús Romero-Barranco (eds.), Corpora and language change in Late Modern English, 109-28.	Lausanne, etc.: Peter Lang.

Dirdal, H. et al.	2022	Design and construction of the Tracking Written Learner Language (TRAWL) Corpus: A longitudinal and multilingual young learner corpus	*Nordic Journal of Language Teaching and Learning* (NJLTL). 10 (2), 115-135.	
Heller, M. et al.	2014	Introduction: Sociolinguistics and tourism– mobilities, markets, multilingualism.	Journal of Sociolinguistics 18/4, 2014: 425–458.	
Le Bruin, et al.	2022	Parallel Corpus Research and Target Language Representativeness: The Contrastive, Typological, and *Translation Mining* Traditions	*Languages*, 7(3), 176	
McEnery, T. & A. Wilson	2001	*Corpus Linguistics*		Edinburgh University Press
Paquot, M. et al.	2022	The Varieties of English for Specific Purposes dAtabase (VESPA): Towards a multi-L1 and multi-register learner corpus of disciplinary writing	*Research in Corpus Linguistics.* 10 (2), 1-15	
Sinclair, J.	2004	*Trust the Text: Language, Corpus and Discourse*		London: Routledge
Šinkūnienė, J.	2018	The power of English: *I* and *we* in Lithuanian, Lithuanian English and British English research writing.	In Pilar Mur Dueñas, & Jolanta Šinkūnienė (Eds.). *Intercultural perspectives on research writing.*	Amsterdam: John Benjamins. 59-79.
Swales, J.	2004	*Research Genres: Explorations and Applications*		Cambridge: CUP
Van Olmen, D. & J. Šinkūnienė	2021	*Pragmatic Markers and Peripheries*		Amsterdam: Benjamins
Yáñez-Bouza, N.	2024	'My dearest Clara … my dear friend' – Personal names and direct address in Mary Hamilton's correspondence	Journal of Historical Sociolinguistics 10.1, 31-62. DOI: 10.1515/jhsl-2022-0038.	

Updated: November 23, 2025