

# Table of Contents

---

## Computational Linguistics

---

### Semantics

*Invited Talk:*

Multiword Expressions: A Pain in the Neck for NLP . . . . .	1
<i>Ivan A. Sag, Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger</i>	

A Hypothesis on the Origin of the Sign Types . . . . .	16
<i>Roland Hausser</i>	

Quantification and Intensionality in Situation Semantics . . . . .	32
<i>Roussanka Loukanova</i>	

Generalized Quantification in Situation Semantics . . . . .	46
<i>Roussanka Loukanova</i>	

Sign Language Translation via DRT and HPSG . . . . .	58
<i>Éva Sáfar and Ian Marshall</i>	

Multilayered Extended Semantic Networks as a Language for Meaning Representation in NLP Systems . . . . .	69
<i>H. Helbig and C. Gnörlich</i>	

Constructing a Sensuous Judgment System Based on Conceptual Processing . . . . .	86
<i>Atsushi Horiguchi, Seiji Tsuchiya, Kazuhide Kojima, Hirokazu Watabe, and Tsukasa Kawaoka</i>	

Towards a Natural Language Driven Automated Help Desk . . . . .	96
<i>Melanie Knapp and Jens Woch</i>	

### Word Sense Disambiguation

*Invited Talk:*

Lexical Tuning . . . . .	106
<i>Yorick Wilks and Roberta Catizone</i>	

A Baseline Methodology for Word Sense Disambiguation . . . . .	126
<i>Ted Pedersen</i>	

An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet . . . . .	136
<i>Satanjeev Banerjee and Ted Pedersen</i>	

Feature Selection Analysis for Maximum Entropy-Based WSD .....	146
<i>Armando Suárez and Manuel Palomar</i>	

Combining Supervised-Unsupervised Methods for Word Sense Disambiguation .....	156
<i>Andrés Montoyo, Armando Suárez, and Manuel Palomar</i>	

A Proposal for WSD Using Semantic Similarity .....	165
<i>Susana Soler and Andrés Montoyo</i>	

## **Anaphora**

### *Invited Talk:*

A New, Fully Automatic Version of Mitkov's Knowledge-Poor Pronoun Resolution Method .....	168
<i>Ruslan Mitkov, Richard Evans, and Constantin Orasan</i>	

Pronominal Anaphora Generation in an English-Spanish MT Approach ..	187
<i>Jesús Peral and Antonio Ferrández</i>	

Using LSA for Pronominal Anaphora Resolution .....	197
<i>Beata Klebanov and Peter Wiemer-Hastings</i>	

## **Syntax and Parsing**

The Spanish Auxiliary Verb System in HPSG .....	200
<i>Ivan Meza and Luis Pineda</i>	

Surface Syntactic Relations in Spanish .....	210
<i>Igor A. Bolshakov</i>	

Parsing Ill-Formed Inputs with Constraint Graphs .....	220
<i>Philippe Blache and David-Olivier Azulay</i>	

## **Part of Speech Tagging**

Part-of-Speech Tagging with Evolutionary Algorithms .....	230
<i>Lourdes Araujo</i>	

Formal Methods of Tokenization for Part-of-Speech Tagging .....	240
<i>Jorge Graña, Fco. Mario Barcala, and Jesús Vilares</i>	

Sepe: A POS Tagger for Spanish .....	250
<i>Héctor Jiménez and Guillermo Morales</i>	

Fuzzy Set Tagging .....	260
<i>Dariusz J. Kogut</i>	

## Lexicon and Corpus

### *Invited Talk:*

Towards a Standard for a Multilingual Lexical Entry: The EAGLES/ISLE Initiative . . . . .	264
<i>Nicoletta Calzolari, Antonio Zampolli, and Alessandro Lenci</i>	
Quantitative Comparison of Homonymy in Spanish EuroWordNet and Traditional Dictionaries . . . . .	280
<i>Igor A. Bolshakov, Sofia N. Galicia-Haro, and Alexander Gelbukh</i>	
Compilation of a Spanish Representative Corpus . . . . .	285
<i>Alexander Gelbukh, Grigori Sidorov, and Liliana Chanona-Hernández</i>	
Aligning Multiword Terms Using a Hybrid Approach . . . . .	289
<i>Arantza Casillas and Raquel Martínez</i>	
Automatic Selection of Defining Vocabulary in an Explanatory Dictionary . . . . .	300
<i>Alexander Gelbukh and Grigori Sidorov</i>	

## Text Generation

Integrated Natural Language Generation with Schema-Tree Adjoining Grammars . . . . .	304
<i>Karin Harbusch and Jens Woch</i>	
Experiments with a Bilingual Document Generation Environment . . . . .	314
<i>Arantza Casillas and Raquel Martínez</i>	

## Morphology

A Computational Model of Change in Politeness with the Addition of Word Endings . . . . .	318
<i>Tamotsu Shirado and Hitoshi Isahara</i>	
Tartar Morphology Implementation . . . . .	327
<i>Dj.Sh. Suleymanov</i>	

## Speech

Automatic Generation of Pronunciation Lexicons for Spanish . . . . .	330
<i>Esmeralda Uraga and Luis Pineda</i>	

---

## Intelligent Text Processing

---

### Spelling

Diacritics Restoration: Learning from Letters versus Learning from Words . . . . .	339
<i>Rada F. Mihalcea</i>	

## Information Extraction and Information Retrieval

A Comparative Study of Information Extraction Strategies . . . . .	349
<i>Ronen Feldman, Yonatan Aumann, Michal Finkelstein-Landau, Eyal Hurvitz, Yizhar Regev, and Ariel Yaroshevich</i>	
Answer Extraction in Technical Domains . . . . .	360
<i>Fabio Rinaldi, Michael Hess, Diego Mollá, Rolf Schwitter, James Dowdall, Gerold Schneider, and Rachel Fournier</i>	
Automatic Extraction of Non-standard Lexical Data for a Metalinguistic Information Database . . . . .	370
<i>Carlos Rodríguez</i>	
Text Segmentation for Efficient Information Retrieval . . . . .	373
<i>Fernando Llopis, Antonio Ferrández, and José Luis Vicedo</i>	
Using Syntactic Dependency-Pairs Conflation to Improve Retrieval Performance in Spanish . . . . .	381
<i>Jesús Vilares, Fco. Mario Barcala, and Miguel A. Alonso</i>	

## Summarization

Multi-document Summarization Using Informative Words and Its Evaluation with a QA System . . . . .	391
<i>June-Jei Kuo, Hung-Chia Wung, Chuan-Jie Lin, and Hsin-Hsi Chen</i>	

## Text Mining

Automated Selection of Interesting Medical Text Documents by the TEA Text Analyzer . . . . .	402
<i>Jan Žižka and Aleš Bourek</i>	

## Text Classification and Categorization

Chinese Documents Classification Based on N-Grams . . . . .	405
<i>Shuigeng Zhou and Jihong Guan</i>	
Cross-Lingual Document Similarity Calculation Using the Multilingual Thesaurus EUROVOC . . . . .	415
<i>Ralf Steinberger, Bruno Pouliquen, and Johan Hagman</i>	
Empirical Formula for Testing Word Similarity and Its Application for Constructing a Word Frequency List . . . . .	425
<i>Pavel Makagonov and Mikhail Alexandrov</i>	

## Document Processing

AutoMarkup: A Tool for Automatically Marking up Text Documents . . . . .	433
<i>Shazia Akhtar, Ronan G. Reilly, and John Dunnion</i>	

Identification of Recurrent Patterns to Extract Definitory Contexts . . . . .	436
<i>Gerardo Sierra and Rodrigo Alarcón</i>	

## **Demo Descriptions**

Specification Marks Method: Design and Implementation . . . . .	439
<i>Sonia Vázquez, M<sup>a</sup> Carmen Calle, Susana Soler, and Andrés Montoyo</i>	

<b>Author Index</b> . . . . .	443
-------------------------------	-----

Computational Linguistics and Intelligent Text  
Processing

Third International Conference, CICLing 2002, Mexico  
City, Mexico, February 17-23, 2002 Proceedings  
Gelbukh, A. (Ed.)

2002, XIV, 450 p., Softcover

ISBN: 978-3-540-43219-7