

Invited Talks

Wednesday, 1st April 2009

	TITLE	LECTURER	PLACE	TIME
Tik1	<i>Slacker Semantics: Why Superficiality, Dependency and Avoidance of Commitment Can Be the Right Way to Go</i>	Ann Copestake (University of Cambridge, UK)	Alexandra Trianti Hall	09:30 -10:30

Friday, 3rd April 2009

	TITLE	LECTURER	PLACE	TIME
Tik2	<i>NLP and the Humanities: The Revival of an Old Liaison</i>	Franciska de Jong (University of Twente, the Netherlands)	Nikos Skalkotas Hall	09:30 -10:30

LECTURER	TITLE - ABSTRACT
TIK1 Ann Copestake (University of Cambridge, UK)	<p data-bbox="616 248 1445 304">Slacker Semantics: Why Superficiality, Dependency and Avoidance of Commitment Can Be the Right Way to Go</p> <p data-bbox="616 338 1445 667">Standard linguistic approaches to compositional semantics need adaptation for use in broadcoverage computational processing. I'll argue that an ideal compositional semantic representation should capture all and only the information available from syntax and productive morphology, allowing for shallow as well as deep syntactic analysis and avoiding over-commitment when ambiguity cannot be resolved. This leads to 'surfacy' semantic representations which must nevertheless allow for enrichment by deeper analysis (including lexical semantics and anaphora resolution), support (robust) inference and have a logically-sound interpretation. Other desiderata are compatibility with generation, statistical ranking of analyses/realisations and distributional semantics, ease of use for consumers of the representation and for human annotators, and cross-linguistic adequacy.</p> <p data-bbox="616 701 1445 947">Several current approaches to computational compositional semantics meet at least some of these goals, but in this talk I'll concentrate on work within the DELPH-IN community using Minimal Recursion Semantics (MRS) and Robust MRS. I'll give an overview to illustrate the extent to which I believe we're currently meeting these aims. I'll go on to show that, under certain assumptions about the grammar, an interconversion is possible between MRS and a variable-free semantic dependency notation (DMRS), which removes redundancy, supports additional forms of underspecification and makes the representation easier for humans to process.</p>
TIK2 Franciska de Jong (University of Twente, the Netherlands)	<p data-bbox="616 969 1244 992">NLP and the Humanities: The Revival of an Old Liaison</p> <p data-bbox="616 1025 1445 1384">The humanities and the field of natural language processing (NLP) have always had common playgrounds. The liaison was never constrained to linguistics; also philosophical, philological and literary studies have had their impact on NLP, and there have always been dedicated conferences and journals for the humanities and the NLP community. The more recent emergence of the field of Computational Humanities demonstrates that the potential for mutual impact has gained in strength and diversity and that important lessons can be learned for other application areas than the humanities. A renewed liaison with the now computational humanities can help NLP to set up an innovative research agenda which covers a wide range of topics including semantic analysis, integration of multimodal information, language-based interaction, performance evaluation, service models, and usability studies.</p> <p data-bbox="616 1417 1445 1697">As will be discussed, the further and combined exploration of these topics will help to develop an infrastructure that will allow content and data driven research domains in the humanities to renew their field and to exploit the potential coming from the wide-scale digitization initiative. To name a few: art history, media studies, oral history, archeology, archiving studies have needs that can be served in novel ways by the matured branches that NLP offers today. Based on an overview of initiatives, an analysis of the characteristics of the interaction between the humanities and NLP, and some telling examples, it will be outlined what could be done to increase the chances for a bright future for the old ties, and how other domains can benefit as well.</p>

Tutorials

Monday, 30 March 2009

	TITLE	LECTURER	PLACE	TIME
Tut1	<i>Constraint Conditional models for NLP</i>	Ming-Wei Chang Lev Ratinov Dan Roth	CSI-Room 1*	09:00 -12:30
Tut2	<i>Statistical Language Modeling for Information Access</i>	Maarten de Rijke Edgar Meij	CSI-Room 1	14:30 -18:00

Tuesday, 31 March 2009

	TITLE	LECTURER	PLACE	TIME
Tut3	<i>Reinforcement Learning for Adaptive Dialogue Systems</i>	Oliver Lemon Verena Rieser	CSI-Room 1	09:00 -12:30
Tut4	<i>Combinatory Categorical Grammars for Robust Natural Language Processing</i>	Mark Steedman	CSI-Room 1	14:30 -18:00

* CSI-Room 1: Conference Suite 1 – Room 1

LECTURER TITLE - ABSTRACT**Tut1** **Ming-Wei
Chang,
Lev Ratinov,
Dan Roth****Constraint Conditional models for NLP**

Making decisions in natural language processing problems often involves assigning values to sets of interdependent variables where the expressive dependency structure can influence, or even dictate, what assignments are possible. This setting is of particular significance in structured learning problems such as semantic role labeling, named entity and relation recognition, co-reference resolution, transliteration, summarization and machine translation, but the approach has a broader set of applications such as textual entailment and question answering. In all these cases, it is natural to either formulate the decision process as a constrained optimization problem, or to break up the complex problem into a set of subproblems and require solutions to be consistent modulo (soft, possibly) constraints. In both cases, the resulting objective function is composed of learned models, subject to domain or problem specific constraints.

Constrained Conditional Models is a learning and inference framework that refers to augmenting the learning of conditional (probabilistic or discriminative) models with declarative constraints (written, for example, using a first-order representation) as a way to support decisions in an expressive output space while maintaining modularity and tractability of training and inference. Models of this kind have recently attracted much attention within the NLP community.

Formulating problems as constrained optimization problems over the output of learned models has several advantages. It allows one to focus on the modeling of problems by providing the opportunity to incorporate problem specific global constraints using a first order language, freeing the developer from (much of the) low level feature engineering, and it can also guarantee exact inference. It provides the freedom of decoupling the stage of model generation (learning) from that of the constrained inference stage, often resulting in simplifying the learning stage and the engineering problem of building an NLP system, while improving the quality of the solutions.

The primary goal of this tutorial is to introduce the framework of Constrained Conditional Models (CCMs) to the broader ACL community, motivate it as a generic framework for learning and inference in global NLP decision problems, present some of the key theoretical and practical issues involved in using CCMs and survey some of the existing applications of it as a way to promote further development of the framework and additional applications. The tutorial will thus be useful for many of the senior and junior researchers that have interest in global decision problems in NLP, providing a concise overview of recent perspectives and research results.

Tut2 **Maarten de
Rijke,
Edgar Meij****Statistical Language Modeling for Information Access**

About a decade ago, statistical language models were introduced in information retrieval (IR). They have proved to be an attractive and effective framework for IR, partly due to their well-defined statistical properties which are built on solid theoretical foundations. Since their introduction to IR, many new models, techniques, and applications have emerged. Language models have, for example, been effectively applied in tasks such as question answering, cross-lingual IR, expert finding, retrieving semi-structured information, topic tracking and detection. Other developments have included leveraging various sources of information, such as feedback documents, thesauri, external corpora, and syntactical clues for the estimation of a query model.

The purpose of this tutorial is to systematically explain the use of statistical language models in information retrieval with an emphasis on the underlying principles and framework, empirically effective models, as well as language models developed for a broad range of retrieval tasks, both traditional and non-traditional, including semi-structured document retrieval, expert finding, question answering, cross-language IR, blog retrieval, topic detection and tracking. Participants can expect to learn the major principles and methods of applying statistical language models to a range of information retrieval tasks, outstanding problems in this area, and to obtain comprehensive pointers to the research literature and available toolkits.

No background in information retrieval is required, but some basic familiarity with statistics and data-driven approaches to language processing is assumed.

LECTURER TITLE - ABSTRACT

Tut3	Oliver Lemon, Verena Rieser	Reinforcement Learning for Adaptive Dialogue Systems <p>This tutorial will provide an overview to the rapidly growing field of Reinforcement Learning for automatic dialogue system development.</p> <p>Designing a spoken dialogue system can be a time-consuming and challenging process. One of the main problems is to design the dialogue management strategy -- what the system should say next, in a particular context. The standard approach to this is to hand-craft a finite state machine or a rule-based strategy, but this method is not data-driven, and has no measurable performance guarantees.</p> <p>To facilitate dialogue strategy development, recent research investigates the use of Reinforcement Learning (RL) methods applied to automatic dialogue strategy optimisation from corpora. In this course we introduce the basics of RL and how to (practically) apply it to adaptive dialogue system development. We describe each step of the development cycle -- from data collection to reward modelling to training and evaluation/user testing. We also demonstrate tools for RL and discuss practical issues, such as corpus requirements, to get people started.</p> <p>In particular, this tutorial will cover:</p> <ul style="list-style-type: none">• A brief introduction to system development and the use of RL.• An outline some of the existing methods and challenges, discussing current work in the field.• Simulation-based training methods, such as simulated users and reward modelling.• Practical examples from dialogue strategy learning and Natural Language Generation from the CLASSiC project (Computational Learning in Adaptive Systems for Spoken Conversation, http://www.classic-project.org/ funded under EC FP7, Call 1).• Tools
-------------	--	---

Tut4	Mark Steedman	Combinatory Categorical Grammars for Robust Natural Language Processing <p>Combinatory Categorical Grammar (CCG) is a "radically lexicalized" grammar formalism, in which all language-specific grammatical information, including subcategorization and specification of construction and word-order, resides in the lexicon. A small number of rule-types, corresponding semantically to the simplest operators of the Combinatory applicative calculus of Curry and Feys, project lexical items monotonically onto derivations for the strings and meaning representations of the language. An intended consequence is a generalized notion of surface syntactic constituency.</p> <p>The statistical modeling techniques developed in the last decade have made wide-coverage CCG parsing (Hockenmaier and Steedman 2007; Clark and Curran 2007) an attractive basis for a number of practical applications in NLP, because of its principled separation between the language-specific lexicon and the mechanism for projecting that lexicon onto sentences including long-range, so-called "non-projective," dependencies. These applications include spoken natural language generation (SNLG, Baker et al. 2004), question answering (QA, Bos et al. 2005), semantically based grammar induction (Zettlemoyer and Collins 2005), and statistical machine translation (SMT, Birch et al. 2007). A number of tools for developing CCG applications are available from the Linguistic Data Consortium (LDC, http://www ldc.upenn.edu/) and OpenCCG (http://openccg.sourceforge.net).</p> <p>This body of work is quite recent, and CCG is a strikingly different formalism from standard computational approaches to grammar. The tutorial will make the approach and the tools more widely accessible to a broad natural language processing linguistic audience, and will also bring those already familiar with the approach up to speed with more recent theoretical and computational developments, including:</p> <ul style="list-style-type: none">• CCG grammars in comparison to other formalisms, illustrated with a variety of constructions using examples from the Penn treebank, drawing on recent work with Baldridge and Bozsahin; questions of expressive power and efficiency;• The problem of inducing a large CCG lexicon and parsing model from a suitably transformed version of the Penn treebank, considering both generative and discriminative models, and the role of "supertagging", drawing on recent work with Hockenmaier, Clark and Curran;• Open research questions on the problem of porting parsers to new domains and to new languages with limited amounts of labeled data, and on generalizing treebank grammars and parsing models using unlabeled text.
-------------	--------------------------	---

Workshops

Monday, 30 March 2009

	WORKSHOP TITLE	ORGANIZERS	PLACE	
W01	<i>The Interaction between Linguistics and Computational Linguistics: Virtuous, Vicious or Vacuous?</i>	Timothy Baldwin Valia Kordoni	MC2	Full day
W03	<i>Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education (LaTeCH-SHELT&R 2009)</i>	Piroska Lendvai Lars Borin Antal van den Bosch Martin Reynaert Caroline Sporleder	MC3.2	Full day
W04	<i>4th Workshop on Statistical Machine Translation (WMT 2009)</i>	Chris Callison-Burch Philipp Koehn Josh Schroeder Christof Monz	Skalkotas	2-day
W05	<i>2nd International Workshop on Semantic Representation of Spoken Language 2009 (SRSL 2009)</i>	Thierry Declerck Manuel Alcantara-Pla	MC3.3	Full day
W06	<i>12th European Workshop on Natural Language Generation (ENLG 2009)</i>	Emiel Krahmer Mariet Theune	MC3	2-day
W10	<i>Computational Linguistic Aspects of Grammatical Inference</i>	Menno van Zaanen Colin de la Higuera	MC3.4	Full day

Tuesday, 31 March 2009

	WORKSHOP TITLE	ORGANIZERS	PLACE	
W02	<i>Geometric Models of Natural Language Semantics (GEMS 2009)</i>	Roberto Basili Marco Pennacchiotti	MC2	Full day
W04	<i>4th Workshop on Statistical Machine Translation (WMT 2009)</i>	Chris Callison-Burch Philipp Koehn Josh Schroeder Christof Monz	Skalkotas	2-day
W06	<i>12th European Workshop on Natural Language Generation (ENLG 2009)</i>	Emiel Krahmer Mariet Theune	MC3	2-day
W07	<i>Language Technologies for African Languages (AfLaT 2009)</i>	Lori Levin John Kiango Judith Klavans Manuela Noske Guy De Pauw Gilles-Maurice de Schryver Peter Waiganjo Wagacha	MC3.4	Full day
W08	<i>Computational Approaches to Semitic Languages</i>	Mike Rosner Shuly Wintner	MC3.3	Full day
W09	<i>Cognitive Aspects of Computational Language Acquisition</i>	Thierry Poibeau Afra Alishahi Aline Villavicencio	MC3.2	Full day

Conference Program

Workshops (30 & 31 March 2009)
Main Conference (1 to 3 April 2009)
 Technical Sessions
 Student Research Workshop
 System Demonstrations

Monday, 30 March 2009 – Workshops

Morning Session

	W01. Workshop on the Interaction between Linguistics and Computational Linguistics: Virtuous, Vicious or Vacuous? (MC2 Room)	W03. Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education (LaTeCH - SHELT&R 2009) (MC3.2 Room)	W04. Fourth Workshop on Statistical Machine Translation (WMT 2009) (Skalkotas Hall)
09:00-10:30	<p><i>Machine Translation and its Philosophical Accounts</i> Stelios Piperidis</p> <p><i>The Annotation Conundrum</i> Mark Liberman</p>	<p>Opening</p> <p>Invited Talk by Martin Doerr</p> <p>Moderated Discussion</p>	<p>Opening Remarks</p> <p>Overview of the Shared Tasks</p> <p><i>Findings of the 2009 Workshop on Statistical Machine Translation</i> Chris Callison-Burch, Philipp Koehn, Christof Monz and Josh Schroeder</p>
10:30-11:00	Coffee break		Booster Session
11:00-12:30	<p><i>How the Statistical Revolution Changes (Computational) Linguistics</i> Mark Johnson</p> <p>Discussion</p>	<p><i>Content Analysis of Museum Documentation in a Transdisciplinary Perspective</i> Guenther Goerz and Martin Scholz</p> <p><i>An Intelligent Authoring Environment for Abstract Semantic Representations of Cultural Object Descriptions</i> Stasinios Konstantopoulos, Vangelis Karkaletsis and Dimitris Bilidas</p> <p><i>Multiple Sequence Alignments in Linguistics</i> Jelena Prokić, Martijn Wieling and John Nerbonne</p> <p><i>Evaluating the Pairwise String Alignment of Pronunciations</i> Martijn Wieling, Jelena Prokić and John Nerbonne</p>	<p>Poster Session: Shared Evaluation Task</p> <p>Syntax-Oriented Evaluation Measures for Machine Translation Output Maja Popovic and Hermann Ney</p> <p><i>A Simple Automatic MT Evaluation Metric</i> Petr Homola, Vladislav Kubon and Pavel Pecina</p> <p><i>Machine Translation Evaluation with Textual Entailment Features</i> Sebastian Pado, Michel Galley, Daniel Jurafsky and Christopher D. Manning</p> <p>Poster Session: Shared System Combinations</p> <p><i>Combining Multi-Engine Translations with Moses</i> Yu Chen, Michael Jellinghaus, Andreas Eisele, Yi Zhang, Sabine Hunsicker, Silke Theison, Christian Federmann and Hans Uszkoreit</p> <p><i>CMU System Combination for WMT'09</i> Almut Silja Hildebrand and Stephan Vogel</p> <p><i>The RWTH System Combination System for WMT 2009</i> Gregor Leusch, Evgeny Matusov and Hermann Ney</p> <p><i>Machine Translation System Combination with Flexible Word Ordering</i> Kenneth Heafield, Greg Hanneman and Alon Lavie</p> <p><i>Incremental Hypothesis Alignment with Flexible Matching for Building Confusion Networks: BBN System Description for WMT09 System Combination Task</i> Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz</p> <p>Invited Talk by Martin Kay</p>
12:30-14:00	Lunch		

Monday, 30 March 2009 - Workshops

Afternoon Session

	W01. Workshop on the Interaction between Linguistics and Computational Linguistics: Virtuous, Vicious or Vacuous?	W03. Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education (LaTeCH - SHELT&R 2009)	W04. Fourth Workshop on Statistical Machine Translation (WMT 2009)
	(MC2 Room)	(MC3.2 Room)	(Skalkotas Hall)
14:00-16:00	<p><i>Computational Linguistics and Generative Linguistics: The Triumph of Hope over Experience</i> Geoffrey Pullum</p> <p>Panel and Discussion</p> <p><i>Linguistics in Computational Linguistics: Observations and Predictions</i> Hans Uszkoreit</p> <p><i>Linguistically Naïve != Language Independent: Why NLP Needs Linguistic Typology</i> Emily M. Bender</p> <p><i>Parsed Corpora for Linguistics</i> Gertjan van Noord</p> <p><i>Computational Linguistics and Linguistics: What Keeps Them together, What Sets Them apart?</i> Gregor Erbach</p> <p><i>What Do Computational Linguists Need to Know about Linguistics?</i> Robert C. Moore</p>	<p>Invited Talk by Tamás Váradi</p> <p><i>A Web-Enabled and Speech-Enhanced Parallel Corpus of Greek-Bulgarian Cultural Texts</i> Voula Giouli, Nikos Glaros, Kiril Simov and Petya Osenova</p> <p><i>The Development of the "Index Thomisticus" Treebank Valency Lexicon</i> Barbara McGillivray and Marco Passarotti</p> <p><i>Applying NLP Technologies to the Collection and Enrichment of Language Data on the Web to Aid Linguistic Research</i> Fei Xia and William Lewis</p>	<p>Panel Discussion</p> <p>Shared Task: Translation</p> <p>Booster Sesssion</p> <p>Poster Session (poster titles and authors on page 10)</p>
16:00-16:30	Coffee break		
16:30-18:00	<p><i>The Interaction of Syntactic Theory and Computational Psycholinguistics</i> Frank Keller</p> <p>Discussion</p> <p>Closing Remarks</p>	<p><i>Instance-Driven Discovery of Ontological Relation Labels</i> Marieke van Erp, Antal van den Bosch, Sander Wubben and Steve Hunt</p> <p><i>The Role of Metadata in the Longevity of Cultural Heritage Resources</i> Milena Dobрева and Nikola Ikonov</p> <p>Moderated Discussion, Closing</p>	

W04. Fourth Workshop on Statistical Machine Translation (WMT 2009)

Shared Task: Translation

Poster Session

The RWTH Machine Translation System for WMT 2009

Maja Popovic, David Vilar, Daniel Stein, Evgeny Matusov and Hermann Ney

Translation Combination using Factored Word Substitution

Christian Federmann, Silke Theison, Andreas Eisele, Hans Uszkoreit, Yu Chen, Michael Jellinghaus and Sabine Hunsicker

NUS at WMT09: Domain Adaptation Experiments for English-Spanish Machine Translation of News Commentary Text

Preslav Nakov and Hwee Tou Ng

The Universität Karlsruhe Translation System for the EACL-WMT 2009

Jan Niehues, Teresa Herrmann, Muntsin Kolss and Alex Waibel

The TALP-UPC Phrase-Based Translation System for EACL-WMT 2009

José A. R. Fonollosa, Maxim Khalilov, Marta R. Costa-jussá, José B. Mariño, Carlos A. Henríguez Q., Adolfo Hernández H. and Rafael E. Banchs

Deep Linguistic Multilingual Translation and Bilingual Dictionaries

Eric Wehrli, Luka Nerima and Yves Scherrer

MATREX: The DCU MT System for WMT 2009

Jinhua Du, Yifan He, Sergio Penkale and Andy Way

LIMSI's Statistical Translation Systems for WMT'09

Alexandre Allauzen, Josep Crego, Aurélien Max and François Yvon

NICT@WMT09: Model Adaptation and Transliteration for Spanish-English SMT

Michael Paul, Andrew Finch and Eiichiro Sumita

Statistical Post Editing and Dictionary Extraction: Systran/Edinburgh submissions for ACL-WMT2009

Loïc Dugast, Jean Senellart and Philipp Koehn

Experiments in Morphosyntactic Processing for Translating to and from German

Alexander Fraser

Improving Alignment for SMT by Reordering and Augmenting the Training Corpus

Maria Holmqvist, Sara Stymne, Jody Foo and Lars Ahrenberg

English-Czech MT in 2008

Ondrej Bojar, David Marecek, Václav Novák, Martin Popel, Jan Ptáček, Jan Rouš and Zdenek Zabokrtsk\y

SMT and SPE Machine Translation Systems for WMT'09

Holger Schwenk, Sadaf Abdul Rauf, Loic Barrault and Jean Senellart

Joshua: An Open Source Toolkit for Parsing-Based Machine Translation

Zhifei Li, Chris Callison-Burch, Chris Dyer, Sanjeev Khudanpur, Lane Schwartz, Wren Thornton, Jonathan Weese and Omar Zaidan

An Improved Statistical Transfer System for French-English Machine Translation

Greg Hanneman, Vamshi Ambati, Jonathan H. Clark, Alok Parlikar and Alon Lavie

The University of Maryland Statistical Machine Translation System for the Fourth Workshop on Machine Translation

Chris Dyer, Hendra Setiawan, Yuval Marton and Philip Resnik

Toward Using Morphology in French-English Phrase-Based SMT

Marine Carpuat

MorphoLogic's Submission for the WMT 2009 Shared Task

Attila Novák

Edinburgh's Submission to all Tracks of the WMT2009 Shared Task with Reordering and Speed Improvements to Moses

Philipp Koehn and Barry Haddow

Monday, 30 March 2009 – Workshops

Morning Session

	W05. Workshop on Semantic Representation of Spoken Language (SRSL 2009)	W06. 12th European Workshop on Natural Language Generation (ENLG 2009)	W10. Workshop on Computational Linguistic Aspects of Grammatical Inference (CLAGI 2009)
	(MC3.3 Room)	(MC3 Room)	(MC3.4 Room)
09:00-10:30	<p><i>Extreme-case formulations in Cypriot Greek</i> Maria Christodoulidou</p> <p><i>On the Segmentation of Requests in Spoken Language</i> Michael Alvarez-Pereyre</p> <p><i>Identifying Segment Topics in Medical Dictations</i> Johannes Matiassek, Jeremy Jancsary, Alexandra Klein and Harald Trost</p>	<p><i>Using NLG to Help Language-Impaired Users Tell Stories and Participate in Social Dialogues</i> Ehud Reiter, Norman Alm, Rolf Black, Martin Dempster, Ross Turner and Annalu Waller</p> <p><i>Towards a Generation-Based Semantic Web Authoring Tool</i> Richard Power</p> <p><i>System Building Cost vs. Output Quality in Data-to-Text Generation</i> Anja Belz and Eric Kow</p>	<p>Opening Remarks Menno van Zaanen and Colin de la Higuera</p> <p>Invited Talk by Damir Ćavar, on <i>Bootstrapping of Linguistic Features for Bootstrapping Grammars</i></p>
10:30-11:00	Coffee break		
11:00-12:30	<p><i>Semantic Representation of Non-Sentential Utterances in Dialog</i> Silvie Cinkova</p> <p><i>Annotating Spoken Dialogs: from Speech Segments to Dialog Acts and Frame Semantics</i> Marco Dinarelli, Silvia Quarteroni, Sara Tonelli, Alessandro Moschitti and Giuseppe Riccardi</p> <p><i>Predicting Concept Types in User Corrections in Dialog</i> Svetlana Stoyanchev and Amanda Stent</p>	<p><i>Is Sentence Compression an NLG Task?</i> Erwin Marsi, Emiel Kraemer, Iris hendrickx and Walter Daelemans</p> <p><i>Probabilistic Approaches for Modeling Text Structure and their Application to Text-to-Text Generation</i> Invited speaker: Regina Barzilay (MIT)</p>	<p><i>Dialogue Act Prediction Using Stochastic Context-Free Grammar Induction</i> Jeroen Geertzen</p> <p><i>Experiments Using OSTIA for a Language Production Task</i> Dana Angluin and Leonor Becerra-Bonache</p> <p><i>GREAT: A Finite-State Machine Translation Toolkit Implementing a Grammatical Inference Approach for Transducer Inference (GIATI)</i> Jorge González and Francisco Casacuberta</p>
12:30-14:00	Lunch		

Monday, 30 March 2009 - Workshops

Afternoon Session

	W05. Workshop on Semantic Representation of Spoken Language (SRSL 2009) (MC3.3 Room)	W06. 12th European Workshop on Natural Language Generation (ENLG 2009) (MC3 Room)	W10. Workshop on Computational Linguistic Aspects of Grammatical Inference (CLAGI 2009) (MC3.4 Room)
14:00-16:00	<p><i>Deeper Spoken Language Understanding for Man-machine Dialogue on Broader Application Domains: a Logical Alternative to Concept Spotting</i> Jeanne Villaneau and Jean-Yves Antoine</p> <p><i>An Integrated Approach to Robust Processing of Situated Spoken Dialogue</i> Pierre Lison and Geert-Jan M. Kruijff</p> <p><i>RUBISC - a Robust Unification-Based Incremental Semantic Chunker</i> Michaela Atterer and David Schlangen</p> <p><i>Incrementality, Speaker-Hearer Switching and the Disambiguation Challenge</i> Ruth Kempson, Eleni Gregoromichelaki and Yo Sato</p>	<p><i>Distinguishable Entities: Definitions and Properties</i> Monique Rolbert and Pascal Pr�ea</p> <p><i>Generating Approximate Geographic Descriptions</i> Ross Turner, Yaji Sripada and Ehud Reiter</p> <p><i>Class-Based Ordering of Prenominal Modifiers</i> Margaret Mitchell</p> <p><i>Referring Expression Generation through Attribute-Based Heuristics</i> Robert Dale and Jette Viethen</p>	<p><i>A Note on Contextual Binary Feature Grammars</i> Alexander Clark, Remi Eyraud and Amaury Habrard</p> <p><i>Language Models for Contextual Error Detection and Correction</i> Herman Stehouwer and Menno van Zaanen</p> <p><i>On Statistical Parsing of French with Supervised and Semi-Supervised Strategies</i> Marie-H�el�ene Candito, Benoit Crabb�e and Djam�e Seddah</p> <p><i>Upper Bounds for Unsupervised Parsing with Unambiguous Non-Terminally Separated Grammars</i> Franco M. Luque and Gabriel Infante-Lopez</p>
16:00-16:30	Coffee break		
16:30-18:00	<p>Summary Thierry Declerck / Manuel Alcantara</p> <p>Discussion and outlook</p>	<p>Introduction to the Generation Challenges 2009</p> <p><i>Report on the First NLG Challenge on Generating Instructions in Virtual Environments (GIVE)</i> Donna Byron, Alexander Koller, Kristina Striegnitz, Kristina Striegnitz, Robert Dale, Johanna Moore and Jon Oberlander</p> <p><i>The TUNA-REG Challenge 2009: Overview and Evaluation Results</i> Albert Gatt, Anja Belz and Eric Kow</p> <p>Discussion</p>	<p><i>Comparing learners for Boolean Partitions: Implications for Morphological Paradigms</i> Katya Pertsova</p> <p>Discussion (Panel)</p>

Tuesday, 31 March 2009 - Workshops

Morning Session

	W02. Workshop on GEMS - GEometric Models of Natural Language Semantics (GEMS 2009) (MC2 Room)	W04. Fourth Workshop on Statistical Machine Translation (WMT 2009) (Skalkotas Hall)	W06. 12th European Workshop on Natural Language Generation (ENLG 2009) (MC3 Room)
09:00-10:30	<p>Invited Talk: Patrick Pantel <i>"Lexical Semantics in the Prime Time: Applications to Web Search"</i></p> <p><i>One distributional memory, many semantic spaces</i> Marco Baroni and Alessandro Lenci</p> <p><i>Word Space Models of Lexical Variation</i> Yves Peirsman and Dirk Speelman</p>	<p>Full Papers Session 1: Use of training data</p> <p><i>Mining a Comparable Text Corpus for a Vietnamese-French Statistical Machine Translation System</i> Thi Ngoc Diep Do, Viet Bac Le, Brigitte Bigi, Laurent Besacier and Eric Castelli</p> <p><i>Improving Arabic-Chinese Statistical Machine Translation using English as Pivot Language</i> Nizar Habash and Jun Hu</p> <p><i>Domain Adaptation for Statistical Machine Translation with Monolingual Resources</i> Nicola Bertoldi and Marcello Federico</p>	<p><i>A Model for Human Readable Instruction Generation Using Level-Based Discourse Planning and Dynamic Inference of Attributes</i> Daniel Dionne, Salvador de la Puente, Carlos León, Pablo Gervás and Raquel Hervás</p> <p><i>Learning Lexical Alignment Policies for Generating Referring Expressions for Spoken Dialogue Systems</i> Srinivasan Janarthanam and Oliver Lemon</p> <p><i>An Alignment-capable Microplanner for Natural Language Generation</i> Hendrik Buschmeier, Kirsten Bergmann and Stefan Kopp</p>
10:30-11:00	Coffee break		
11:00-12:30	<p><i>Unsupervised Classification with Dependency Based Word Spaces</i> Klaus Rothenhausler and Hinrich Schuetze</p> <p><i>A Study of Convolution Tree Kernel with Local Alignment</i> Lidan Zhang and Kwok-Ping Chan</p> <p><i>BagPack: A general framework to represent semantic relations</i> Amaç Herdagdelen and Marco Baroni</p> <p><i>Positioning for Conceptual Development using Latent Semantic Analysis</i> Fridolin Wild, Bernhard Hoisl and Gaston Burek</p> <p><i>Semantic similarity: which is the best method to compute it? Extrinsic evaluation based on the choice of distractors in multiple-choice tests</i> Ruslan Mitkov, Le An Ha, Andrea Varga and Luz Rello</p>	<p>Full Papers Session 2: Reordering</p> <p><i>Chinese Syntactic Reordering for Adequate Generation of Korean Verbal Phrases in Chinese-to-Korean SMT</i> Jin-Ji Li, Jungi Kim, Dong-II Kim and Jong-Hyeok Lee</p> <p><i>A Quantitative Analysis of Reordering Phenomena</i> Alexandra Birch, Phil Blunsom and Miles Osborne</p> <p><i>A POS-Based Model for Long-Range Reorderings in SMT</i> Jan Niehues and Muntsin Kolss</p>	<p>Poster session (poster titles and authors on page 16)</p>
12:30-14:00	Lunch		

Tuesday, 31 March 2009 - Workshops

Afternoon Session

	W02. Workshop on GEMS - GEometric Models of Natural Language Semantics (GEMS 2009) (MC2 Room)	W04. Fourth Workshop on Statistical Machine Translation (WMT 2009) (Skalkotas Hall)	W06. 12th European Workshop on Natural Language Generation (ENLG 2009) (MC3 Room)
14:00-16:00	<p><i>Paraphrase assessment in structured vector space: Exploring parameters and datasets</i> Katrin Erk and Sebastian Pado</p> <p><i>SVD Feature Selection for Probabilistic Taxonomy Learning</i> Francesca Fallucchi and Fabio Massimo Zanzotto</p> <p><i>Unsupervised and Constrained Dirichlet Process Mixture Models for Verb Clustering</i> Andreas Vlachos, Anna Korhonen and Zoubin Ghahramani</p> <p><i>A Non-negative Tensor Factorization Model for Selectional Preference Induction</i> Tim Van de Cruys</p> <p><i>A Graph-Theoretic Algorithm for Automatic Extension of Translation Lexicons</i> Beate Dorow, Florian Laws, Lukas Michelbacher, C. Scheible and J. Utt</p> <p><i>Handling Sparsity for Verb Noun Multi-Word Expression Token Classification</i> Mona Diab and Madhav Krishna</p>	<p>Full Papers Session 3: Linguistic modeling</p> <p><i>Disambiguating "DE" for Chinese-English Machine Translation</i> Pi-Chuan Chang, Daniel Jurafsky and Christopher D. Manning</p> <p><i>A Systematic Analysis of Translation Model Search Spaces</i> Michael Auli, Adam Lopez, Hieu Hoang and Philipp Koehn</p> <p><i>A Deep Learning Approach to Machine Transliteration</i> Thomas Deselaers, Saša Hasan, Oliver Bender and Hermann Ney</p>	<p><i>Investigating Content Selection for Language Generation using Machine Learning</i> Colin Kelly, Ann Copestake and Nikiforos Karamanis</p> <p><i>Generating Clausal Coordinate Ellipsis Multilingually: A Uniform Approach Based on Postediting</i> Karin Harbusch and Gerard Kempen</p> <p><i>Towards Empirical Evaluation of Affective Tactical NLG</i> Ielka van der Sluis and Chris Mellish</p>
16:00-16:30	Coffee break		
16:30-18:00	<p><i>Semantic Density Analysis: Comparing word meaning across time and phonetic space</i> Eyal Sagi, Stefan Kaufmann and Brady Clark</p> <p><i>Context-theoretic Semantics for Natural Language: an Overview</i> Daoud Clarke</p> <p>Panel</p> <p>Best Paper</p>	<p>Full Papers Session 4: Error metrics and tuning</p> <p><i>Stabilizing Minimum Error Rate Training</i> George Foster and Roland Kuhn</p> <p><i>On the Robustness of Syntactic and Semantic Features for Automatic MT Evaluation</i> Jesús Giménez and Lluís Màrquez</p> <p><i>Fluency, Adequacy, or HTER? Exploring Different Human Judgments with a Tunable MT Metric</i> Matthew Snover, Nitin Madnani, Bonnie Dorr and Richard Schwartz</p>	<p><i>What game theory can do for NLG: The case of vague language</i> Invited speaker: Kees van Deemter (University of Aberdeen)</p> <p>Final words / Closing of ENLG 2009</p>

W06. 12th European Workshop on Natural Language Generation (ENLG 2009)

Posters

A Wizard-of-Oz Environment to Study Referring Expression Generation in a Situated Spoken Dialogue Task

Srinivasan Janarthanam and Oliver Lemon

The Effect of Linguistic Devices in Information Presentation Messages on Recall and Comprehension

Martin I. Tietze, Andi Winterboer and Johanna D. Moore

A Situated Context Model for Resolution and Generation of Referring Expressions

Hendrik Zender, Geert-Jan M. Kruijff and Ivana Kruijff-Korbayova

A Hearer-oriented Evaluation of Referring Expression Generation

Imtiaz Hussain Khan, Kees van Deemter, Graeme Ritchie, Albert Gatt and Alexandra A. Cleland

Generating Natural Language Descriptions of Ontology Concepts

Niels Schütte

Clustering and Matching Headlines for Automatic Paraphrase Acquisition

Sander Wubben, Antal van den Bosch, Emiel Kraemer and Erwin Marsi

Towards a Game-Theoretic Approach to Content Determination

Ralf Klabunde

Collection and Analysis of Referring Expressions Used in Collaboration

Philipp Spanger, Yasuhara Masaaki, Tokunaga Takenobu and Iida Ryu

Precision and Mathematical Form in First and Subsequent Mentions of Numerical Facts and their Relation to Document Structure

Sandra Williams and Richard Power

SimpleNLG: A realisation engine for practical applications

Albert Gatt and Ehud Reiter

Generation Challenges Posters

Realizing the Costs: Template-Based Surface Realisation in the GRAPH Approach to Referring Expression Generation

Ivo Brugman, Mariët Theune, Emiel Kraemer and Jette Viethen

Generation of Referring Expression with an Individual Imprint

Bernd Bohnet

Evolutionary and Case-Based Approaches to REG: NIL-UCM-EvoTAP, NIL-UCM-ValuesCBR and NIL-UCM-EvoCBR

Raquel Hervás and Pablo Gervás:

USP-EACH: Improved Frequency-based Greedy Attribute Selection

Diego Jesus de Lucena and Ivandré Paraboni

A Probabilistic Model of Referring Expressions for Complex Objects

Kotaro Funakoshi, Philipp Spanger, Mikio Nakano and Takenobu Tokunaga

Finding landmarks or counting steps: two strategies for navigation instructions in virtual environments

Kristina Striegnitz and Filip Majda

Instruction Giving based on Graphical World Abstractions

Daniel Dionne, Salvador de la Puente, Carlos León, Raquel Hervás, Pablo Gervás

Two Approaches to GIVE: Dynamic Level Adaptation versus Playfulness

Roan Boer Rookhuiszen, Michel Obbink, Mariët Theune

Tuesday, 31 March 2009 - Workshops

Morning Session

	W07. Workshop on Language Technologies for African Languages (AfLaT 2009) (MC3.4 Room)	W08. Workshop on Computational Approaches to Semitic Languages (MC3.3 Room)	W09. Workshop on Cognitive Aspects of Computational Language Acquisition (MC3.2 Room)
09:00-10:30	<p>Invited Talk: «<i>African Language Families and their Structural Properties</i>» by Sonja Bosch</p> <p><u>Note:</u> poster session is open at the coffee break (poster titles and authors on page 20)</p>	<p>Opening + Keynote</p> <p><i>Finite State Morphology for Ethiopian Semitic languages</i> Prof. Michael Gasser, Indiana University</p> <p><i>How to Establish a Verbal Paradigm on the Basis of Ancient Syriac Manuscripts</i> Wido van Peursen</p>	<p>Welcome</p> <p>Invited Talk: «<i>Conceptual Descriptions: Evidence from Corpora, the Mind, and the Brain</i>» Massimo Poesio (University of Essex and University of Trento)</p>
10:30-11:00	Coffee break		
11:00-12:30	<p><i>Collecting and Evaluating Speech Recognition Corpora for Nine Southern Bantu Languages</i> Jaco Badenhorst, Charl Van Heerden, Marelle Davel and Etienne Barnard</p> <p><i>The SAWA Corpus: a Parallel Corpus English – Swahili</i> Guy De Pauw, Peter Waiganjo Wagacha and Gilles-Maurice de Schryver</p> <p><i>Information Structure in African Languages: Corpora and Tools</i> Christian Chiarcos, Ines Fiedler, Mira Grubic, Andreas Haida, Katharina Hartmann, Julia Ritz, Anne Schwarz, Amir Zeldes and Malte Zimmermann</p> <p><u>Note:</u> poster session is open at the lunch break (poster titles and authors on page 20)</p>	<p><i>The Karamel System and Semitic Languages: Structured Multi-Tiered Morphology</i> Francois Barthelemy</p> <p><i>Revisiting Multi-tape Automata for Semitic Morphological Analysis and Generation</i> Mans Hulden</p> <p><i>A Hybrid Approach for Building Arabic Diacritizer</i> Khaled Shaalan, Hitham M. Abo Bakr and Ibrahim Ziedan</p>	<p><i>Towards a formal view of corrective feedback</i> Staffan Larsson and Robin Cooper</p> <p><i>A Collaborative Tool for the Computational Modelling of Child Language Acquisition</i> Kris Jack</p> <p><i>What's in a Message?</i> Stergos Afantenos and Nicolas Hernandez</p>
12:30-14:00	Lunch		

Tuesday, 31 March 2009 - Workshops

Afternoon Session

	W07. Workshop on Language Technologies for African Languages (AfLaT 2009) (MC3.4 Room)	W08. Workshop on Computational Approaches to Semitic Languages (MC3.3 Room)	W09. Workshop on Cognitive Aspects of Computational Language Acquisition (MC3.2 Room)
14:00-16:00	<p><i>A Computational Approach to Yorùbá Morphology</i> Raphael Finkel and Odetunji Ajadi Odejobi</p> <p><i>Using Technology Transfer to Advance Automatic Lemmatisation for Setswana</i> Hendrik Johannes Groenewald</p> <p><i>Part-of-Speech Tagging of Northern Sotho: Disambiguating Polysemous Function Words</i> Gertrud Faaß, Ulrich Heid, Elsabé Taljard and Daan J. Prinsloo</p> <p><i>Development of an Amharic Text-to-Speech System Using Cepstral Method</i> Tadesse Anberbir and Tomio Takara</p> <p><i>Note: poster session is open at the coffee break (poster titles and authors on page 20)</i></p>	<p><i>Unsupervised Concept Discovery In Hebrew Using Simple Unsupervised Word Prefix Segmentation for Hebrew and Arabic</i> Elad Dinur, Dmitry Davidov and Ari Rappoport</p> <p><i>Automatic Treebank-Based Acquisition of Arabic LFG Dependency Structures</i> Lamia Tounsi, Mohammed Attia and Josef van Genabith</p> <p><i>Spoken Arabic Dialect Identification Using Phonotactic Modeling</i> Fadi Biadisy, Julia Hirschberg and Nizar Habash</p>	<p>Invited Talk: <i>"Treebank Parsing and Knowledge of Language: A Cognitive Perspective"</i> Robert Berwick (Massachusetts Institute of Technology, Joint work with Prof. Sandiway Fong, University of Arizona)</p> <p><i>Another look at indirect negative evidence</i> Alexander Clark and Shalom Lappin</p> <p><i>Categorizing Local Contexts as a Step in Grammatical Category Induction</i> Markus Dickinson and Charles Jochim</p>
16:00-16:30	Coffee break		
16:30-18:00	<p><i>Building Capacities in Human Language Technology for African Languages</i> Tunde Adegbola</p> <p><i>Initial Fieldwork for LWAZI: A Telephone-Based Spoken Dialog System for Rural South Africa</i> Tebogo Gumede and Madelaine Plauché</p> <p>Discussion</p>	<p><i>Structure-based evaluation of an Arabic semantic Query Expansion using the JIRS system</i> Lahsen Abouenour, Karim Bouzoubaa and Paolo Rosso</p> <p><i>Syntactic Reordering for English-Arabic Phrase-Based Machine Translation</i> Jakob Elming and Nizar Habash</p> <p>Closing</p> <p>Discussion</p>	<p><i>Darwinised Data-Oriented Parsing - Statistical NLP with added sex and death</i> Dave Cochran</p> <p><i>Language Diversity across the Consonant Inventories: A Study in the Framework of Complex Networks</i> Monojit Choudhury, Animesh Mukherjee, Anupam Basu, Niloy Ganguly, Ashish Garg and Vaibhav Jalan</p>

W07. Workshop on Language Technologies for African Languages (AfLaT 2009)

Poster Session (During Coffee and Lunch Breaks)

Setswana Tokenisation and Computational Verb Morphology: Facing the Challenge of a Disjunctive Orthography
Rigardt Pretorius, Ansu Berg, Laurette Pretorius and Biffie Viljoen

Interlinear Glossing and its Role in Theoretical and Descriptive Studies of African and other Lesser-Documented Languages
Dorothee Beermann and Pavel Mihaylov

Towards an Electronic Dictionary of Tamajaq Language in Niger
Chantal Enguehard and Issouf Modi

A Repository of Free Lexical Resources for African Languages: The Project and the Method
Piotr Bański and Beata Wójtowicz

Exploiting Cross-linguistic Similarities in Zulu and Xhosa Computational Morphology
Laurette Pretorius and Sonja Bosch

Methods for Amharic Part-of-Speech Tagging Björn Gambäck, Fredrik Olsson, Atelach Alemu Argaw and Lars Asker

An Ontology for Accessing Transcription Systems (OATS)
Steven Moran

Wednesday, 1 April 2009 – Main Conference

Morning Session

09:00–09:30	Plenary Session: Opening (Trianti Hall)			
09:30–10:30	Plenary Session: Invited Talk (Trianti Hall)			
	Ann Copestake, University of Cambridge <i>Slacker Semantics: Why Superficiality, Dependency and Avoidance of Commitment Can Be the Right Way to Go</i>			
10:30–11:00	Break (Muses Foyer & Skalkotas Foyer)			
11:00–11:30	T1. Semantics <u>Chair:</u> Robin Cooper (Skalkotas Hall) <i>Deriving Generalized Knowledge from Corpora Using WordNet Abstraction</i> Benjamin Van Durme, Phillip Michalak, and Lenhart Schubert	T2. Parsing <u>Chair:</u> Giorgio Satta (MC2 Room) <i>Parsing Mildly Non-Projective Dependency Structures</i> Carlos Gómez-Rodríguez, David Weir and John Carroll	T3. Phonology <u>Chair:</u> Veronique Hoste (MC3 Room) <i>Discovering Global Patterns in Linguistic Networks through Spectral Analysis: A Case Study of the Consonant Inventories</i> Animesh Mukherjee, Monojit Choudhury and Ravi Kannan	T4. Translation <u>Chair:</u> Dan Roth (Trianti Hall) <i>Correcting Automatic Translations through Collaborations between MT and Monolingual Target-Language Users</i> Joshua Albrecht, Rebecca Hwa and G. Elisabeta Marai
11:30–12:30	Q1. Generation, Summarization and Discourse <u>Chair:</u> Johanna Moore (Skalkotas Hall) <i>Text Summarization Model based on Maximum Coverage Problem and its Variant</i> Hiroya Takamura and Manabu Okumura <i>Company-Oriented Extractive Summarization of Financial News</i> Katja Filippova, Mihai Surdeanu, Massimiliano Ciaramita and Hugo Zaragoza <i>Predicting the Fluency of Text with Shallow Structural Features: Case Studies of Machine Translation and Human-Written Text</i> Jieun Chae and Ani Nenkova <i>Optimization in Coreference Resolution is not Needed: A Nearly-Optimal Algorithm with Intensional Constraints</i> Manfred Klenner and Étienne Ailloud <i>Cognitively Motivated Features for Readability Assessment</i> Lijun Feng, Noémie Elhadad and Matt Huenerfauth	Q2. Syntax and Parsing <u>Chair:</u> David Weir (MC2 Room) <i>Treebank Grammar Techniques for Non-Projective Dependency Parsing</i> Marco Kuhlmann and Giorgio Satta <i>Incremental Parsing Models for Dialog Task Structure</i> Srinivas Bangalore and Amanda Stent <i>Rich Bitext Projection Features for Parse Reranking</i> Alexander Fraser, Renjing Wang and Hinrich Schuetze <i>Correcting Dependency Annotation Errors</i> Markus Dickinson <i>Learning Efficient Parsing</i> Gertjan van Noord <i>Deterministic Shift-Reduce Parsing for Unification-Based Grammars by Using Default Unification</i> Takashi Ninomiya, Takuya Matsuzaki, Nobuyuki Shimizu and Hiroshi Nakagawa	Q3. Morphology, Tagging and Word Segmentation <u>Chair:</u> Eva Hajicova (MC3 Room) <i>TBL-Improved Non-Deterministic Segmentation and POS Tagging for a Chinese Parser</i> Martin Forst and Ji Fang <i>Weakly Supervised Part-of-Speech Tagging for Morphologically-Rich, Resource-Scarce Languages</i> Kazi Saidul Hasan and Vincent Ng <i>Character-Level Dependencies in Chinese: Usefulness and Learning</i> Hai Zhao <i>Correcting a PoS-Tagged Corpus Using Three Complementary Methods</i> Hrafn Loftsson <i>Semitic Morphological Analysis and Generation Using Finite State Transducers with Feature Structures</i> Michael Gasser	Q4. Lexical Semantics <u>Chair:</u> Eneko Agirre (Trianti Hall) <i>An Empirical Study on Class-Based Word Sense Disambiguation</i> Rubén Izquierdo, Armando Suárez and German Rigau <i>Data-Driven Semantic Analysis for Multilingual WSD and Lexical Selection in Translation</i> Marianna Apidianaki <i>Predicting Strong Associations on the Basis of Corpus Data</i> Yves Peirsman and Dirk Geeraerts <i>Measuring Frame Relatedness</i> Marco Pennacchiotti and Michael Wirth
12:30–14:00	Lunch (Muses Foyer)			

Wednesday, 1 April 2009 – Main Conference

Afternoon Session

14:00–15:00	Poster Session (Muses Foyer)			
	P1. Generation, Summarization and Discourse See Q1.	P2. Syntax and Parsing See Q2.	P3. Morphology, Tagging and Word Segmentation See Q3	P4. Lexical Semantics See Q4.
15:00–16:00	T5. Information Extraction <u>Chair:</u> Claire Grover (Skalkotas Hall) <i>Automatic Single-Document Key Fact Extraction from Newswire Articles</i> Itamar Kastner and Christof Monz <i>Structural, Transitive and Latent Models for Biographic Fact Extraction</i> Nikesh Garera and David Yarowsky	T6. Summarization <u>Chair:</u> Noemie Elhadad (MC2 Room) <i>Sentiment Summarization: Evaluating and Learning User Preferences</i> Kevin Lerman, Sasha Blair-Goldensohn and Ryan McDonald <i>Performance Confidence Estimation for Automatic Summarization</i> Annie Louis and Ani Nenkova	T7. Speech <u>Chair:</u> Srinivas Bangalore (MC3 Room) <i>Lattice Parsing to Integrate Speech Recognition and Rule-Based Machine Translation</i> Selçuk Köprü and Adnan Yazici <i>Re-Ranking Models for Spoken Language Understanding</i> Marco Dinarelli, Alessandro Moschitti and Giuseppe Riccardi	T8. Translation <u>Chair:</u> Mike Rosner (Trianti Hall) <i>On the Use of Comparable Corpora to Improve SMT Performance</i> Sadaf Abdul-Rauf and Holger Schwenk <i>N-Gram-Based Statistical Machine Translation versus Syntax Augmented Machine Translation: Comparison and System Combination</i> Maxim Khalilov and José A. R. Fonollosa
16:00–16:30	Break (Muses Foyer & Skalkotas Foyer)			
16:30–18:00	T9. Lexical Semantics <u>Chair:</u> Hans Uszkoreit (Skalkotas Hall) <i>Personalizing PageRank for Word Sense Disambiguation</i> Eneko Agirre and Aitor Soroa <i>Supervised Domain Adaption for WSD</i> Eneko Agirre and Oier Lopez de Lacalle <i>Using Cycles and Quasi-Cycles to Disambiguate Dictionary Glosses</i> Roberto Navigli	T10. Generation <u>Chair:</u> Richard Power (MC2 Room) <i>Natural Language Generation as Planning Under Uncertainty for Spoken Dialogue Systems</i> Verena Rieser and Oliver Lemon <i>Human Evaluation of a German Surface Realisation Ranker</i> Aoife Cahill and Martin Forst <i>Improving Grammaticality in Statistical Sentence Generation: Introducing a Dependency Spanning Tree Algorithm with an Argument Satisfaction Model</i> Stephen Wan, Mark Dras, Robert Dale and Cécile Paris	T11. Named Entity Recognition <u>Chair:</u> Menno van Zaanen (MC3 Room) <i>Clique-Based Clustering for Improving Named Entity Recognition Systems</i> Julien Ah-Pine and Guillaume Jacquet <i>Analysing Wikipedia and Gold-Standard Corpora for NER Training</i> Joel Nothman, Tara Murphy and James R. Curran <i>Learning-Based Named Entity Recognition for Morphologically-Rich, Resource-Scarce Languages</i> Kazi Saidul Hasan, Md. Altaf ur Rahman and Vincent Ng	T12. Translation <u>Chair:</u> Rebecca Hwa (Trianti Hall) <i>MINT: A Method for Effective and Scalable Mining of Bilingual Named Entity Transliterations from Large Comparable Corpora</i> Raghavendra Udupa, K Saravanan, A Kumaran and Jagadeesh Jagarlamudi <i>Rule Filtering by Pattern for Efficient Hierarchical Translation</i> Gonzalo Iglesias, Adrià de Gispert, Eduardo R. Banga and William Byrne <i>Lightly Supervised Transliteration for Machine Translation</i> Amit Kirschenbaum and Shuly Wintner

Thursday, 2 April 2009 – Main Conference

Morning Session

09:00–10:30				
Student Research Workshop				
<p>Track 1 (Trianti Hall)</p> <p><i>A Chain-starting Classifier of Definite NPs in Spanish</i> Marta Recasens</p> <p><i>Structural Correspondence Learning for Parse Disambiguation</i> Barbara Plank</p> <p><i>Extraction of definitions using grammar-enhanced machine learning</i> Eline Westerhout</p>	<p>Track 2 (Skalkotas Hall)</p> <p><i>Aligning Medical Domain Ontologies for Clinical Query Extraction</i> Pinar Wennerberg</p> <p><i>A Generalized Vector Space Model for Text Retrieval Based on Semantic Relatedness</i> George Tsatsaronis and Vicky Panagiotopoulou</p> <p><i>Finding Word Substitutions Using a Distributional Similarity Baseline and Immediate Context Overlap</i> Aurelie Herbelot</p>	<p>Track 3 (MC3 Room)</p> <p><i>Speech Emotion Recognition with TGI+ classifier</i> Julia Sidorova</p> <p><i>A Memory-Based Approach to the Treatment of Serial Verb Construction in Combinatory Categorical Grammar</i> Prachya Boonkwan</p> <p><i>A Comparison of Merging Strategies for Translation of German Compounds</i> Sara Stymne</p>	<p>Track 4 (MC2 Room)</p> <p><i>Combining a Statistical Language Model with Logistic Regression to Predict the Lexical and Syntactic Difficulty of Texts for FFL</i> Thomas François</p> <p><i>Modelling Early Language Acquisition Skills: Towards a General Statistical Learning Mechanism</i> Guillaume Aimetti</p>	
10:30–11:00 Break (Muses Foyer & Skalkotas Foyer)				
<p>11:00–11:30</p> <p>T13. Information Retrieval <u>Chair:</u> Ani Nenkova (Skalkotas Hall)</p> <p><i>Using Non-lexical Features to Identify Effective Indexing Terms for Biomedical Illustrations</i> Matthew Simpson, Dina Demner-Fushman, Charles Sneiderman, Sameer K. Antani and George R. Thoma</p>	<p>T14. Question Answering <u>Chair:</u> Timothy Baldwin (MC2 Room)</p> <p><i>Flexible Answer Typing with Discriminative Preference Ranking</i> Christopher Pinchak, Dekang Lin and Davood Rafiei</p>	<p>T15. Text Categorization <u>Chair:</u> Lori Levin (MC3 Room)</p> <p><i>Syntactic and Semantic Kernels for Short Text Pair Categorization</i> Alessandro Moschitti</p>	<p>T16. Machine Learning <u>Chair:</u> Vincent Ng (Trianti Hall)</p> <p><i>Cube Summing, Approximate Inference with Non-Local Features, and Dynamic Programming without Semirings</i> Kevin Gimpel and Noah A. Smith</p>	
<p>11:30–12:30</p> <p>Q5. Semantics, Pragmatics and Learning <u>Chair:</u> Ido Dagan (Skalkotas Hall)</p> <p><i>Contextual Phrase-Level Polarity Analysis using Lexical Affect Scoring and Syntactic N-grams</i> Apoorv Agarwal, Fadi Biadisy and Kathleen McKeown</p> <p><i>Inference Rules and their Application to Recognizing Textual Entailment</i> Georgiana Dinu and Rui Wang</p> <p><i>A Robust and Extensible Exemplar-Based Model of Thematic Fit</i> Bram Vandekerckhove, Dominiek Sandra and Walter Daelemans</p> <p><i>An Alignment Algorithm Using Belief Propagation and a Structure-Based Distortion Model</i> Fabien Cromieres and Sadao Kurohashi</p> <p><i>Sequential Labeling with Latent Variables: An Exact Inference Algorithm and Its Efficient Approximation</i> Xu Sun and Jun'ichi Tsujii</p>	<p>Q6. Resources <u>Chair:</u> Roberto Navigli (MC2 Room)</p> <p><i>Empirical Evaluations of Animacy Annotation</i> Lilja Øvrelid</p> <p><i>Large-Coverage Root Lexicon Extraction for Hindi</i> Cohan Sujay Carlos, Monojit Choudhury and Sandipan Dandapat</p> <p><i>Growing Finely-Discriminating Taxonomies from Seeds of Varying Quality and Size</i> Tony Veale, Guofu Li and Yanfen Hao</p> <p><i>Semi-Supervised Polarity Lexicon Induction</i> Delip Rao and Deepak Ravichandran</p> <p><i>Outclassing Wikipedia in Open-Domain Information Extraction: Weakly-Supervised Acquisition of Attributes over Conceptual Hierarchies</i> Marius Pasca</p>	<p>Q7. Dialogue and Speech <u>Chair:</u> Amanda Stent (MC3 Room)</p> <p><i>User Simulations for Context-Sensitive Speech Recognition in Spoken Dialogue Systems</i> Oliver Lemon and Ioannis Konstas</p> <p><i>Learning to Interpret Utterances Using Dialogue History</i> David DeVault and Matthew Stone</p> <p><i>Incremental Dialogue Processing in a Micro-Domain</i> Gabriel Skantze and David Schlangen</p> <p><i>Frequency Matters: Pitch Accents and Information Status</i> Katrin Schweitzer, Michael Walsh, Bernd Möbius, Arndt Riestler, Antje Schweitzer and Hinrich Schuetze</p> <p><i>Web Augmentation of Language Models for Continuous Speech Recognition of SMS Text Messages</i> Mathias Creutz, Sami Virpioja and Anna Kovaleva</p>	<p>Q8. Translation <u>Chair:</u> Robert Moore (Trianti Hall)</p> <p><i>Bilingually Motivated Domain-Adapted Word Segmentation for Statistical Machine Translation</i> Yanjun Ma and Andy Way</p> <p><i>Lexical Morphology in Machine Translation: A Feasibility Study</i> Bruno Cartoni</p> <p><i>Improving Mid-Range Re-Ordering using Templates of Factors</i> Hieu Hoang and Philipp Koehn</p> <p><i>Feature-Based Method for Document Alignment in Comparable News Corpora</i> Thuy Vu, Ai Ti Aw and Min Zhang</p> <p><i>End-to-End Evaluation in Simultaneous Translation</i> Olivier Hamon, Christian Fügen, Djamel Mostefa, Victoria Arranz, Muntsin Kolss, Alex Waibel, and Khalid Choukri</p>	
12:30–14:00 Lunch (Muses Foyer)				

Thursday, 2 April 2009 – Main Conference

Afternoon Session

Poster Session (Muses Foyer)				
14:00–15:00	P5. Semantics, Pragmatics and Learning See Q5.	P6. Resources See Q6.	P7. Dialogue and Speech See Q7.	P8. Translation See Q8.
15:00–16:00	T17. Dialogue Chair: Oliver Lemon (Skalkotas Hall) <i>Who is “You”? Combining Linguistic and Gaze Features to Resolve Second-Person References in Dialogue</i> Matthew Frampton, Raquel Fernandez, Patrick Ehlen, Mario Christoudias, Trevor Darrell and Stanley Peters <i>A General, Abstract Model of Incremental Dialogue Processing</i> David Schlangen and Gabriel Skantze	T18. Tagging Chair: Stephen Clark (MC2 Room) <i>Tagging Urdu Text with Parts of Speech: A Tagger Comparison</i> Hassan Sajjad and Helmut Schmid <i>Semi-supervised Training for the Averaged Perceptron POS Tagger</i> Drahomíra “johanka” Spoustová, Jan Hajič, Jan Raab and Miroslav Spusta	T19. Speech Chair: Alessandro MoSchitti (MC3 Room) <i>Reconstructing False Start Errors in Spontaneous Speech Text</i> Erin Fitzgerald, Keith Hall and Frederick Jelinek <i>Effects of Word Confusion Networks on Voice Search</i> Junlan Feng and Srinivas Bangalore	T20. Translation Chair: Tony Veale (Trianti Hall) <i>Translation as Weighted Deduction</i> Adam Lopez <i>Improvements in Analogical Learning: Application to Translating Multi-Terms of the Medical Domain</i> Philippe Langlais, François Yvon and Pierre Zweigenbaum
16:00–16:30	Break (Muses Foyer & Skalkotas Foyer)			
16:30–18:00	T21. Lexical Semantics Chair: Ana Korhonen (Skalkotas Hall) <i>Bayesian Word Sense Induction</i> Samuel Brody and Mirella Lapata <i>Co-Dispersion: A Windowless Approach to Lexical Association</i> Justin Washtell <i>Using Lexical and Relational Similarity to Classify Semantic Relations</i> Diarmuid Ó Séaghdha and Ann Copestake	T22. Syntax/Parsing Chair: Josef van Genabith (MC2 Room) <i>Dependency Trees and the Strong Generative Capacity of CCG</i> Alexander Koller and Marco Kuhlmann <i>Parsing Coordinations</i> Sandra Kübler, Erhard Hinrichs, Wolfgang Maier and Eva Klett <i>Enhancing Unlexicalized Parsing Performance Using a Wide Coverage Lexicon, Fuzzy Tag-Set Mapping, and EM-HMM-Based Lexical Probabilities</i> Yoav Goldberg, Reut Tsarfaty, Meni Adler and Michael Elhadad	T23. Resources Chair: James Curran (MC3 Room) <i>Language-Independent Bilingual Terminology Extraction from a Multilingual Parallel Corpus</i> Els Lefever, Lieve Macken and Veronique Hoste <i>Generating a Non-English Subjectivity Lexicon: Relations That Matter</i> Valentin Jijkoun and Katja Hofmann <i>Language ID in the Context of Harvesting Language Data off the Web</i> Fei Xia, William Lewis and Hoifung Poon	T24. Translation Chair: Shuly Wintner (Trianti Hall) <i>Word Lattices for Multi-Source Translation</i> Josh Schroeder, Trevor Cohn and Philipp Koehn <i>Translation and Extension of Concepts Across Languages</i> Dmitry Davidov and Ari Rappoport <i>Syntactic Phrase Reordering for English-to-Arabic Statistical Machine Translation</i> Ibrahim Badr, Rabih Zbib and James Glass

Friday, 3 April 2009 – Main Conference

Morning Session

09:30–10:30	Plenary Session: Invited Talk (Skalkotas Hall)		
	Franciska de Jong, Universiteit Twente <i>NLP and the Humanities: The Revival of an Old Liaison</i>		
10:30–11:00	Break (Muses Foyer & Skalkotas Foyer)		
11:00–12:30	Demo Session (Muses Foyer)		
	<i>Frolog: an Accommodating Text-Adventure Game</i> Luciana Benotti	<i>A Comparison of Clausal Coordinate Ellipsis in Estonian and German: Remarkably Similar Elision Rules Allow a Language-Independent Ellipsis-Generation Module</i> Karin Harbusch, Mare Koit and Haldur Öim	<i>Matching Readers' Preferences and Reading Skills with Appropriate Web Texts</i> Eleni Miltsakaki
	<i>CBSEAS, a Summarization System – Integration of Opinion Mining Techniques to Summarize Blogs</i> Aurélien Bossard, Michel Génèreux and Thierry Poibeau	<i>Foma: a Finite-State Compiler and Library</i> Mans Hulden	<i>A Text-Based Search Interface for Multimedia Dialectics</i> Katerina Pastra and Eirini Balta
	<i>Grammatical Framework Web Service</i> Björn Bringert, Krasimir Angelov and Aarne Ranta	<i>The Software Architecture for the First Challenge on Generating Instructions in Virtual Environments</i> Alexander Koller, Donna Byron, Justine Cassell, Robert Dale, Johanna Moore, Jon Oberlander and Kristina Striegnitz	<i>Grammar Development in GF</i> Aarne Ranta, Krasimir Angelov and Björn Bringert
	<i>Gossip Galore – A Self-Learning Agent for Exchanging Pop Trivia</i> Xiwen Cheng, Peter Adolphs, Feiyu Xu, Hans Uszkoreit and Hong Li	<i>Adaptive Natural Language Interaction</i> Stasinios Konstantopoulos, Athanasios Tegos, Dimitrios Bilidas, Ion Androutsopoulos, Gerasimos Lampouras, Colin Matheson, Olivier Deroo and Prodromos Malakasiotis	<i>Three BioNLP Tools Powered by a Biological Lexicon</i> Yutaka Sasaki, Paul Thompson, John McNaught and Sophia Ananiadou
	<i>An Open-Source Natural Language Generator for OWL Ontologies and Its Use in Protégé and Second Life</i> Dimitrios Galanis, George Karakatsiotis and Ion Androutsopoulos	<i>Parsing, Projecting & Prototypes: Repurposing Linguistic Data on the Web</i> William Lewis and Fei Xia	<i>A Mobile Health and Fitness Companion Demonstrator</i> Olov Ståhl, Björn Gambäck, Markku Turunen and Jaakko Hakulinen
	<i>eHumanities Desktop – An Online System for Corpus Management and Analysis in Support of Computing in the Humanities</i> Rüdiger Gleim, Ulli Waltinger, Alexandra Ernst, Alexander Mehler, Tobias Feith and Dietmar Esch	<i>A Tool for Multi-Word Expression Extraction in Modern Greek Using Syntactic Parsing</i> Athina Michou and Violeta Seretan	
12:30–14:00	Lunch (Muses Foyer)		

Friday, 3 April 2009 – Main Conference

Afternoon Session

14:00–15:30	<p>T25. Discourse <u>Chair:</u> Kathleen McKeown</p> <p>(Skalkotas Hall)</p> <p><i>EM Works for Pronoun Anaphora Resolution</i> Eugene Charniak and Micha Elsner</p> <p><i>Person Identification from Text and Speech Genre Samples</i> Jade Goldstein-Stewart, Ransom Winder and Roberta Sabin</p> <p><i>Unsupervised Recognition of Literal and Non-Literal Use of Idiomatic Expressions</i> Caroline Sporleder and Linlin Li</p>	<p>T26. Syntax and Parsing <u>Chair:</u> Sandra Kuebler</p> <p>(MC2 Room)</p> <p><i>Fast Full Parsing by Linear-Chain Conditional Random Fields</i> Yoshimasa Tsuruoka, Jun'ichi Tsujii and Sophia Ananiadou</p> <p><i>Unsupervised Methods for Head Assignments</i> Federico Sangati and Willem Zuidema</p> <p><i>Incremental Parsing with Parallel Multiple Context-Free Grammars</i> Krasimir Angelov</p>	<p>T27. Semantics <u>Chair:</u> Ann Copestake</p> <p>(MC3 Room)</p> <p><i>A Logic of Semantic Representations for Shallow Parsing</i> Alexander Koller and Alex Lascarides</p> <p><i>Semi-Supervised Semantic Role Labeling</i> Hagen Fürstenau and Mirella Lapata</p> <p><i>Evaluating the Inferential Utility of Lexical-Semantic Resources</i> Shachar Mirkin, Ido Dagan and Eyal Shnarch</p>
15:30–16:00	Break (Muses Foyer & Skalkotas Foyer)		
16:00–17:00	Plenary Session: EACL Business Meeting (Skalkotas Hall)		
17:00–17:15	Plenary Session: Closing (Skalkotas Hall)		