

LT3220: CORPUS LINGUISTICS

Effective Term

Semester B 2025/26

Part I Course Overview

Course Title

Corpus Linguistics

Subject Code

LT - Linguistics and Translation

Course Number

3220

Academic Unit

Linguistics and Translation (LT)

College/School

College of Liberal Arts and Social Sciences (CH)

Course Duration

One Semester

Credit Units

3

Level

B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

Nil

Precursors

Nil

Equivalent Courses

CTL3220 Corpus Linguistics

Exclusive Courses

Nil

Part II Course Details

Abstract

The main aim of this course is to introduce the basic concepts, principles and methodologies in corpus linguistics, with an emphasis on the construction and use of organic collections of texts for linguistic knowledge and practical applications, particularly in relation to the basic understanding of large language models.

Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Describe corpus linguistics and its history of development.		x	x	x
2	Discuss major uses of linguistic corpora for research and application.		x	x	x
3	Perform some of the basic linguistic computations.		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	Lecturing	Students will engage in formal lectures to gain knowledge about basic theories and practices.	1, 2, 3	3 hours/week
2	Peer discussion	Students will participate in groups to consolidate their learning as they produce four reports (in the form of PPT slides) and will actively engage as audience members during peers' presentations in order to expand their own knowledge.	1, 2, 3	

Assessment Tasks / Activities (ATs)

ATs	CILO No.	Weighting (%)	Remarks ("- " for nil entry)	Allow Use of GenAI?	
1	Group Projects: Independent group-based discussions and reports on assigned topics	1, 2, 3	50	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

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Assessment Rubrics (AR)**Assessment Task**

1. Group projects

Criterion

Capacity for self-directed learning and discovery to understand the principles of corpus linguistics

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Insufficient

Assessment Task

2. Examination

Criterion

Mastery of concepts and techniques, ability to describe and analyze linguistic features.

Excellent (A+, A, A-)

Excellent

Good (B+, B, B-)

Good

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Insufficient

Part III Other Information**Keyword Syllabus**

History of corpus linguistics

Design and creation of corpora

Lexical analysis with raw corpora

Grammatical analysis with tagged corpora

Syntactic analysis with parsed corpora

Software tools (concordancer and SPSS)

Reading List**Compulsory Readings**

Title	
1	Biber, Douglas, Susan Conrad and Randi Reppen. Corpus Linguistics – Investigating Language Structure and Use. Cambridge: Cambridge University Press, 1998.
2	Butler, Christopher. Statistics in Linguistics. Oxford: Blackwell, 1985.
3	Fang, Alex Chengyu. English Corpora and Automated Grammatical Analysis. Beijing: The Commercial Press, 2007.
4	Fang, Alex Chengyu and Jing Cao. 2015. Text Genres and Registers: The Computation of Linguistic Features. Berlin and Heidelberg: Springer.

5	Garside, Roger, Geoffrey Leech and Geoffrey Simpson (eds.). The Computational Analysis of English: A Corpus-based Approach. London: Longman, 1987.
6	McEnery, Tony and Adrew Wilson. Corpus Linguistics. Edinburgh: Edinburgh University Press, 1996.

Additional Readings

Title	
1	Literary and Linguistic Computing
2	Language Resource and Evaluation
3	International Journal of Corpus Linguistics