

**Workshop on
WORDNET FOR DRAVIDIAN LANGUAGES
2 - 3, June 2003**

**ANNA UNIVERSITY
CHENNAI**



Organised by

**AU-KBC Research Centre
MIT Campus, Anna University**

Jointly with

**Department of Linguistics
Tamil University, Thanjavur**

&

**Central Institute of Indian Languages
Mysore**

Background:

A WordNet is a lexical database whose design is inspired by current psycholinguistic theories of human lexical memory. In WordNet, there two kinds of relations that are recognized: lexical and semantic. Lexical relations are between word forms; semantic relations are between word meanings. Basic categories such as nouns, verbs, adjectives and adverbs are organized into synonym sets, each one of which represents an underlying lexical concept. WordNets represent a useful linguistic resource in that they can help in enhancing the performance of Machine Translation Systems and Search Engines.

WordNets are being globally established for a number of languages, such as EuroWordNet for European languages. In Indian languages, WordNets are being developed at IIT Bombay (Hindi and Marathi), MS University, Baroda (Gujarati), IIT Kharagpur (Bengali), Utkal University, Bhubaneswar (Oriya) and AU-KBC Research Centre, Chennai (Tamil).

Analogous to the EuroWordNet, building an IndoWordNet needs WordNets in all major Indian languages. Among the Indian languages, the South Indian (Dravidian) family of languages such as Tamil, Telugu, Kannada and Malayalam have close association in many linguistic features such as morphology and semantics. They also share a lot of culture-specific features. Building WordNets in these languages will make easier the task of a linked IndoWordNet.

The present Workshop is being organised on the basis of the experience gained in building a Tamil WordNet. The Workshop aims to explore the common features of building WordNets for South Indian languages.

Objectives of the Workshop:

- To present ongoing work on WordNets in Dravidian languages.
- To explore the common features of Dravidian languages that are useful in a WordNet.
- To exchange ideas, experience, tools and resources in building WordNets.
- To evolve a program for the development of WordNets in the four major Dravidian languages.

Workshop Details:

In addition to covering the background as well as theoretical and computational aspects of WordNets and their construction, there will be presentation and demonstration on the Tamil WordNet. Reference material on WordNet will be sent to the participating groups once they confirm their participation. We will also provide the participants with approximately 1000 Tamil words (chosen from the broad semantic domain of Inanimate Entities) along with their English equivalents well ahead of the Workshop, so that they can generate the equivalents in their respective languages before the Workshop. The organisers will provide them with the necessary WordNet interfaces for entering equivalents in the various languages, so that the Workshop would have a strong hands-on component.

The Workshop will be conducted by Dr.S.Rajendran, Department of Linguistics, Tamil University, Thanjavur in coordination with the other organising institutions.

There will also be presentations by other experts who have done work in this area.

Target Audience:

Linguists / Computer Scientists / Researchers / Students from Universities or Institutions engaged in or interested in the area of language processing in Dravidian languages.

Registration:

Interested persons may register through e-mail or on plain paper, giving full details. TA limited to III AC and DA as per rules will be available to a limited number of participants.

For registration or further information, contact:

Dr.S.Arulmozi
AU-KBC Research Centre
MIT Campus, Anna University
Chromepet, Chennai 600044
Ph.: 044-22232711 / 22234885
Fax: 044-22231034
E-Mail: arulmozi@au-kbc.org