

**EACL 2009**

**Proceedings of the  
EACL 2009  
Workshop on the  
Interaction between  
Linguistics and  
Computational Linguistics:  
Virtuous, Vicious or Vacuous?**

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

This workshop is an attempt to bring together a range of linguists and computational linguists who operate across or near the computational “divide”, to reflect on the relationship between the two fields, including the following questions:

- What contributions has computational linguistics made to linguistics, and vice versa?
- What are examples of success/failure of marrying linguistics and computational linguistics, and what can we learn from them?
- How can we better facilitate the virtuous cycle between computational linguistics and linguistics?
- Is modern-day computational linguistics relevant to current-day linguistics, and vice versa? If not, should it be made more relevant, and how?
- What do linguistics and computational linguistics stand to gain from greater cross-awareness between the two fields?
- What untapped areas/aspects of linguistics are ripe for cross-fertilisation with computational linguistics, and vice versa?

On the basis of exploring answers to these and other questions, the workshop aims to explore possible trajectories for linguistics and computational linguistics, in terms of both concrete low-level tasks and high-level aspirations/synergies.

In its infancy, computational linguistics drew heavily on theoretical linguistics. There have been numerous examples of co-development successes between computational and theoretical linguistics over the years (e.g. syntactic theories, discourse processing and language resource development), and significant crossover with other areas of linguistics such as psycholinguistics and corpus linguistics.

Throughout the history of the field, however, there has always been a subset of computational linguistics which has openly distanced itself from theoretical linguistics, perhaps most famously in the field of machine translation (MT) where there is relatively little in the majority of “successful” MT systems that a linguist would identify with. In the current climate of hard-core empiricism within computational linguistics it is appropriate to reflect on where we have come from and where we are headed relative to the various other fields of linguistics. As part of this reflection, it is timely to look beyond theoretical linguistics to the various other fields of linguistics which have traditionally received less exposure in computational linguistics, including sociolinguistics, historical linguistics, neurolinguistics and evolutionary linguistics.

We would like to thank all of our invited speakers and panelists for agreeing to participate in the workshop and help shape the debate. We would also like to thank the workshop chairs and local organisers of EACL 2009 for all of their behind-the-scenes efforts, without which this workshop would not have been possible. The workshop is endorsed by the Erasmus Mundus European Masters Program in Language and Communication Technologies (LCT; <http://lct-master.org>).

Timothy Baldwin  
Valia Kordoni



# Organizers

## Workshop Chairs:

Timothy Baldwin, University of Melbourne (Australia)  
Valia Kordoni, DFKI and Saarland University (Germany)

## Invited Speakers:

Mark Johnson, Brown University (USA)  
Frank Keller, University of Edinburgh (UK)  
Mark Liberman, University of Pennsylvania (USA)  
Stelios Piperidis, Institute for Language and Speech Processing (Greece)  
Geoffrey Pullum, University of Edinburgh (UK)

## Panelists:

Emily Bender, University of Washington (USA)  
Gregor Erbach, European Union (Belgium)  
Bob Moore, Microsoft Research (USA)  
Gertjan van Noord, University of Groningen (Netherlands)  
Hans Uszkoreit, DFKI and Saarland University (Germany)



## Table of Contents

<i>Machine Translation and its Philosophical Accounts</i>	
Stelios Piperidis .....	1
<i>The Annotation Conundrum</i>	
Mark Liberman .....	2
<i>How the Statistical Revolution Changes (Computational) Linguistics</i>	
Mark Johnson .....	3
<i>Computational Linguistics and Generative Linguistics: The Triumph of Hope over Experience</i>	
Geoffrey Pullum .....	12
<i>Linguistics in Computational Linguistics: Observations and Predictions</i>	
Hans Uszkoreit .....	22
<i>Linguistically Naïve != Language Independent: Why NLP Needs Linguistic Typology</i>	
Emily M. Bender .....	26
<i>Parsed Corpora for Linguistics</i>	
Gertjan van Noord and Gosse Bouma .....	33
<i>Computational Linguistics and Linguistics: What Keeps Them together, What Sets Them apart?</i>	
Gregor Erbach .....	40
<i>What Do Computational Linguists Need to Know about Linguistics?</i>	
Robert C. Moore .....	41
<i>The Interaction of Syntactic Theory and Computational Psycholinguistics</i>	
Frank Keller .....	43





## Workshop Program

### Monday, March 30, 2009

- 8:55–9:00      Opening Remarks
- 09:00-09:45    *Machine Translation and its Philosophical Accounts*  
Stelios Piperidis
- 09:45-10:30    *The Annotation Conundrum*  
Mark Liberman
- 10:30-11:00    Coffee Break
- 11:00-11:45    *How the Statistical Revolution Changes (Computational) Linguistics*  
Mark Johnson
- 11:45-12:30    Discussion
- 12:30-14:00    Lunch Break
- 14:00-14:45    *Computational Linguistics and Generative Linguistics: The Triumph of Hope over Experience*  
Geoffrey Pullum
- 14:45-16:00    Panel and Discussion
- Linguistics in Computational Linguistics: Observations and Predictions*  
Hans Uszkoreit
- Linguistically Naïve != Language Independent: Why NLP Needs Linguistic Typology*  
Emily M. Bender
- Parsed Corpora for Linguistics*  
Gertjan van Noord and Gosse Bouma
- Computational Linguistics and Linguistics: What Keeps Them together, What Sets Them apart?*  
Gregor Erbach
- What Do Computational Linguists Need to Know about Linguistics?*  
Robert C. Moore
- 16:00-16:30    Coffee Break

**Monday, March 30, 2009 (continued)**

16:30-17:15 *The Interaction of Syntactic Theory and Computational Psycholinguistics*  
Frank Keller

17:15-17:55 Discussion

17:55-18:00 Closing Remarks