

HLT-NAACL 2006

**SIGPHON 2006:
Eighth Meeting of the
ACL Special Interest Group
on Computational Phonology**

Proceedings of the Workshop

8 June 2006
New York City, USA

Production and Manufacturing by
Omnipress Inc.
2600 Anderson Street
Madison, WI 53704

©2006 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

Introduction

We are pleased to present the Proceedings of the Eighth Meeting of the ACL Special Interest Group on Computational Phonology (SIGPHON) to be held on June 8 in New York City. This is the first time that the SIGPHON workshop has been collocated with the HLT-NAACL conference. Previous meetings were held in conjunction with ACL and COLING in Las Cruces (1994), Santa Cruz (1996), Madrid (1997), Quebec (1998), Luxembourg (2000), Philadelphia (2002), and Barcelona (2004).

One of the missions of SIGPHON is to encourage interaction between work in computational linguistics and work in theoretical phonology, in the hope that both fields will profit from the interaction. In addition, SIGPHON continues to promote work in computational morphology, seeking to fill in for the absence of an analogous SIGMORPH group. Our recent meetings have been successful in both regards, and we anticipate this will continue in 2006. Many mainstream phonologists are employing computational tools and models that are of considerable interest to computational linguists more generally, and our intention is that this workshop should be a forum to bring this work to the attention of a wider range of computational linguists.

The submissions were reviewed by a program committee composed of eighteen experts in the field. We are grateful to them for their timely, thoughtful, and thorough reviews.

We hope you enjoy this year's meeting!

Greg Kondrak
Richard Wicentowski
June 2006

Organizers:

Richard Wicentowski, Swarthmore College
Greg Kondrak, University of Alberta

SIGPHON Executive Committee:

Jason Eisner, The Johns Hopkins University, President
Richard Wicentowski, Swarthmore College, Secretary
Adam Albright, Massachusetts Institute of Technology
Katrín Kirchoff, University of Washington
Eric Fosler-Lussier, The Ohio State University

Program Committee Members:

Adam Albright, University of California, Santa Cruz
Paul Boersma, University of Amsterdam
Anja Belz, University of Brighton
Steven Bird, University of Melbourne
Julie Carson-Berndsen, University College Dublin
John Coleman, University of Oxford
Mathias Creutz, Helsinki University of Technology
Jason Eisner, The Johns Hopkins University
John Goldsmith, University of Chicago
Sharon Goldwater, Brown University
Lauri Karttunen, Palo Alto Research Center
Greg Kondrak, University of Alberta
Mike Maxwell, Linguistic Data Consortium
Kemal Oflazer, Sabanci University
Gerald Penn, University of Toronto
Vito Pirrelli, Istituto di Linguistica Computazionale
Jason Riggle, University of Chicago
Richard Sproat, University of Illinois at Urbana-Champaign

Workshop Website:

<http://nlp.cs.swarthmore.edu/sigphon06/>

Table of Contents

<i>A Combined Phonetic-Phonological Approach to Estimating Cross-Language Phoneme Similarity in an ASR Environment</i>	
Lynette Melnar and Chen Liu	1
<i>Improving Syllabification Models with Phonotactic Knowledge</i>	
Karin Müller	11
<i>Learning Quantity Insensitive Stress Systems via Local Inference</i>	
Jeffrey Heinz	21
<i>Invited Talk: Universal Constraint Rankings Result from Learning and Evolution</i>	
Paul Boersma	31
<i>Exploring variant definitions of pointer length in MDL</i>	
Aris Xanthos, Yu Hu and John Goldsmith	32
<i>Improved morpho-phonological sequence processing with constraint satisfaction inference</i>	
Antal van den Bosch and Sander Canisius	41
<i>Richness of the Base and Probabilistic Unsupervised Learning in Optimality Theory</i>	
Gaja Jarosz	50
<i>Morphology Induction from Limited Noisy Data Using Approximate String Matching</i>	
Burcu Karagol-Ayan, David Doermann and Amy Weinberg	60
<i>Learning Probabilistic Paradigms for Morphology in a Latent Class Model</i>	
Erwin Chan	69
<i>A Naive Theory of Affixation and an Algorithm for Extraction</i>	
Harald Hammarström	79

Conference Program

Thursday, June 8, 2006

- 9:00–9:30 *A Combined Phonetic-Phonological Approach to Estimating Cross-Language Phoneme Similarity in an ASR Environment*
Lynette Melnar and Chen Liu
- 9:30–10:00 *Improving Syllabification Models with Phonotactic Knowledge*
Karin Müller
- 10:00–10:30 *Learning Quantity Insensitive Stress Systems via Local Inference*
Jeffrey Heinz
- 10:30–11:00 Break
- 11:00–12:30 *Invited Talk: Universal Constraint Rankings Result from Learning and Evolution*
Paul Boersma
- 12:30–14:00 Lunch
- 14:00–14:30 *Exploring variant definitions of pointer length in MDL*
Aris Xanthos, Yu Hu and John Goldsmith
- 14:30–15:00 *Improved morpho-phonological sequence processing with constraint satisfaction inference*
Antal van den Bosch and Sander Canisius
- 15:00–15:30 *Richness of the Base and Probabilistic Unsupervised Learning in Optimality Theory*
Gaja Jarosz
- 15:30–16:00 Break
- 16:00–16:30 *Morphology Induction from Limited Noisy Data Using Approximate String Matching*
Burcu Karagol-Ayan, David Doermann and Amy Weinberg
- 16:30–17:00 *Learning Probabilistic Paradigms for Morphology in a Latent Class Model*
Erwin Chan
- 17:00–17:30 *A Naive Theory of Affixation and an Algorithm for Extraction*
Harald Hammarström

