

LT5423: QUANTITATIVE METHODS FOR LINGUISTIC DATA ANALYSIS

Effective Term

Semester A 2026/27

Part I Course Overview

Course Title

Quantitative Methods for Linguistic Data Analysis

Subject Code

LT - Linguistics and Translation

Course Number

5423

Academic Unit

Linguistics and Translation (LT)

College/School

College of Liberal Arts and Social Sciences (CH)

Course Duration

One Semester

Credit Units

3

Level

P5, P6 - Postgraduate Degree

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

Nil

Precursors

Nil

Equivalent Courses

Nil

Exclusive Courses

Nil

Part II Course Details

Abstract

This is an introductory course which highlights some of the basic and yet most widely used statistical measures for linguistic analysis. In particular, this course will introduce quantitative notions about corpus data and explain how such data can be statistically described, analysed and tested. The students will be required to be able to apply the quantitative methods to perform statistical analysis of linguistic data.

Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify the importance of a quantitative approach the study of human language.	20	x	x	x
2	Explain the quantitative underpinnings of language on the one hand and of the multifaceted linguistic form on the other.	20	x	x	x
3	Reflect critically on the complex relations between language and language use via discovering and identifying the quantifiable aspects of language.	20	x	x	x
4	Conduct and critique quantitative analysis of linguistic data.	20	x	x	x
5	Hypothesize and formulate generalizations about human language based on quantitative observations and analysis.	20	x	x	x

A1: Attitude

Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability

Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments

Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

LTAs		Brief Description	CILO No.	Hours/week (if applicable)
1	Lectures	Students will engage in presentation and explanation of basic theories and methods	1, 2, 3, 4, 5	
2	Class Activities	Students will engage in exercises, group discussions and presentations	1, 2, 3, 4, 5	

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks ("- for nil entry)	Allow Use of GenAI?
1	One presentation	1, 2, 3, 4, 5	10	Students can use GenAI tools to brainstorm, gather information, solve problems, format references, and edit language. Students should give proper acknowledgment to all the content from GenAI tools that they incorporated into the assignment and keep a journal of their GenAI conversations. Students should not present texts generated or translated by GenAI tools as their own original work.	Yes
2	One assignment	1, 2, 3, 4, 5	30	Students can use GenAI tools to brainstorm, gather information, solve problems, format references, and edit language. Students should give proper acknowledgment to all the content from GenAI tools that they incorporated into the assignment and keep a journal of their GenAI conversations. Students should not present texts generated or translated by GenAI tools as their own original work.	Yes
3	One quiz	1, 2, 3, 4, 5	10	-	No

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Assessment Rubrics (AR)

Assessment Task

Oral Presentation (10%)

Criterion

Demonstration of understanding basic concepts Demonstration of ability to apply basic concepts

Excellent (A+, A, A-)

- 1) comprehensive understanding of the course contents and being able to apply the contents;
- 2) Demonstrating accurate and critique analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Good (B+, B, B-)

- 1) comprehensive understanding of the course contents;
- 2) demonstrating accurate analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Fair (C+, C, C-)

- 1) moderate or limited understanding on topics and contents of the course;
- 2) demonstrating basic ability to analyze linguistic instances;
- 3) appropriate format in the writing, but writing content merely shows some understanding of the differences between academic and non-academic style of writing and put that understanding to practice.

Marginal (D)

- 1) little understanding on topics and contents of the course;
- 2) insufficient knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Failure (F)

- 1) no understanding on topics and contents of the course;
- 2) incorrect knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Assessment Task

Assignment (30%)

Criterion

Demonstration of understanding basic concepts Demonstration of ability to apply basic concepts

Excellent (A+, A, A-)

- 1) comprehensive understanding of the course contents and being able to apply the contents;
- 2) Demonstrating accurate and critique analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Good (B+, B, B-)

- 1) comprehensive understanding of the course contents;

- 2) demonstrating accurate analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Fair (C+, C, C-)

- 1) moderate or limited understanding on topics and contents of the course;
- 2) demonstrating basic ability to analyze linguistic instances;
- 3) appropriate format in the writing, but writing content merely shows some understanding of the differences between academic and non-academic style of writing and put that understanding to practice.

Marginal (D)

- 1) little understanding on topics and contents of the course;
- 2) insufficient knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Failure (F)

- 1) no understanding on topics and contents of the course;
- 2) incorrect knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Assessment Task

Quiz (10%)

Criterion

Demonstration of understanding basic concepts Demonstration of ability to apply basic concepts

Excellent (A+, A, A-)

- 1) comprehensive understanding of the course contents and being able to apply the contents;
- 2) Demonstrating accurate and critique analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Good (B+, B, B-)

- 1) comprehensive understanding of the course contents;
- 2) demonstrating accurate analysis on linguistic instances;
- 3) writing format is appropriate in an academic style and standard.

Fair (C+, C, C-)

- 1) moderate or limited understanding on topics and contents of the course;
- 2) demonstrating basic ability to analyze linguistic instances;
- 3) appropriate format in the writing, but writing content merely shows some understanding of the differences between academic and non-academic style of writing and put that understanding to practice.

Marginal (D)

- 1) little understanding on topics and contents of the course;
- 2) insufficient knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Failure (F)

- 1) no understanding on topics and contents of the course;
- 2) incorrect knowledge on course contents;
- 3) inappropriate writing style and format for academic paper.

Assessment Task

Examination (50%)

Criterion

Same as Coursework but under time constraint

Excellent (A+, A, A-)

- 1) accurate and clear answers to question;
- 2) demonstrating integrative and comprehensive understanding from question answers;
- 3) strategic use of ideas and information from referenced articles, course handouts and other sources;
- 4) argumentation strategies and contents are logic, excellent and substantiated;
- 5) excellent command of the English language.

Good (B+, B, B-)

- 1) accurate and clear answers to question;
- 2) accurate and clear understanding on specific terms taught in the course;
- 3) good use of ideas and information from referenced articles, course handouts and other sources;
- 4) accurate and precise analysis on behavioral data of languages;
- 5) argumentation strategies and contents are logic, good and substantiated;
- 6) good command of the English language.

Fair (C+, C, C-)

- 1) accurate and yet insufficient answers to questions;
- 2) appropriate and yet limited understanding on specific terms taught in the course;
- 3) good summary of ideas and information from referenced articles, course handouts and other sources;
- 4) accurate and yet unspecified analysis on the behavioral data of language;
- 5) argumentation strategies and contents are logic but merely appropriate;
- 6) adequate command of the English language.

Marginal (D)

- 1) very limited critical answers to questions;
- 2) little understanding on specific terms taught in the course;
- 3) unclear answers due to mere collection of ideas and information from referenced articles, course handouts and other sources;
- 4) inadequate analysis behavioral data of language;
- 5) poor argumentations;
- 6) poor command of English.

Failure (F)

- 1) inappropriate critical answers to questions;
- 2) insufficient understanding on specific terms taught in the course;
- 3) incorrect answers;
- 4) incorrect analysis of behavioral data of language;
- 5) inappropriate argumentations;
- 6) poor or little command of English.

Part III Other Information

Keyword Syllabus

types, tokens and word frequencies; concordance, collocations, and word co-occurrences; variance and central tendency; group scores, correlations; corpus annotation, inter-subject agreement; clustering methods for finding structure in data.

Reading List

Compulsory Readings

Title	
1	Oakes, Michael. 1998. Statistics for Corpus Linguistics. Edinburgh: Edinburgh University Press.
2	Heiman, Gary W. 2014. Basic Statistics for the Behavioral Sciences. Cengage Learning.

Additional Readings

Title	
1	Baayen R.H. 2008. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge: Cambridge University Press.
2	Gries Stefan. 2009. Quantitative Corpus Linguistics with R: A Practical Introduction. London: Routledge.
3	Butler, Christopher. 1985. Statistics in Linguistics. Oxford: Blackwell.