

## RANLP 2005

### Programme

#### **Day 1 (21 September)**

8.45 Opening

9.00-10.00 Keynote speech: Ido Dagan

Morning Session 1a (regular papers) – parallel session to 1b and 1c

#### *Textual entailment*

10.00-10.30

Tree Edit Distance for Textual Entailment

Milen Kouylekov and Bernardo Magnini

10.30-11.00

Lexico-Syntactic Subsumption for Textual Entailment

Vasile Rus, Art Graesser and Kirtan Desai

Morning Session 1b (regular papers) – parallel session to 1a and 1c

#### *Document classification*

10.00-10.30

Document Classification Using Semantic Networks with An Adaptive Similarity Measure

Filip Ginter, Sampo Pyysalo and Tapio Salakoski

10.30-11.00

Using the Essence of Texts to Improve Document Categorization

Rada Mihalcea and Samer Hassan

Morning Session 1c (regular papers) – parallel session to 1a and 1b

#### *Parallel corpora*

10.00-10.30

Lexical Transfer Selection Using Annotated Parallel Corpora

Stelios Piperidis, Panagiotis Dimitrakis and Irene Balta

10.30-11.00

Parallel corpora for medium density languages

Dániel Varga, Péter Halácsy, András Kornai, Viktor Nagy, László Németh and Viktor Trón

Coffee break 11.00-11.30

Morning session 2a (short papers) – parallel session to 2b and 2c

*Named entity recognition and classification*

11.30-11.55

Noun Phrases and Named Entities in Biomedical Texts: Does Domain Change without Retraining Matter?

Udo Hahn, Joachim Wermter and Juliane Fluck

11.55-12.20

Towards large-scale, open-domain and ontology-based named entity classification

Philipp Cimiano and Johanna Voëlker

12.20-12.45

A Proposal For An Arabic Named Entity Tagger Leveraging a Parallel Corpus

Doaa Samy, Antonio Moreno and Josè M<sup>a</sup> Guirao

Morning session 2b (short papers) – parallel session to 2a and 2c

*Summarisation*

11.30-11.55

Concept-centred summarisation: producing glossary entries for terms using summarisation methods

Le An Ha, Constantin Orasan

11.55-12.20

Summarizing Reports on Evolving Events; Part I: Linear Evolution

Stergos D. Afantenos, Vangelis Karkaletsis, Panagiotis Stamatopoulos

12.20-12.45

Automation of Summary Evaluation by the Pyramid Method

Aaron Harnly, Ani Nenkova, Rebecca Passonneau, Owen Rambow

Morning session 2c (short papers) – parallel session to 2a and 2b

*Word sense disambiguation*

11.30-11.55

An Integrated approach to Word Sense Disambiguation

Jordi Atserias, Lluís Padró, German Rigau

11.55-12.20

Word Sense Disambiguation Using Co-Occurrence Statistics on Random Labels

Martin Hassel

12.20-12.45

Exploring feature spaces with SVD and unlabeled data for Word Sense Disambiguation

Eneko Agirre, Oier Lopez de Lacalle, David Martínez

Lunch break 12.45-14.30

14.30- 15.30 Keynote speech: Robert Dale

Afternoon session 1a (short papers) – parallel session to 1b and 1c

*Educational applications*

15.30-15.55

Automatic Identification of Cognates and False Friends in French and English

Diana Inkpen, Oana Frunza, Grzegorz Kondrak

15.55-16.20

About the effects of using Anaphora Resolution in assessing free-text student answers

Diana Pérez, Oana Postolache, Enrique Alfonseca, Dan Cristea and Pilar Rodriguez

Afternoon session 1b (short papers) – parallel session to 1a and 1c

*Spelling correction and grammar checking*

15.30-15.55

Faking Errors to Avoid Making Errors: Very Weakly Supervised Learning for Error Detection in Writing

Jonas Sjöbergh, Ola Knutsson

15.55-16.20  
Context-Based Ranking of Suggestions for Spelling Correction  
Julia Trushkina

Afternoon session 1c (short papers) – parallel session to 1a and 1b

*Semantic parsing*

15.30-15.55  
Assigning Function Labels to Unparsed Text  
Gabriele Musillo and Paola Merlo

15.55-16.20  
Selecting Features for Semantic Roles in QA Systems  
Paloma Moreda and Manuel Palomar

Coffee break 16.20-16.50

Poster session 1 (*Techniques*): 16.50-18.00

**Day 2 (22 September)**

9.00-10.00 Keynote speech: Makoto Nagao

Morning Session 3a (regular papers) – parallel session to 3b and 3c

*Semantic similarity*

10.00-10.30  
Text Semantic Similarity, with Applications  
Courtney Corley, Andras Csomai, Rada Mihalcea

10.30-11.00  
A Simple WWW-based Method for Semantic Word Class Acquisition ???  
Keiji Shinzato, Kentaro Torisawa

Morning Session 3b (regular papers) – parallel session to 3a and 3c

*Discourse processing*

10.00-10.30

Exploiting Linguistic Cues to Classify Rhetorical Relations  
Caroline Sporleder and Alex Lascarides

10.30-11.00

A Data-driven Approach to Pronominal Anaphora Resolution for German  
Erhard W. Hinrichs, Katja Filippova and Holger Wunsch

Morning Session 3c (regular papers) – parallel session to 3a and 3b

*Ontologies*

10.00-10.30

Semantic Indexing using Minimum Redundancy Cut in Ontologies  
Florian Seydoux, Jean-Cédric Chappelier

10.30-11.00

Automatic Building of Wordnets  
Eduard Barbu, Verginica Barbu Mititelu

Coffee break 11.00-11.30

Morning session 4a (short papers) – parallel session to 4b and 4c

*Tagging and parsing*

11.30-11.55

Target-Language-Driven Agglomerative Part-of-Speech Tag Clustering for Machine Translation  
Felipe Sánchez-Martínez, Juan Antonio Pérez-Ortiz and Mikel L. Forcada

11.55-12.20

Knowledge-Poor Approach to Dependency Parsing: Dependency parsing based on Morpho-Syntactic Information  
Julia Trushkina

12.20-12.45

Robust stochastic parsing using optimal maximum coverage  
Vladimír Kadlec, Marita Ailomaa, Jean-Cédric Chappelier, Martin Rajman

Morning session 4b (short papers) – parallel session to 4a and 4c

*Text classification and text normalisation*

11.30-11.55

Feature Selection for Electronic Negotiation Texts

Marina Sokolova, Vivi Nastase, Mohak Shah, Stan Szpakowicz

11.55-12.20

Optimizing the subtasks in the double classification approach to information extraction

Viktor Pekar, Richard Evans

12.20-12.45

Truecasing for the PORTAGE system

Akakpo Agbago, Roland Kuhn, George Foster

Morning session 4c (short papers) – parallel session to 4a and 4b

*Syntax*

11.30-11.55

Persian word order is free but not discontinuous

Allan Ramsay, Najmeh Ahmed and Vahid Mirzaeian

11.55-12.20

Estimating Grammaticality from a Small Corpus

Asad B. Sayeed and Stan Szpakowicz

12.20-12.45

Syntactic Identification of Attribution in the RST Treebank

Peter Rossen Skadhauge and Daniel Hardt

Lunch break 12.45-14.30

14.30- 15.30 Keynote speech: John Nerbonne

Afternoon session 2a (regular papers) – parallel session to 2b and 2c

*Development of annotated resources*

15.30-16.00

Exploiting Parallel Texts to Produce a Multilingual Sense Tagged Corpus for Word Sense Disambiguation

Lucia Specia, Maria das Graças V. Nunes, Mark Stevenson

16.00-16.30

The SenSem Project: syntactico-semantic annotation of sentences in Spanish

Laura Alonso, Joan Antoni Capilla, Irene Castellón, Ana Fernández-Montraveta, Gloria Vázquez

Afternoon session 2b (regular papers) – parallel session to 2a and 2c

*Word distribution analysis and applications*

15.30-16.00

Disambiguating Coordinations Using Word Distribution Information

Francis Chantree, Adam Kilgarriff, Anne de Roeck, Alistair Willis

16.00-16.30

Very Frequent Function Words do not Distribute Homogeneously

Avik Sarkar, Anne De Roeck, Paul H Garthwaite

Afternoon session 2c (regular papers) – parallel session to 2a and 2b

*Grammars/parsing*

15.30-16.00

One Step toward a Richer Model of Unsupervised Grammar Induction

Heshaam Feili, Gholam-Reza Ghassem-Sani

16.00-16.30

How Do Treebank Annotation Schemes Influence Parsing Results? Or How Not to Compare Apples And Oranges

Sandra Kübler

Coffee break 16.30-17.00

Poster session 2 (*Applications*): 17.00-18.10

### **Day 3 (23 September)**

9.00-10.00 Keynote speech: Ralph Grishman

Morning session 5a (regular papers) – parallel session to 5b

#### *Document indexing*

10.00-10.30

Terms Representation with Generalized Latent Semantic Analysis

Irina Matveeva, Gina-Anne Levow, Ayman Farahat and Christiaan Royer

10.30-11.00

Indexing and Querying Linguistic Metadata and Document Content

Niraj Aswani, Valentin Tablan, Kalina Bontcheva and Hamish Cunningham

Morning session 5b (regular papers) – parallel session to 5a

#### *Applications (Machine Translation, Question Answering)*

10.00-10.30

Multi-Perspective Evaluation of the FAME Speech-to-Speech Translation System for Catalan, English and Spanish

Victoria Arranz, Elisabet Comelles, David Farwell

10.30-11.00

Optimizing information retrieval in question answering using syntactic annotation

Jörg Tiedemann

Coffee break 11.00-11.30

Morning session 6a (short papers) – parallel session to 6b

#### *Summarisation and question answering*

11.30-11.55

Adapting Lexical Chaining to Summarize Conversational Dialogues

Iryna Gurevych and Thade Nahnsen

11.55-12.20

The Use of Causal Expressions for Abstracting and Question Answering

Christopher Paice and William Black



12.20-12.45

Integrating NLP Tools to Support Information Access to News Archives

H. Saggion, E. Barker, R. Gaizauskas, Jonathan Foster and A. Dali

Morning session 6b (short papers) – parallel session to 6a

*Semantic similarity/closeness and semantic interpretation*

11.30-11.55

Using context-window overlapping in synonym discovery and ontology population

María Ruiz-Casado, Enrique Alfonseca, Pablo Castells

11.55-12.20

Exploring Features to Identify Semantic Nearest Neighbours: A Case Study on German Particle Verbs

Sabine Schulte im Walde

12.20-12.45

Extracting Predicate Structures from Parse Trees

Manfred Klenner

Lunch break 12.45-14.30

14.30- 15.30 Keynote speech: Anne de Roeck

Afternoon session 3a (short papers) – parallel session to 3b

*Morphology and lexicography*

15.30-15.55

Unsupervised, Knowledge-Free Morpheme Boundary Detection

Stefan Bordag

15.55-16.20

Building a Cross-lingual Referential Knowledge Database using Dictionaries

Shigeko Nariyama, Takaaki Tanaka, Eric Nichols, Francis Bond, Hiromi Nakaiwa

16.20-16.45

Multilingual Lexical Acquisition by Bootstrapping Cognate Seed Lexicons

Kornél Markó, Stefan Schulz, Udo Hahn

Afternoon session 3b (short papers) – parallel session to 3a

*Temporal processing, PP attachment disambiguation and linguistic annotations*

15.30-15.55

Temporally Ordering Event Instances in Natural Language Texts  
Argyrios Vasilakopoulos, William J. Black

15.55-16.20

Enough is Enough --- Estimating Upper Bounds of the Size of Training Corpora for  
Unsupervised PP Attachment Disambiguation  
Michael Poprat

16.20-16.45

A framework for representing and managing linguistic annotations based on typed  
feature structures  
X. Artola, A. Díaz de Ilarraza, N. Ezeiza, K. Gojenola, G. Labaka, A. Sologaitoa, A.  
Soroa

Coffee break 16.45-17.15

Afternoon session 4a (short papers) – parallel session to 4b

*Using large corpora and the web*

17.15-17.40

Comparing methods for automatic acquisition of Topic Signatures  
Montse Cuadros, Lluís Padro, German Rigau

17.40-18.05

On the Instability of Using Search Engine Page Hits as a Proxy for n-gram  
Frequencies  
Preslav Nakov and Marti Hearst

Afternoon session 4b (short papers) – parallel session to 4a

*Speech synthesis*

17.15-17.40

SpeechLab 2.0 -- A High-Quality Text-to-Speech System for Bulgarian  
Maria Andreeva, Ivaylo Marinov and Stoyan Mihov

18.05 Closure

***Poster session 1***

Finite-State Morphology of Estonian: Two-Levelness Extended

*Heli Uibo*

On Measurement of Semantic Information and Valency of the English Written Text

*Nataliya Fedyushkina, Nadya Vishnyakova*

A corpus-based model for bridging anaphora resolution in Italian.

*Tommaso Caselli, Irina Prodanof*

Composite Topics in Discourse

*Frédéric Bilhaut*

Automatic Acquisition of Expressions Representing Preparation and Utilization of an Object

*Kentaro Torisawa*

Automatic Import of Verbal Syntactic Relations Using Parallel Corpora

*Verginica Barbu Mititelu, Radu Ion*

Knowledge based Feature Engineering using Text Sense Representation Trees

*Ronald Winnemöller*

Knowledge Acquisition Based on Automatically-Extracted Word Hierarchies from Domain-Specific Texts

*Eiko Yamamoto and Hitoshi Isahara*

Adapting a general parser to a sublanguage

*Sophie Aubin, Adeline Nazarenko, Claire Nédellec*

Visual Parser Builder

*Dimitar Blagoev, George Totkov*

EAGLES compliant tagset for the morphosyntactic tagging of Esperanto

*Antonio Toral, Sergio Ferrández, Andrés Montoyo*

An interactive environment for creating and validating syntactic rules

*Panagiotis Bouros, Aggeliki Fotopoulou, Nicholas Glaros*

Approaching Sequential NLP Tasks with an Automata Acquisition Algorithm

*Muntsa Padró, Lluís Padró*

Lexicalization of Grammars with Parameterized Graphs

*Olivier Blanc, Matthieu Constant*

A Core Model of Systemic Linguistic Analysis  
*Sylviane Cardey, Peter Greenfield*

An advanced approach for English-Vietnamese syntactic tree transfer  
*DO Xuan Quang, NGUYEN Luu Ngan, DINH Dien*

OLE - A New Ontology Learning Platform  
*Vít Nováček*

Investigating the features that affect cue usage of non-native speakers of English  
*Xinyu Deng and Nakamura Jun-ichi*

Towards the Automatic Derivation of Lexical Prototypes  
*Aleem Hossain, Mark G. Lee*

Finite State Morphological Analysis of Amharic for Parallel Text Alignment  
*Saba Amsalu, Dafydd Gibbon*

## ***Poster session 2***

Discovering Topic Boundaries for Text Summarization based on Word Co-occurrence  
*Gaël Dias, Elsa Alves*

Partial Forest Transfer for Spoken Language Translation  
*Tetsuro Chino, Satoshi Kamatani*

Parameter reduction in unsupervised trained sliding-window part-of-speech taggers  
*Enrique Sánchez-Villamil, Mikel L. Forcada, Rafael C. Carrasco*

Spelling Correction in Context  
*Guillaume Pinot, Chantal Enguehard*

Recognition of Personal Names in Serbian Texts  
*Cvetana Krstev, Duško Vitas, Sandra Gucul*

Guidelines for the syntactico-semantic annotation of a corpus in Spanish  
*Gloria Vázquez, Ana Fernández-Montraveta, Laura Alonso*

Handling Corrupt Input in Domain-specific Spoken Dialogue Systems  
*Wei-Lin Wu, Ru-Zhan Lu, Feng Gao, Hui Liu*

Analysis and Visualization for Daily Newspaper Corpora  
*Matthias Richter*

Pattern Ambiguity and its Resolution in English to Hindi Translation  
*Niladri Chatterjee, Shailly Goyal, Anjali Naithani*

Using Language Portable Detection for Spanish Named Entity Recognition  
*Zornitsa Kozareva, Oscar Ferrández, Andres Montoyo, Rafael Muñoz*

Self-training and Co-training applied to Spanish Named Entity Recognition  
*Zornitsa Kozareva, Boyan Bonev, Andres Montoyo*

The Intersection of Latent Semantic Analysis and Educational Applications: A Research Taxonomy  
*Debra Trusso Haley, Pete Thomas, Anne DeRoeck, Marian Petre*

An Algorithm Detecting and Tracing Errors in ICALL Systems  
*Veit Reuer, Kai-Uwe Kühnberger*

Evaluating corpus query systems on functionality and speed: TIGERSearch and Emdros  
*Ulrik Petersen*

Weak Named Entities Recognition using morphology and syntax  
*Antonio Toral, Rafael Muñoz, Andrés Montoyo*

Comparison of Topic Identification methods for Arabic language  
*M. Abbas, K. Smaili*

Rapid methods for optimal text selection  
*Rahul Chitturi, Sebsibe H Mariam, Rohit Kumar*

Knowledge Acquisition for Fine-grained Named Entity Classification  
*Olga Uryupina*

Using NLP to build back-of-the-book indexes  
*Touria AÏT MEKKI, Adeline NAZARENKO*

A syntactic strategy for filtering sentences in a question answering  
*Vincent Barbier, Anne-Laure Ligozat*

Using WordNet Similarity and Antonymy Relations to Aid Document  
*Thomas de Simone, Dimitar Kazakov*