

EFLU-SCHOOL OF LANGUAGE SCIENCES
MA Computational Linguistics PROGRAMME (SEMESTERS I AND III August – November 2018)

List of Courses

LS 131 – Morphology I	-	Prof. Shruti Sircar/ Prof. P. Madhavan
LS 141 – Syntax I	-	Prof. M. Hari Prasad
LS 152 – Semantics I	-	Prof. Utpal Lahiri
LS 175 – Quantitative Methods in Linguistics	-	Dr. Indranil Dutta
LS 176 – NLP with Python	-	Dr. Atreyee Sharma
LS 181 – Natural Language Processing I	-	Dr. Rahul Balusu
LS 221 – Phonology II	-	Prof. Hemalatha Nagarajan
LS 286 – Quantitative Corpus Linguistics	-	Dr. Rahul Balusu
LS 386 – Analyzing Large Speech Corpora	-	Dr. Indranil Dutta

Dr. Indranil Dutta
Programme Coordinator

Prof. M. Hari Prasad
Dean, School of Language Sciences

Course Title	Morphology I
Course Code	LS 131
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Prof. Shruti Sircar / Prof. P. Madhavan
Course Descriptions:	<p>The course will examine morphological phenomena from across the world's languages and try to discuss the relationship of morphology to phonology and other areas of grammar, such as syntax and semantics. The course will help learners</p> <ul style="list-style-type: none"> • to identify the meaningful subparts of words and perform morphological analysis on unfamiliar languages • to understand the major morphological phenomena found in the world's languages • to analyze morphological and phonological patterns in particular languages, and their interactions • to represent morphosyntactic structure diagrammatically, and comprehend the relationship between such structure and meaning
Evaluation Scheme	<p><i>Internal (40%)</i> Two sit down open-book tests, which will be a mixture of multiple choice, problem-solving and short-answer questions One final project based on original work on some morphological aspect of an understudied language</p> <p><i>External: sit down examination (60%)</i></p>

Course Title	Syntax - I
Course Code	LS 141
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Prof. M. Hari Prasad
Course Descriptions:	<p>The aim is to convey the concepts and tools of analysis of the theory and for students to think of data from their own languages.</p> <p>Basic notions (<i>Aspects</i>): competence-performance, idealization, mentalism, the notion of generativity, creative aspects of language use, levels of adequacy.</p> <p><i>GB</i>: Theories of Case, Bounding, Government, Theta roles, Binding, Case; LF, DS, SS, the inverted Y grammar; Projection principle, empty categories, c-command X-bar theory; constituents, movements, case theory; traces and chain; Binding principles; empty categories; LF. Pro drop; LF and quantifier scope, Split INFL hypothesis, VP internal subject, DP analysis</p>
Evaluation Scheme	Midterm: Final: 40: 60

Course Title	Semantics – I
Course Code	LS 152
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Dr. Utpal Lahiri
Course Descriptions:	<p>An introduction to fundamental concepts in formal semantics at an intermediate to advanced level and greater technical sophistication than what was covered in LS-151. Topics covered will include inferential relationships that form the data of semantics (entailment, presupposition, implicature, synonymy, antonymy, etc.), reference and truth, predication, Quantification, lambda Calculus, Generalized Quantifiers. Issues relating to Syntax-semantics mapping will be treated in some detail.</p> <p>Textbook: Chierchia, G. and S. McConnell-Ginet: Meaning and Grammar (2nd edition). 2001, MIT Press.</p>
Evaluation Scheme	Home Works: Mid Term: Final

Course Title	Quantitative Methods in Linguistics
Course Code	LS 175
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Dr. Indranil Dutta
Course Descriptions:	<p>This course is an introduction to using statistical methods for linguistic analysis. In the past decade or so, linguists of all persuasions have relied extensively on statistical methods, both for unearthing patterns and seeking generalizations. This includes attempts to model linguistic behavior in quantitative terms, and also using model-theoretic approaches along with empirical methods. In that respect, the primary goal of this course is for us to learn statistical methods, tests and techniques to further this ‘statistical’ turn in linguistic analyses. We will follow the primary textbook that outlines the uses of statistical methods for most all sub-disciplines of linguistics. In addition, we will read supplementary material from the recommended readings. The expectation at the conclusion of the course is that you will be able to design your own studies, perform statistical analyses, present your results and use advanced data visualization techniques to advance your arguments. The topics that we will cover will include the following but are not limited to:</p> <ul style="list-style-type: none"> • Central limit function • Normal, chi-square and F-distribution • Analysis of Variance (ANOVA) • Linear and logistic regression • Principal component analysis (PCA) • Multi-dimensional scaling (MDS) • Mixed-effects modeling
Evaluation Scheme	Weekly quiz 10%, Weekly homework 10%, Midterm examination 20%, Final examination 60%

Course Title	NLP with Python
Course Code	LS 176
Semester	I/III
No. of Credits	5
Name of Faculty Member(s)	Dr. Atreyee Sharma
Course Description: 150/200 words	<p>The aim of this course is to learn basic natural language data manipulation using the Natural Language Toolkit (NLTK), a set of open source libraries which are a part of the Python development environment. Students will learn basic Python functions to achieve simple text processing and manipulation tasks. These will involve regular expressions for normalizing and tokenizing text; word and sentence level segmentation of large unannotated corpora; Part-of-Speech (POS) tagging algorithms and implementation; supervised classification of text and evaluation of classification methods.</p> <p>Reading list: Bird, Steven, Ewan Klein, and Edward Loper. 2009. Natural Language Processing with Python. O'reilly Publishing. Perkins, Jacob. 2010. Python Text Processing with NLTK 2.0 Cookbook. Packt Publishing.</p>
Evaluation Scheme	Mid-term: Final::40:60

Course Title	Natural Language Processing I
Course Code	LS 181
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Dr. Rahul Balusu
Course Descriptions:	<p>A largely non-technical introduction to computational morphology, syntax & semantics and information retrieval. Survey of natural language processing applications for parsing and information processing. Presentation of tools and resources needed for NLP applications.</p> <p>Topics: Essential algorithms for processing word classes and structured relationships among words, part-of- speech taggers based on HMMs and transformation-based learning, CYK and Early algorithms for parsing, unification and typed feature structures, lexicalized and probabilistic parsing, analytical tools like the Chomsky hierarchy and the pumping lemma, first order predicate calculus and other ways of representing meaning, applications to information retrieval, extraction, speech understanding, and machine translation.</p>
Evaluation Scheme	Mid term: Final: 40: 60

Course Title	Phonology II
Course Code	LS 221
Semester	III
No. of Credits	5
Name of Faculty Member(s)	Prof. Hemalatha Nagarajan
Course Descriptions:	<p>This course aims to familiarize students with:</p> <p>a. Lexical phonology and b. Optimality theory</p> <p>Lexical phonology examines the interaction between morphology and phonology and explains the morphologically-conditioned phonological rules in terms of different strata or levels.</p> <p>Optimality theory was introduced in the early 1990s as an alternative model of the organization of natural human language sound systems. This model does not have rules or rule-ordering but instead introduces <i>constraints</i> that are <i>ordered</i> and which are <i>violable</i>. In this course, the basic principles of optimality theory will be introduced and explained (GEN, CON, and EVAL). Three important constraint families will be explored (Faithfulness, Alignment, and Markedness). The course examines recurring issues in phonological theory and shows how optimality theory might account for them.</p> <p>This course will be give you hands-on experience in analyzing data (from known and unknown languages) and equip you for research in the area of phonology.</p> <p>Pre-requisite: Basic Issues in Phonology (but not mandatory)</p>
Evaluation Scheme	<p>40 % Internal: 3 Tests (Open Book)</p> <p>60% External-Semester-end examination-(Open Book)</p>

Course Title	Quantitative Corpus Linguistics
Course Code	LS 286
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Dr. Rahul Balusu
Course Descriptions:	<p>This course is a comprehensive introduction to corpus linguistics and corpus-based methods using R. A part of the course will be devoted to learning about usage-based investigations of linguistic representations. Additionally, while learning about how to collect and work with available text and speech corpora, we will also learn corpus based methods that have enriched our understanding of semantics, pragmatics, sociolinguistics, discourse analysis, forensics, statistical methods in speech and language processing, diachronic linguistics, lexicographic and lexical studies, and second language studies. To that end, apart from familiarizing ourselves with various text and speech analysis software's we will also learn minimal programming in R and Perl. This course will also involve using the tools and techniques learnt in the course to develop projects where linguistically-informed hypotheses will be put to test on large Indian language corpora.</p> <p><i>Prerequisite: LS181</i></p>
Evaluation Scheme	Mid term: Final: 40: 60

Course Title	Analyzing Large Speech Corpora
Course Code	LS 386
Semester	I/III
No. of Credits	5
Name of Faculty Member(S)	Dr. Indranil Dutta
Course Descriptions:	In this course we will learn techniques and tools for large scale exploration of speech corpora. We will integrate knowledge from signal processing techniques, large scale annotation schemas, querying databases and machine learning classification techniques. Topics will include: Building and querying speech databases, large scale processing of speech signals, querying annotation structures, formant and spectral analyses, and using Machine Learning algorithms for classification of speech signals
Evaluation Scheme	Mid-term: Final::40:60