


Annamalai University
Centre of Advanced Study in Linguistics

PG. Diploma in Computational Linguistics (Part Time)
Programme Code: LLIN13

Curriculum

(For students admitted from the academic year 2019-2020)

Programme Outcomes

PO1:	Apply the knowledge of language fundamentals and various literatures in society, computers, psychology, cognitive science and medicine.
PO2:	Formulate, solve and analyze complex problems in variety of domains that constitute the core of language and literature knowledge, including familiarity with diverse questions of interest in the areas of (and interfaces between) structures of language and aesthetics of literature.
PO3:	Apply the acquired knowledge for analyzing language and writing in appropriate genres and modes for a variety of purposes and audiences and provide solutions to societal and environmental contexts for problems related to language change, policy and planning.
PO4:	Design and conduct research, analyse and interpret data to provide valid conclusions in the field of literature and in the descriptive as well as applied language studies.
PO5:	Select and apply appropriate modern theories and techniques including cognitive, psychological, biological, cultural, and social factors for language study and research.
PO6:	Gain exposure to attain knowledge and understand interdisciplinary and multidisciplinary linguistic and literary approaches.
PO7:	Acquire professional and intellectual integrity, code of conduct and ethics on communicational practices, understanding responsibilities and norms for sustainable development of society.
PO8:	Interact with the specific linguistic community and with society at large, through critical conversations and prepare, organize, and deliver their work to the public through speaking and writing.
PO9:	Understand the aesthetic and scientific concepts of language and demonstrate the knowledge as a skilled person in teams and multidisciplinary tasks in their profession.
PO10:	Appreciate the need for self-preparation and life-long learning independently in the broadest context of language challenges in the context of multilingualism and globalization.

Programme Specific Outcomes

PSO1:	To understand the computing Human Languages
PSO2:	To analyse human language components
PSO3:	To know the various models for computing human languages.
PSO4:	To have competence of linguistic names which are indispensable for computing natural languages
PSO5:	To differentiate human language and programming language

Course-1. COMPUTER APPLICATIONS TO LANGUAGE STUDIES

Learning Objectives

- ❖ To understand uses of computer, applications of computer in various tasks.
- ❖ To equip sound knowledge on programming language
- ❖ To understand application of computer in accomplishing linguistics tasks
- ❖ To instruct linguistics nuances for computing human languages
- ❖ To understand various models for creating software based on linguistics

UNIT – 1: Introduction to Computers:

Computer generations; Type of computers – Mainframe, mini, and personal computers; Hardware – computer anatomy: Input and Output devices, Central Processing UNIT – Memories: ROM, RAM; Software; Programming Languages; Databases.

UNIT – 2: Introduction to Linguistics:

Definition of Language and Linguistics; Levels in Language analysis – Phonetics, Phonology, Morphology, Syntax, Semantics and Discourse; Theoretical Vs applied; Branches of Linguistics.

UNIT – 3: Computer Applications to Language Studies:

Word processing and DTP (Desk Top Publishing); Text processing; Translation; Lexicography; Language Learning and Teaching; Speech signal processing; Tools for Linguistic analysis – Sorting, Indexing, Frequency counting, KWIC (Key Word In Context).

UNIT – 4: Tools and Techniques for Computer Assisted Text Processing:

Preparing Text for computer analysis; Reusing text and encoding standard; Packages for Description of Texts; Databases; Text Retrieval System; Object Oriented DBMS and Hypertext.

UNIT – 5: Computer Corpus – Definition:

Corpus creation; Corpora of English; - Lancaster – Oslo/ Bergen (LOB) Corpus; Brown Corpus; London – Lund Corpus; Corpora vs Collection; Use of Corpora – in Linguistics, NLP, Lexicography, Speech Processing, Machine Translation, etc. Annotation of Corpora – grammatical tagging, phrase tagging, etc.

Text Books

1. Burnard Lou, Tools and Techniques for Computer – Assisted Text Processing in Computers and Written Texts (ed.) Christopher S. Butler, Blackwell: Oxford. 1992.
2. Hockett, C.F. A course in Modern Linguistics. New Delhi: Oxford and IBH Publishing Co. Verma, S. K. and Krishnamurthy 1989 Modern Linguistics. Oxford University Press. 1973.
3. Leech, Geoffrey, 'Corpus Annotation Schemes' in Literary and Linguistic Computing, Vol.8 No.4. 1993.
4. Leech, Geoffrey, and Steven Fligelstone, Computers and Corpus Analysis in Computers and Written Texts (ed.) Christopher S. Butler, Blackwell: Oxford. 1992.
5. Sproat, Richard, Morphology and Computation. Cambridge: MA: MIT Press. 1992.

Supplementary Reading

1. Annamalai, E, Corpora Development in Indian Languages in Technology and Languages (ed.) B.B. Rajapurohit, Mysore: CILL.1994.
2. Dash, N. S, Corpus Linguistics and Language Technology. New Delhi: Mittal Publications. 2005.
3. Ganesan, M. 'A Scheme for Grammatical Tagging in Indian Languages' in Technology and Languages (ed.) B.B. Rajapurohit, Mysore: CILL. 1994.
4. Leech, Geoffrey, 'Corpus Annotation Schemes' in Literary and Linguistic Computing, Vol.8 No.4. 1993.

Outcome:

After completion of the course, the students will be able to:

CO1: To describe fundamentals of computers and the basics of linguistics

CO2: To explain various applications of computers in linguistics and language studies.

CO3: To create various Tools for Linguistic analysis

CO4: To differentiate various corpora found in linguistics analysis

CO5: To compile e- dictionaries

Outcome Mapping

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PSO1	PSO2	PSO3	PSO4	PSO5
CO1		√	√	√		√	√	√	√	√	√		√		√
CO2	√	√	√	√	√		√		√	√		√	√	√	√
CO3		√		√		√	√	√			√	√			√
CO4	√		√		√	√			√	√		√	√	√	
CO5	√	√		√	√		√	√		√	√		√	√	

Course- 2: NATURAL LANGUAGE PROCESSING**Learning Objectives**

- ❖ To incorporate technology in human language processing
- ❖ To compile machine readable dictionaries
- ❖ To develop morphological and syntactical parsers
- ❖ To introduce the tools for word sense disambiguation

UNIT – 1. Morphological Analysis:

Introduction to Morphology; Inflectional, Derivational and Compositional morphology; Morphological Analysis; Representation of morphological information; MRD(Machine Readable Dictionary) for stems, for suffixes; Word structure rules – context free rules – features and categories – word structure formalism; Unification and its varieties; role of word grammar

UNIT – 2. Natural Language Processing:

Parsing – Definition; Parsing in traditional grammar; Parsing in formal linguistics; Parsing in Artificial Intelligence; Classification of Parsing – Top-down Parsing, Bottom-up Parsing; Transformational Parsers; Augmented Transition Network (ATN); Lexical Functional Grammar (LFG); Paninian frame work.

UNIT – 3. Computational Lexicography:

Introduction to Lexicography; Dictionary Information; Stages in Dictionary preparation – 1) data collection 2) entry selection 3) entry construction 4) entry arrangement – role of computer in each stage; Computer based Dictionary making; MRD(machine Readable Dictionary); Computational Lexicon; Lexical Resources; Language corpus and Lexicography – use of corpus in dictionary compilation – Tools for Lexicographers

UNIT – 4. Semantic Analysis:

Semantics and Logical forms; Word sense and ambiguity; Defining semantic structure – Model theory; Semantic interpretation and Compositionally; A simple grammar and lexicon with semantic interpretation; Semantic interpretation with feature unification

UNIT – 5. Application of Computational Linguistics:

Machine Translation (MT) – History of MT system – different approaches – direct, inter lingual, and transfer; Problems in lexical transfer; Electronic Dictionary (ED) – Advantages of ED – features of ED; Computer Aided Learning / Teaching (CALT) – role of computational linguistics in language teaching

Text Books:

1. Akshar Bharati, Vineet Chaitanya, and Rajeev Sangal, Natural Language Processing: A Paninian Perspective. New Delhi: PHI. 1995.
2. Allen, J. Natural Language Understanding. The Benjamin Company. 1995.
3. Dowdy R. David, Lauri Karttunen and Arnold M. Z. wicky, Natural Language Parsing (Psychological, computational, and theoretical perspective) Cambridge: CUP. 1985.
4. Hutchins W. John, The Evaluation of Machine Translation System in Practical Experience of Machine Translation System (ed.) V. Lawoon, North-Holland Publishing company. 1982.
5. Kenning, M. J. & Kenning, M. M, An Introduction to Computer Assisted Language Teaching. UK: OUP. 1983.

Supplementary Readings:

1. Ganesan, M. ‘Compilation of Electronic Dictionary for Tamil’ in Proceedings of Tamil Internet 2000 Conference, Singapore. 2000.
2. Grishman, Ralph. Computational Linguistics: An Introduction, London: CUP. 1986.
3. Kaplan, R.M. and John Bresnan, ‘Lexical Functional Grammar: A Formal System of Grammatical Representation’ in The Mental Representation of Grammatical Relations (ed.) John Bresnan, MIT. 1982
4. Meijs, Willem, Linguistic Corpora and Lexicography in Annual Review of Applied Linguistics. Vol. No. 16 pp. 99-114. 1996.
5. Ritchie, D. Grame, Alan W. Black, Graham J. Russell and Stepher G. Pulman, Computational Morphology. England: MIT. 1992.

Outcome

At the end of the course, the student will be able to

- CO1:To make the student to incorporate technology in human language processing.
- CO2:To understand the application of machine in language leaning and teaching
- CO3:To compile electronic dictionaries including various steps in dictionary compilation
- CO4:To create cognitive model for artificial Intelligence.
- CO5:To integrate various disciplines while creating software for various tasks.

Outcome Mapping

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	√	√	√		√		√		√	√	√	√		√	
CO2	√	√		√		√		√		√	√		√	√	√
CO3	√		√		√	√		√	√			√	√	√	√
CO4		√		√	√		√		√	√		√	√	√	
CO5	√	√		√		√		√		√	√	√	√		√
