University of Gdańsk

Subject card

Subject name and code	Statistics for linguists, PG_00188668		
Field of study	English Studies		
Date of commencement of studies	October 2026	Academic year of realisation of subject	2026/2027
Education level	Master's studies	Subject group	Obligatory subject group in the field of study Optional subject group Subject group related to scientific research in the field of study
Mode of study	full-time studies	Mode of delivery	at the university
Year of study	1	Language of instruction	English
Semester of study	2	ECTS credits	1.0
Learning profile	academic	Assessment form	credit
Conducting unit	Division of Corpus Linguistics and Glottodidactics -> Institute of English and American Studies -> Faculty of Languages -> Rector		
Name and surname of lecturer (lecturers)	Subject supervisor	dr hab. Anna Bączkowska, prof. UG	
	Teachers		

Lesson types

Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
Number of study hours	15.0	0.0	0.0	0.0	0.0	15

E-learning hours included: 0.0

Learning activity and number of study hours

Learning activity	Participation in didactic classes included in study plan	Participation in consultation hours	Self-study	SUM
Number of study hours	15	1.0	9.0	25

Subject objectives

The aim of the module is to introduce the fundamentals of statistics and to familiarize students with how select statistical tests can be applied to an analysis of language, in particular in line with Corpus Linguistics and data/text mining.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[FAMU2_U04] Can select and apply information and communication techniques (ICT) during the acquisition and processing of information for research and professional purposes within the framework of English Studies.	The student can choose and apply simple statistical methods for data analysis in the field of linguistic research	[SU4] test/exam - oral or written
[FAMU2_W12] Knows and understands the main development trends in literature, in particular with regard to research into English-language literature.	The student understands methods of using ICT in statistics in the field of linguistic research.	[SW4] test/exam - oral or written
[FAMU2_U05] Can use and present orally and in writing knowledge in the field of auxiliary and related sciences, which is a context for English linguistic and literary studies and supports the professional work of an English philologist.	The student can use and present statistical knowledge applied in the field of linguistic	[SU2] presentation/project/paper/ report [SU4] test/exam - oral or written
[FAMU2_K02] Is prepared to recognise the importance of knowledge and skills in English Studies in solving cognitive and practical problems and to seek the advice of a supervisor in their chosen place of work in the event of difficulty in solving problems on their own.	The student is ready to recognize the significance of knowledge and skills in the field of English studies for solving cognitive and practical problems, and to seek the opinion of a scientific supervisor or mentor at the chosen workplace in case of difficulties with independent problem-solving	[SK8] observation of student's independent or team work
[FAMU2_K01] Is ready to critically evaluate the extent of their knowledge and skills, in particular in the field of English-language linguistics and literature and English.	The student is ready to critically assess their own knowledge in the field of computational data analysis and statistics, they understand the role of this knowledge in cognitive and practical problem solving occurring in social life.	[SK1] oral statement/conversation/ discussion

Subject contents

The students will acquire basic information about general and corpus statistics, inter alia: types of variables, hypothesis testing, type I and type II errors, central tendency, types of data, (non-)parametric tests, chi square test, t test, ANOVA, MANOVA, correlation tests, linear and logistic regression, tests for non-parametric data, variance, degrees of freedom, post-hoc test, Cronbach's alpha, significance level, effect size, cluster analysis; collocation association measures, dispersion measures, distance measures, lexical density and variation, lexical, stylistic and syntactic parameters of texts, etc.

Prerequisites and co-requisites

Level of English at B2 or higher CEFR level

Basic course in linguistics

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
observation	100.0%	10.0%
final and mid-semester tests 60.0% 70.0%	60.0%	70.0%
presentation	60.0%	10.0%
discussion	100.0%	10.0%

Recommended reading

Basic literature

Brezina, V. 2018. Statistics for Corpus Linguistics. Cambridge: CUP.

Woods, A., Fletcher, P., Hughes A., 1989. Statistics in Language Studies. Cambridge: CUP.

	Supplementary literature	Bączkowska, A. 2019. Readability and lexical sophistication of colon cancer websites a corpus-assisted assessment of online educational materials for patients, Forum Filologiczne Ateneum 1(7), str. 925. ; https://doi.org/10.36575/2353-2912/1(7)2019.009 https://www.ateneum.edu.pl/assets/Uploads/003-baczkowska.pdf Bedyńska, S., Cupryańska, M. 2013. Statystyczny drogowskaz 1. Warszawa: SWPS. Brezina, V. 2018. Statistics for Corpus Linguistics. Cambridge: CUP. Glenberg, A., Andrzejewski, M. 2007. Learning from Data: An Introduction to Statistical Reasoning. London: Routledge. Tatarzycki, P. 2007. Statystyka po ludzku. Gliwice: Złote Myśli. Woods, A., Fletcher, P., Hughes A., 1989. Statistics in Language Studies. Cambridge: CUP.
	eResources addresses	
Example issues/ example questions/ tasks being completed		
Work placement	Not applicable	

Document generated electronically. Does not require a seal or signature.