

**10th Conference of the European Chapter of the Association for
Computational Linguistics
EACL03**

Programme
Main Conference

15-17 April 2003, Budapest, Hungary

The people behind EACL 03:

Programme Co-Chairs	Ann Copestake (United Kingdom) Jan Hajic (Czech Republic)
Research Note and Demo Chair	Alberto Lavelli (Italy)
Tutorial Chair	Dan Cristea (Romania)
Publication Chair	Patrick Paroubek (France)
Student Workshop Chair	EACL Student Board (M. Gabsdil, J. Hockenmaier, J. Herring)
Workshop Chair	Steven Krauwer (The Netherlands)
Local Organisation Chair	Ferenc Kiefer (Hungary)

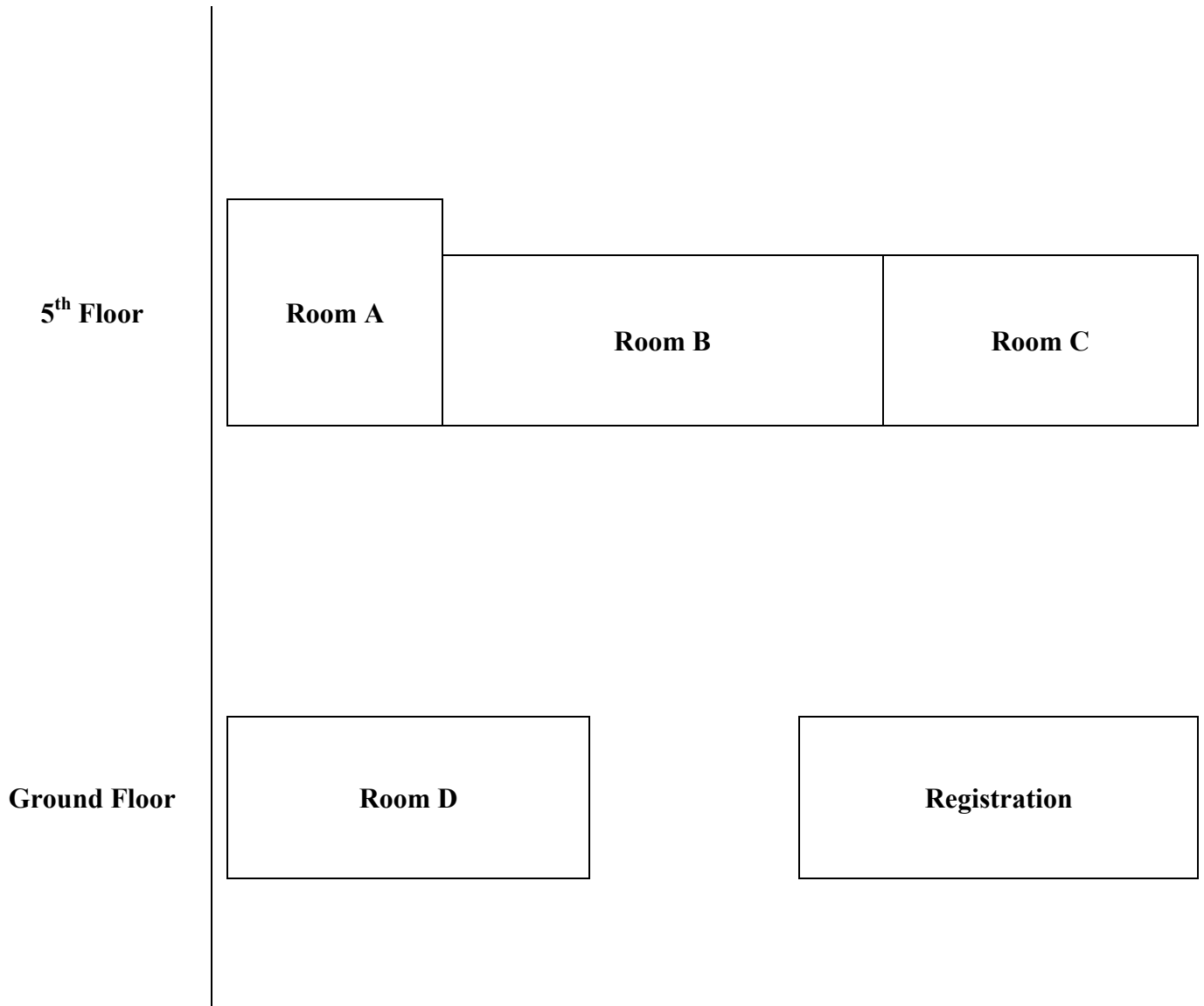
The Conference, the Workshops and the Tutorials are sponsored by:



Contents

Main Conference Meeting Rooms in Hotel Agro	2
Main Conference Programme-at-a-Glance	3
Main Conference Programme	4
Day 1: Tuesday, 15 April 2003	4
Day 2: Wednesday, 16 April 2003	6
Day 3: Thursday, 17 April 2003	7
General Information	9
Registration and Information Desk	9
Book Exhibition	9
Meals	9
Badge	9
Computer Room, Internet Access, Voltage	9
Information for Authors	9
Official Language	9
Social Events	10
Getting to the Conference Site	10
Tickets	10
Bank, Currency, Credit Cards	10
Insurance	10

Main Conference Meeting Rooms in Hotel Agro



Main Conference Programme-at-a-Glance

Tuesday, 15 April 2003

	Room A	Room B	Room C
09.00 – 09.30	Introduction		
09.30 – 10.30	Invited Talk		
10.30 – 11.00	<i>Coffee break</i>		
11.00 – 12.30	Parsing	Machine Learning	Machine Translation
12.30 – 13.50	<i>Lunch</i>		
13.50 – 14.50	RN: IE	RN: Corpora and Treebank	Demos 1 / Room D
14.50 – 15.00	<i>Mini break</i>		
15.00 – 16.00	RN: NLG	RN: Lexical Resources 1	Demos 2 / Room D
16.00 – 16.30	<i>Coffee break</i>		
16.30 – 18.00	Parsing	Compositional Semantics	Machine Translation

Wednesday, 16 April 2003

09.00 – 10.30	Machine Translation and Named Entities	Grammar Formalism	Lexicon
10.30 – 11.00	<i>Coffee break</i>		
11.00 – 12.30	Student Research Workshop: Track 1	Student Research Workshop: Track 2	
12.30 - 13.50	<i>Lunch</i>		
12.30 - 13.50	Student Business Meeting		
13.50 – 14.35	EACL Business Meeting		
14.40 – 16.00	RN: Grammar and Syntax	RN: Speech and Machine Translation	RN: Dictionary
16.00 – 16.30	<i>Coffee break</i>		
16.30 – 18.00	Prosody and Information Structure	Algorithms	Classification

Thursday, 17 April 2003

09.00 – 10.00	RN: Q/A Summarization	RN: IE and Metaphor	Demos 3 / Room D
10.00 – 10.30	<i>Coffee break</i>		
10.30 – 12.30	POS Tagging and Corpus Tools	Dialogue	Lexicon
12.30 – 14.00	<i>Lunch</i>		
14.00 – 15.00	RN: Lexical Resources 2	RN: PoS Tagging and Parsing	Demos 4 / Room D
15.00 – 15.30	<i>Coffee break</i>		
15.30 – 16.30	Invited Talk		
16.30 – 17.00	Closing and Best Paper Award		

Main Conference Programme

Tuesday, 15 April 2003

Day 1

	Room A	Room B	Room C
09.00 09.30	Introduction		
09.30 10.30	<p style="text-align: center;">Invited Talk</p> <p>Multilingual Access to Large Spoken Archives <i>Doug Oard, University of Maryland</i></p>		
	<i>Coffee break</i>		
	Parsing	Machine Learning	Machine Translation
11.00 12.30	<p>Neural Network Probability Estimation for Broad Coverage Parsing <i>James Henderson</i></p> <p>An Efficient Implementation of a New DOP Model <i>Rens Bod</i></p> <p>Bootstrapping Statistical Parsers from Small Datasets Mark Steedman, Anoop Sarkar, Miles Osborne, Rebecca Hwa, Stephen Clark, Julia Hockenmaier, Paul Ruhlen, Steve Baker, Jeremiah Crim</p>	<p>Toward Evaluation of Writing Style: Overly Repetitious Word Use <i>Jill Burstein, Magdalena Wolska</i></p> <p>French Amalgam: A Quick Adaptation of a Sentence Realization System to French <i>Martine Smets, Michael Gamon, Simon Corston-Oliver, Eric Ringger</i></p> <p>Combining Distributional and Morphological Information for Part of Speech Induction <i>Alexander Clark</i></p>	<p>Czech-English Dependency Tree-based Machine Translation <i>Jan Cuvrin, Martin Cmejrek, Jiri Havelka</i></p> <p>Using POS Information for SMT into Morphologically Rich Languages <i>Nicola Ueffing, Hermann Ney</i></p> <p>Automatic Evaluation for a Palpable Measure of a Speech Translation System's Capability <i>Keiji Yasuda, Fumiaki Sugaya, Toshiyuki Takezawa, Seiichi Yamamoto, Masuzo Yanagida</i></p>
12.30 13.50	<i>Lunch</i>		
	Research Notes and Demonstrations		
	IE	Corpora & Treebank	Demos 1 / Room D
13.50 14.10	<p>Mining Web Sites Using Unsupervised Adaptive Information Extraction <i>Alexiei Dingli, Fabio Ciravegna, David Guthrie, Yorick Wilks</i></p>	<p>Annotated Hungarian National Corpus <i>Zoltán Alexin et al</i></p>	<p>Talking through Procedures: An Intelligent Space Station Procedure Assistant <i>Greg Aist et al</i></p> <p>A Dialogue System with Contextually Appropriate Spoken Output Intonation <i>Ivana Kruijff-Korbayova, Kepa Joseba Rodriguez, Elena Karagjosova, Stina Ericsson</i></p>
14.10 14.30	<p>Effect of Utilizing Terminology on Extraction of Protein-Protein Interaction Information from Biomedical Literature <i>Junko Hosaka, Judice Koh, Akihiko Konagaya</i></p>	<p>Arabic Syntactic Trees: from Constituency to Dependency <i>Zdenek Zabokrtsky, Otakar Smrz</i></p>	<p>WASPBENCH: A Lexicographer's Workbench Incorporating State-of-the-Art Word Sense Disambiguation <i>Adam Kilgarriff, Roger Evans, Rob Koeling, David Tugwell</i></p> <p>An Open-Source Environment for Compiling Typed Unification Grammars into Speech Recognisers <i>Manny Rayner, Beth Ann Hockey, John Dowding</i></p> <p>A Cross Language Document Retrieval System Based on Semantic Annotation <i>Bogdan Sacaleanu, Paul Buitelaar, Martin Volk</i></p>
14.30 14.50	<p>The Multilingual Named Entity Recognition Framework <i>Thierry Poibeau</i></p>	<p>The Corpora Management System Based on Java and Oracle Technologies Serge A. Yablonsky</p>	<p>Horacio Saggion, Kalina Robust Generic and Query-based Summarization <i>Kalina Bontcheva, Hamish Cunningham</i></p>
	<i>Minibreak</i>		

Main Conference Programme

	Room A	Room B	Room C
	Research Notes and Demonstrations		
	NLG	Lexical Resources 1	Demos 2 / Room D
15.00 15.20	A Best-First Search Algorithm for Generating Referring Expressions <i>Helmut Horacek</i>	Beyond Lexical Units: Enriching WordNets with Phrasets <i>Luisa Bentivogli, Emanuele Pianta</i>	WYSIWYM - Building User Interfaces with Natural Language Feedback <i>Roger Evans, Richard Power</i> Towards an Adaptive Communication Aid with Ambiguous Text Input <i>Karin Harbusch, Michael Kühn</i> Multilingual Adaptations of a Reusable Information Extraction Tool
15.20 15.40	A Flexible Pragmatics-Driven Language Generator for Animated Agents <i>Paul Piwek</i>	Discovering Corpus-Specific Word Senses <i>Beate Dorow, Dominic Widdows</i>	<i>Diana Maynard, Hamish Cunningham, Kalina Bontcheva</i> AGORA. Multilingual Multiplatform Architecture for the Development of Natural Language Voice Services <i>Jose Relación-Gil, M. Carmen Rodríguez-Gancedo, Luis Hernandez Gomez</i> Event-Coreference across Multiple, Multi-lingual Sources in the Mumis Project
15.40 16.00	Summarizing Neonatal Time Series Data <i>Somayajulu G. Sripada, Ehud Reiter, Jim Hunter, Jin Yu</i>	Creating a Multilingual Collocations Dictionary from Large Text Corpora <i>Luka Nerima, Violeta Seretan, Eric Wehrli</i>	<i>Horacio Saggion, Hamish Cunningham, Jan Kuper, Thierry Declerck, Peter Wittenburg</i> Question Answering Using Web News as Knowledge Base <i>Zhiping Zheng</i>
	<i>Coffee break</i>		
	Parsing	Compositional Semantics	Machine Translation
16.30 18.00	Using Grammatical Relations to Compare Parsers <i>Judita Preiss</i> AVM Description Compilation using Types as Modes <i>Gerald Penn</i> Describing Syntax with Star-Free Regular Expressions <i>Anssi Yli-Jyrä</i>	Underspecification Formalisms: Hole Semantics as Dominance Constraints, <i>Alexander Koller, Joachim Niehren, Stefan Thater</i> Well-Nested Parallelism Constraints for Ellipsis Resolution <i>Katrin Erk, Joachim Niehren</i> Semantic Construction in F-TAG <i>Claire Gardent, Laura Kallmeyer</i>	Effect of Cross-Language IR in Bilingual Lexicon Acquisition from Comparable Corpora <i>Takehito Utsuro, Takashi Horiuchi, Takeshi Hamamoto, Kohei Hino, Takeaki Nakayama</i> Combining Clues for Word Alignment <i>Jörg Tiedemann</i> Automatic Construction of Machine Translation Knowledge Using Translation Literalness <i>Kenji Imamura, Eiichiro Sumita, Yuji Matsumoto</i>

Main Conference Programme

Wednesday, 16 April 2003

Day 2

	Room A	Room B	Room C
	Machine Translation and Named Entities	Grammar Formalisms	Lexicon
09.00 10.30	Efficient Search for Interactive Statistical Machine Translation <i>Richard Zens, Franz Josef Och, Hermann Ney</i>	Rigid Grammars in the Associative-Commutative Lambek Calculus are not Learnable <i>Christophe Costa Florencio</i>	Evaluating and Combining Approaches to Selectional Preference Acquisition <i>Carsten Brockmann, Mirella Lapata</i>
	Learning Translations of Named-Entity Phrases from Parallel Corpora <i>Robert C. Moore</i>	Multi-Modal Combinatory Categorical Grammar <i>Geert-Jan M. Kruijff, Jason Baldrige</i>	Experiments on the Choice of Features for Learning Verb Classes <i>Sabine Schulte im Walde</i>
	Named Entity Recognition For Catalan Using Only Spanish Resources and Unlabelled Data <i>Xavier Carreras, Lluís Màrquez, Lluís Padró</i>	Topological Parsing <i>Gerald Penn, Mohammad Haji-Abdolhosseini</i>	A General Feature Space for Automatic Verb Classification <i>Eric Joanis, Suzanne Stevenson</i>
	<i>Coffee break</i>		
	Student Research Workshops		
	Track 1	Track 2	
11.00 12.30	Learning to Identify Fragmented Words in Spoken Discourse <i>Piroska Lendvai</i>	Integrating Cohesion and Coherence for Automatic Summarization <i>Laura Alonso Alemany, Maria Fuentes Fort</i>	
	Clustering Adjectives for Class Discovery <i>Gemma Boleda, Laura Alonso Alemany</i>	Focusing on Scenario Recognition in Information Extraction <i>Milena Yankova, Svetla Boytcheva</i>	
	A Dynamic Logic Formalisation of the Dialogue Gameboard <i>Raquel Fernandez</i>	Reversing Controlled Document Authoring to Normalize Documents <i>Aurélien Max</i>	
12.30 13.50	<i>Lunch</i>		
12.30 13.50	Student Business Meeting		
13.50 14.35	EACL Business Meeting		
	Research Notes and Demonstrations		
	Grammar & Syntax	Speech & Machine Translation	Dictionary
14.40 15.00	Lexicon Acquisition with a Large-Coverage Unification-Based Grammar <i>Frederik Fouvry</i>	Meaningful Conversation with a Mobile Robot <i>Johan Bos, Ewan Klein, Tetsushi Oka</i>	Contents and Evaluation of the First Slovenian-German Online Dictionary <i>Birte Lönneker, Primo Jakopin</i>
15.00 15.20	The Key Role of Semantics in the Development of Large-Scale Grammars of Natural Language <i>Valia Kordoni</i>	Rule-Based Lexical Modelling of Foreign-Accented Pronunciation Variants <i>Stefan Schaden</i>	Oxford Dictionary of English - Current Developments <i>James McCracken, Adam Kilgarriff</i>
15.20 15.40	Lexicalized Grammar Acquisition <i>Yusuke Miyao, Takashi Ninomiya, Jun'ichi Tsujii</i>	A Corpus-Centered Approach to Spoken Language Translation <i>Eiichiro Sumita et al</i>	
15.40 16.00	A Machine Learning Approach to the Identification of WH Gaps <i>Derrick Higgins</i>	Using Noisy Bilingual Data for Statistical Machine Translation <i>Stephan Vogel</i>	
	<i>Coffee break</i>		

Main Conference Programme

	Room A	Room B	Room C
	Prosody and Information Structure	Algorithms	Classification
16.30 18.00	<p>Learning PP Attachment for Filtering Prosodic Phrasing <i>Olga van Herwijnen, Antal van den Bosch, Jacques Terken, Erwin Marsi</i></p> <p>Information Structure in Topological Dependency Grammar <i>Geert-Jan M. Kruijff, Denys Duchier</i></p> <p>Producing Contextually Appropriate Intonation in an Information-State Based Dialogue System <i>Ivana Kruijff-Korbayova, Kepa Joseba Rodriguez, Stina Ericsson, Elena Karagjosova</i></p>	<p>Constraint Based Integration of Deep and Shallow Parsing Techniques <i>Michael Daum, Kilian A. Foth, Wolfgang Menzel</i></p> <p>Comparison of Alignment Templates and Maximum Entropy Models for NLP <i>Oliver Bender, Klaus Macherey, Franz Josef Och, Hermann Ney</i></p> <p>Linear Text Segmentation Using a Dynamic Programming Algorithm <i>Athanasios Kehagias, Pavlina Fragkou, Vassilios Petridis</i></p>	<p>Language Independent Authorship Attribution with Character Level N-Grams <i>Fuchun Peng, Dale Schuurmans, Vlado Keselj, Shaojun Wang</i></p> <p>Classifying the Hungarian Web <i>Andras Kornai, Marc Krellenstein, Michael Mulligan, David Twomey, Fruzsina Veress, Alec Wysoker</i></p> <p>A Comparison of Event Models for Naive Bayes Anti-Spam E-Mail Filtering <i>Karl-Michael Schneider</i></p>

Thursday, 17 April 2003

Day 3

	Room A	Room B	Room B
Research Notes and Demonstrations			
	Q/A Summarization	IE & Metaphor	Demos 3 / Room D
09.00 09.20	<p>WEBCOOP: A Cooperative Question Answering System on the Web <i>Farah Benamara, Patrick Saint Dizier</i></p>	<p>Automatic Acquisition of Script Knowledge from a Text Collection <i>Toshiaki Fujiki, Hidetsugu Nanba, Manabu Okumura</i></p>	<p>Suregen-2: A Shell System for the Generation of Clinical Documents <i>Dirk Kraus</i></p> <p>Integrating Natural Language Generation with the Apache Cocoon XML Server <i>Graham Wilcock</i></p> <p>A Flexible Pragmatics-Driven Language Generator for Animated Agents <i>Paul Piwek</i></p> <p>Development of Corpora within the CLaRK System: The BulTreeBank Project Experience <i>Kiril Simov et al</i></p> <p>The Corpora Management System Based on Java and Oracle Technology <i>Serge A. Yablonsky</i></p>
09.20 09.40	<p>How to Build a QA System in Your Back-Garden: Application for Romanian <i>Constantin Orasan, Doina Tatar, Gabriela Seban, Dana Avram, Adrian Onet</i></p>	<p>NLP for Indexing and Retrieval of Captioned Photographs <i>Horacio Saggion, Katerina Pastra, Yorick Wilks</i></p>	
09.40 10.00	<p>CAST: A Computer-Aided Summarisation Tool <i>Constantin Orasan, Ruslan Mitkov, Laura Hasler</i></p>	<p>Domain-Transcending Mappings in a System for Metaphorical Reasoning <i>John Barnden, Sheila Glasbey, Mark Lee, Alan Wallington</i></p>	
<i>Coffee break</i>			

Main Conference Programme

	Room A	Room B	Room C
	POS Tagging and Corpus Tools	Dialogue	Lexicon
10.30 12.30	<p>Detecting Errors in Part-of-Speech Annotation <i>Markus Dickinson, Detmar Meurers</i></p> <p>Investigating GIS and Smoothing for Maximum Entropy Taggers <i>James R. Curran, Stephen Clark</i></p> <p>Finite Structure Query: A Tool for Querying Syntactically Annotated Corpora <i>Stephan Kepser</i></p> <p>An Integrated Term-Based Corpus Query System <i>Kostas Manios, Goran Nenadic, Irena Spasic, Sophia Ananiadou</i></p>	<p>A Model of Back-Channel Acknowledgements in Spoken Dialogue <i>Nicola Cathcart, Jean Carletta, Ewan Klein</i></p> <p>The Role of Initiative in Tutorial Dialogue <i>Mark G. Core, Johanna D. Moore, Claus Zinn</i></p> <p>Targeted Help for Spoken Dialogue Systems <i>Beth Ann Hockey, Oliver Lemon, Ellen Campana, Laura Hiatt, Gregory Aist, Jim Hieronymus, John Dowding</i></p> <p>Transparent Combination of Rule-Based and Data-Driven Approaches in Speech Understanding <i>Manny Rayner, Beth Ann Hockey</i></p>	<p>QUALIFIER: Question Answering by Lexical Fabric and External Resources <i>Hui Yang, Tat-Seng Chua</i></p> <p>Generalised PP-attachment Disambiguation Using Corpus-based Linguistic Diagnostics <i>Paola Merlo</i></p> <p>Detecting Novel Compounds: The Role of Distributional Evidence <i>Mirella Lapata, Alex Lascarides</i></p> <p>Empirical Methods for Compound Splitting <i>Philipp Koehn, Kevin Knight</i></p>
12.30 14.00	<i>Lunch</i>		
14.00 15.00	Research Notes and Demonstrations		
	Lexical Resources 2	PoS Tagging & Parsing	Demos 4 / Room D
14.00 14.20	Experiments on Candidate Data for Collocation Extraction <i>Stefan Evert, Hannah Kermes</i>	Categorial Fluidity in Chinese and its Implications for Part-of-speech Tagging <i>Oi Yee Kwong</i>	Controlled Authoring of Biological Experiment Reports <i>Caroline Brun, Marc Dymetman, Eric Fanchon, Stanislas Lhomme</i>
14.20 14.40	Word Classification Based on Combined Measures of Distributional and Semantic Similarity <i>Viktor Pekar, Steffen Staab</i>	PEAS, the First Instantiation of a Comparative Framework for Evaluating Parsers of French <i>Véronique Gendner, Gabriel Illouz, Michèle Jardino, Laura Monceaux, Patrick Paroubek, Isabelle Robba, Anne Vilnat</i>	CarSim: An Automatic 3D Text-to-Scene Conversion System Applied to Road Accident Reports <i>Ola Åkerberg, Hans Svensson, Bastian Schulz, Pierre Nugues</i>
14.40 15.00	Interactive Word Alignment for Language Engineering <i>Lars Ahrenberg, Magnus Merkel, Michael Petterstedt</i>	A Cascaded Finite-State Parser for German <i>Michael Schiehlen</i>	The Multilingual Named Entity Recognition Framework <i>Thierry Poibeau</i>
	<i>Coffee break</i>		
15.30 16.30	Invited talk Linguistic Variation and Computation <i>John Nerbonne</i>		Creating a Multilingual Collocations Dictionary from Large Text Corpora <i>Violeta Seretan</i>
16.30 17.00	Closing and best paper award		

General Information

Registration and Information Desk

The Registration and Information Desk will operate in the Hotel Agro Conference Center as follows:

Friday, 11 April	08:00-12:00
Saturday, 12 April	08:00-17:00
Sunday, 13 April	08:00-17:00
Monday, 14 April	08:00-17:00
Tuesday, 15 April	08:00-17:00
Wednesday, 16 April	08:00-17:00
Thursday, 17 April	08:00-14:00

Any questions regarding the Conference logistics, the programme and assistance of speakers can be directed to the Registration and Information Desk.

Book Exhibition

The book exhibition takes place on the 5th floor and is open during the Conference hours on the Main Conference (15 - 17 April).

Meals

- Breakfast is included in the hotel price.
- Optional sandwich lunches are available on April 12 - 17 (3 sandwiches, refreshment, fruit, coffee) and are served in the Restaurant of the Hotel Agro Conference Center. Each ticket costs HUF 2000/lunch. In case you wish to order the lunch on-site, you are requested to do so as early as possible but not later than the previous day. Please, note that the availability of tickets is limited.
- Coffee, tea and refreshments are served during the breaks.

Badge

Every conference event you participate in is indicated with badges in different colour in your badge holder (Tutorials: green, Workshops on 13 April: red, Workshops on 14 April: blue, Main Conference: white). Badges need to be changed accordingly. In case of badge loss, we charge EUR 10 for the substitution.

Computer Room, Internet Access, Voltage

A number of PCs and plugs for your laptop with Internet access are available:

- 12-17 April: 5th Floor corridor.
- 15-17 April: Room D on the Ground Floor (Except for Demos hours).

The electricity supply in Hungary is 230 V AC (50 Hz).

Information for Authors

For the benefit of both the audience and the next speaker, all the speakers are kindly requested to strictly adhere to the time limit that otherwise will be enforced by session chairs.

Data beamers, overhead projectors and flip charts will be available in the session rooms. Authors are kindly asked to use their own laptops. In addition, computers are also available in the session rooms during the Main Conference.

Official Language

The official language of the Conference is English.

General Information

Social Events

- **Welcome Reception** (19.00, Monday, April 14) in the Restaurant of Hotel Agro Conference Center. All registered participants of the Main Conference are cordially invited. (*Please, wear your white-colour badges.*)
- **Banquet** (19.30, Wednesday, April 16): Cruising on the Danube you can enjoy the evening colours and night-lights of Budapest with live music and dance. The conference banquet will be held on the Europa Boat (address: Budai Alsó rakpart, Batthyany Dezső tér). Special bus transfer will take the participants from Hotel Agro to the Europa Boat and back. The bus departs at 19.00 from Hotel Agro Conference Center. The banquet costs HUF 12 500.

Getting to the City Center

From Hotel Agro the City can be reached by bus #21 (written in black!). The bus stop is in a 2-3 minute walking distance at Hotel Olympia. It takes you to Moszkva tér (square), where you can choose either tram #6 or tram #4 or Metro line 2 (red) to get into the City center.

Tickets for Public Transportation

Tickets are available at METRO stations, news stands or major bus/tram termini and must be validated inside the vehicles or at the entrance of the METRO. Please, keep your ticket with you until the end of the journey/exit (Metro).

Tickets for public transportation can also be bought at the Registration and Information Desk.

Telephone Card

Telephone cards are available at the Registration and Information Desk.

Bank, Currency, Credit Cards

The unit of currency is the Hungarian Forint (HUF), denoted as “Ft” by the Hungarians. International credit cards (AmEx, EC/MC, Visa) are accepted at most hotels, restaurants and shops. ATMs are available all over the city (1 EUR is approx. 250 HUF).

Insurance

The Organizers of the Conference do not provide insurance and do not take responsibility for any loss, accident or illness that might occur during the Conference or in the course of travel to or from the meeting site. It is, therefore, the responsibility of the participants to check their coverage with their insurance provider.