

JRF in Linguistics 2012

PAPER I (RLNI)

General Linguistics

SYLLABUS

Theoretical issues of corpus linguistics – definition and types – features of corpus – typology of corpus – corpus designing principles and methodologies – text, speech and multimodal corpus generation – corpus data archiving and management – linguistic data and information representation and retrieval - corpus editing techniques and tools – text sanitation and management – extralinguistic and intralinguistic annotation – orthographic annotation – prosodic annotation – grammatical annotation – lexical annotation – semantic annotation – anaphoric annotation – discourse annotation – rhetoric annotation – frequency counts – lexical sorting – concordance – lemmatisation – local-word-grouping – key-word-in-context – type-token analysis – lexical generativity – lexical collocation – morphological processing – morphological generation – chunking – multiword decomposition and paraphrasing – sentence analysis and interpretation – parsing – contextual meanings of words – word sense disambiguation – semantic net – WordNet - Synset – synonymy – polysemy – homonymy – hypernymy – hyponymy – social variation of language use – regional variations – register variations – diachronic variations – gender variation – text variations – discourse – pragmatics – conversation analysis – dialogue interpretation – speech acts – stylometrics – first and second language teaching – computer assisted language teaching – data driven learning – online language teaching – lexical resource generation – printed dictionary – digital dictionary – online dictionary – online thesaurus – usage based grammars – spoken language grammar – online interactive grammar – machine translation – machine-aided translation – domain-based machine translation – models of machine translation – theories of machine translation – approaches of machine translation, etc.

Time: 2 hours

Full Marks: 100

It contains twenty (20) short questions. You have to answer all the questions – each one in about fifty (50) words. Each question carries five (5) marks (5 x 20 = 100 marks)

1. How does speech differ from text?
2. What is linguistic information retrieval?
3. What are common text processing techniques?
4. What is transcription?
5. What is collocation?
6. What is lexical gap?

7. What is allomorph?
8. What is lexicography?
9. What is hypercorrection?
10. What is Rule Ordering?

JRF in Linguistics 2012

PAPER II (RLNII)

Computational Linguistics

Time: 2 hours

Full Marks: 100

It contains ten (10) descriptive questions. You have to answer all the questions – each one in about three hundred (300) words. Each question carries ten (10) marks (10 x 10 = 100 marks)

- (1) Define the basic features of a speech corpus.
- (2) Show difference between speech and spoken corpus with examples.
- (3) Show with examples how POS tagging is done.
- (4) How does concordance help on word sense disambiguation?
- (5) What is context? How does it contribute in understanding sense of words?
- (6) What is WordNet? What do you know about IndoWordNet?
- (7) Show with examples how does polysemy differ from homonymy?
- (8) What is example based MT? Show from English to Bengali.
- (9) What are the salient features of an on-line dictionary?
- (10) What is the difference between bottom-up and top-down parsing? Explain.