

**EACL 2009**

**Proceedings of the  
EACL 2009 Workshop on  
Computational Linguistic  
Aspects of  
Grammatical Inference**

**CLAGI 2009**

30 March – 2009

Megaron Athens International Conference Centre  
Athens, Greece

Production and Manufacturing by  
*TEHNOGRAFIA DIGITAL PRESS*  
7 Ektoros Street  
152 35 Vrilissia  
Athens, Greece



©2009 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

## Preface

We are delighted to present you with this volume containing the papers accepted for presentation at the 1<sup>st</sup> workshop CLAGI 2009, held in Athens, Greece, on March 30<sup>th</sup> 2009.

We want to acknowledge the help of the PASCAL 2 network of excellence.

Thanks also to Damir Čavar for giving an invited talk and to the programme committee for the reviewing and advising.

We are indebted to the general chair of EACL 2009, Alex Lascarides, to the publication chairs, Kemal Oflazer and David Schlangen, and to the Workshop chairs Stephen Clark and Miriam Butt, for all their help. We are also grateful for Thierry Murgue's assistance when putting together the proceedings.

Wishing you a very enjoyable time at CLAGI 2009!

Menno van Zaanen and Colin de la Higuera  
CLAGI 2009 Programme Chairs



## CLAGI 2009 Program Committee

### Program Chairs:

Menno van Zaanen, Tilburg University (The Netherlands)

Colin de la Higuera, University of Saint-Étienne, (France)

### Program Committee Members:

Pieter Adriaans, University of Amsterdam (The Netherlands)

Srinivas Bangalore, AT&T Labs-Research (USA)

Leonor Becerra-Bonache, Yale University (USA)

Rens Bod, University of Amsterdam (The Netherlands)

Antal van den Bosch, Tilburg University (The Netherlands)

Alexander Clark, Royal Holloway, University of London (UK)

Walter Daelemans, University of Antwerp (Belgium)

Shimon Edelman, Cornell University (USA)

Jeroen Geertzen, University of Cambridge (UK)

Jeffrey Heinz, University of Delaware (USA)

Colin de la Higuera, University of Saint-Étienne (France)

Alfons Juan, Polytechnic University of Valencia (Spain)

Frantisek Mraz, Charles University (Czech Republic)

Georgios Petasis, National Centre for Scientific Research (NCSR) "Demokritos" (Greece)

Khalil Sima'an, University of Amsterdam (The Netherlands)

Richard Sproat, University of Illinois at Urbana-Champaign (USA)

Menno van Zaanen, Tilburg University (The Netherlands)

Willem Zuidema, University of Amsterdam (The Netherlands)



## Table of Contents

<i>Grammatical Inference and Computational Linguistics</i>	
Menno van Zaanen and Colin de la Higuera .....	1
<i>On Bootstrapping of Linguistic Features for Bootstrapping Grammars</i>	
Damir Čavar .....	5
<i>Dialogue Act Prediction Using Stochastic Context-Free Grammar Induction</i>	
Jeroen Geertzen .....	7
<i>Experiments Using OSTIA for a Language Production Task</i>	
Dana Angluin and Leonor Becerra-Bonache .....	16
<i>GREAT: A Finite-State Machine Translation Toolkit Implementing a Grammatical Inference Approach for Transducer Inference (GIATI)</i>	
Jorge González and Francisco Casacuberta .....	24
<i>A Note on Contextual Binary Feature Grammars</i>	
Alexander Clark, Remi Eyraud and Amaury Habrard .....	33
<i>Language Models for Contextual Error Detection and Correction</i>	
Herman Stehouwer and Menno van Zaanen .....	41
<i>On Statistical Parsing of French with Supervised and Semi-Supervised Strategies</i>	
Marie Candito, Benoit Crabbé and Djamé Seddah .....	49
<i>Upper Bounds for Unsupervised Parsing with Unambiguous Non-Terminally Separated Grammars</i>	
Franco M. Luque and Gabriel Infante-Lopez .....	58
<i>Comparing Learners for Boolean partitions: Implications for Morphological Paradigms</i>	
Katya Pertsova .....	66

