

LT5411: COMPUTATIONAL LINGUISTICS

Effective Term

Semester A 2025/26

Part I Course Overview

Course Title

Computational Linguistics

Subject Code

LT - Linguistics and Translation

Course Number

5411

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

CTL5411 Computational Linguistics

Exclusive Courses

Nil

Part II Course Details

Abstract

This course aims at introducing students with various academic backgrounds to the quantitative aspects of language and the basic concepts of the application of computational techniques in language processing. The course will introduce quantitative

notions about corpus data and explain how such data can be statistically described, analysed and tested. It will also cover topics selected from a wide range of issues and challenges of the field, including but not limited to multilingual computing, using computers in linguistic studies, natural language processing and the use of large corpora in natural language applications. This course will also provide students with a foundation for further research in relevant areas.

Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify key computational linguistic concepts and issues in major areas of study in the field and recognize the importance of a quantitative approach to the study of human language.	30	x	x	x
2	Explain the theoretical basis underlying major computational approaches to handling these issues, conduct and critique quantitative analysis of linguistic data.	40	x	x	x
3	Review the development of one or more major area of research in computational linguistics and devise innovative solutions to the critical issues based on quantitative observations and analysis.	30	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 learn about major issues and concepts in computational linguistics and introduce different approaches to their solution. Interaction between instructor and students is expected.	1, 2	
2	Discussion	Students will discuss the theories and algorithms, to reinforce concepts learned during lectures.	1, 2	

3	Readings	Students will read relevant book chapters, articles and other kinds of supplementary materials.	1, 2, 3	
4	Programming Exercises	Students will engage in hands-on practical exercises involving the use of computational tools for various language processing tasks.	1, 2, 3	

Assessment Tasks / Activities (ATs)

ATs		CILO No.	Weighting (%)	Remarks ("- for nil entry)	Allow Use of GenAI?
1	Assignments on theories and applications in computational linguistics	1, 2, 3	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
2	Quiz(zes) on concepts and issues in computational linguistics, and theoretical basis underlying some computational approaches to their solutions	1, 2, 3	30	-	No

Continuous Assessment (%)

60

Examination (%)

40

Examination Duration (Hours)

2

Assessment Rubrics (AR)

Assessment Task

Assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

Ability to demonstrate competence in theories and applications in computational linguistics

Excellent

(A+, A, A-) High

Good

(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

Quiz(zes) (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

Ability to demonstrate competence in concepts and issues in computational linguistics

Excellent

(A+, A, A-) High

Good

(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

Examination (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

Ability to demonstrate knowledge on theory and practice of computational linguistics

Excellent

(A+, A, A-) Demonstrate excellent knowledge on theory and practice of computational linguistics

Good

(B+, B, B-) Demonstrate significant knowledge on theory and practice of computational linguistics

Fair

(C+, C, C-) Demonstrate fair knowledge on theory and practice of computational linguistics

Marginal

(D) Demonstrate basic knowledge on theory and practice of computational linguistics

Failure

(F) Not even demonstrating marginal knowledge on theory and practice of computational linguistics

Assessment Task

Assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Ability to demonstrate competence in theories and applications in computational linguistics

Excellent

(A+, A, A-) High

Good

(B+, B) Significant

Marginal

(B-, C+, C) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

Quiz(zes) (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Ability to demonstrate competence in concepts and issues in computational linguistics

Excellent

(A+, A, A-) High

Good

(B+, B) Significant

Marginal

(B-, C+, C) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

Examination (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

Ability to demonstrate knowledge on theory and practice of computational linguistics

Excellent

(A+, A, A-) Demonstrate excellent knowledge on theory and practice of computational linguistics

Good

(B+, B) Demonstrate significant knowledge on theory and practice of computational linguistics

Marginal

(B-, C+, C) Demonstrate basic knowledge on theory and practice of computational linguistics

Failure

(F) Not even demonstrating marginal knowledge on theory and practice of computational linguistics

Part III Other Information

Keyword Syllabus

Natural language processing: Tokenisation, Morphological analysis, Part-of-speech tagging, Context-free rules, Parsing, Semantic representation, Disambiguation, Rule-based methods, Corpus-based methods, Statistical methods

Linguistic computing: Frequency counts, Quantitative methods in linguistic studies, Linguistic corpora, Text markup, Corpus annotation, Concordance

Natural language applications: Machine translation, Information retrieval, Information extraction, Natural language generation

Multilingual computing: Character encoding, Input and display, Internationalization and localization

Reading List**Compulsory Readings**

Title	
1	Manning, C.D. and Schütze H. (1999). Foundations of Statistical Natural Language Processing. Cambridge, MA: The MIT Press.
2	Jurafsky D. and Martin, J.H. (2009). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (2nd Edition). Upper Saddle River, NJ: Pearson Prentice Hall.
3	Allen, J. (1995). Natural Language Understanding. Redwood City, CA: Benjamin/Cummings.
4	Oakes, M. (1998). Statistics for Corpus Linguistics. Edinburgh: Edinburgh University Press.

Additional Readings

Title	
1	Grishman R. (1986) Computational Linguistics: An Introduction. Cambridge, UK: Cambridge University Press.
2	Indurkha N. and Damerau F.J. (2010) Handbook of Natural Language Processing. Boca Raton, FL: Chapman & Hall.
3	McEnery, T. and Wilson, A. (1996) Corpus Linguistics. Edinburgh, UK: Edinburgh University Press.
4	Mitkov R. (2003) Oxford Handbook of Computational Linguistics. Oxford: Oxford University Press.
5	黃昌寧、李涪子 (2002) 《語料庫語言學》 北京：商務印書館

6	俞士汶 (2003) 《計算語言學概論》 北京：商務印書館
7	Journal and conference papers at ACL Anthology http://www.aclweb.org/anthology-index/