

ACL-08: HLT

SIGMORPHON 2008

Tenth Meeting of the ACL Special Interest Group on Computational Morphology and Phonology

Proceedings of the Workshop

June 19, 2008

The Ohio State University
Columbus, Ohio, USA

Production and Manufacturing by
Omnipress Inc.
2600 Anderson Street
Madison, WI 53707
USA

©2008 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

ISBN 978-1-932432-12-1

Introduction

We are pleased to present the Proceedings of the Tenth Meeting of the ACL Special Interest Group on Computational Morphology and Phonology (SIGMORPHON), to be held on June 19, 2008 at Ohio State University in Columbus, Ohio.

The purpose of SIGMORPHON is to foster computational research on the phonological, morphological, and phonetic properties of human language. All three of these sub-areas deal largely with the local structure of words and so share many technical methods. Furthermore, computational work that models empirical data must often draw on at least two of these areas, with explicit consideration of the morphology-phonology or phonology-phonetics interface.

Morphology and phonetics were officially added to the SIG's charter only in 2006, when the SIG membership voted to amend the SIG's constitution and change its name from SIGPHON. This expansion of the SIG's mission beyond phonology was supported and subsequently approved by the ACL. The new name (suggested by Johanna Moore) is pronounced "SIG *more* fun," which we hope is an accurate assessment.

Since SIGMORPHON's 2007 workshop was a special-topic workshop on computing and historical phonology, the present 2008 workshop is the first that reflects the breadth of the new charter. In particular, we are pleased to include two papers on unsupervised morphological analysis, and an invited talk on articulatory modeling for speech recognition.

We are grateful to the program committee for their careful and thoughtful reviews and discussions of the papers submitted this year. Just over half of the submissions were accepted on first review, with an additional submission accepted after revisions. We also thank this year's invited speakers, Karen Livescu and Jason Riggle, for presenting their noteworthy work to the SIGMORPHON community.

We hope that you enjoy the workshop and these proceedings.

Jason Eisner
Jeffrey Heinz

Organizers:

Jason Eisner, Johns Hopkins University
Jeffrey Heinz, University of Delaware

Program Committee Members:

Adam Albright, Massachusetts Institute of Technology
Lynne Cahill, University of Brighton
Mathias Creutz, Helsinki University of Technology
Mark Ellison, University of Western Australia
Eric Fosler-Lussier, Ohio State University
John Goldsmith, University of Chicago
Sharon Goldwater, Stanford University
Katrín Kirchoff, University of Washington
Greg Kondrak, University of Alberta
Ying Lin, University of Arizona
Mike Maxwell, University of Maryland
John Nerbonne, University of Groningen
Kemal Oflazer, Sabanci University
Jason Riggle, University of Chicago
Richard Sproat, University of Illinois
Richard Wicentowski, Swarthmore University
Shuly Wintner, University of Haifa

Invited Speakers:

Karen Livescu, Toyota Technological Institute at Chicago
Jason Riggle, University of Chicago

Table of Contents

<i>Invited talk: Phonological Models in Automatic Speech Recognition</i> Karen Livescu	1
<i>Bayesian Learning over Conflicting Data: Predictions for Language Change</i> Rebecca Morley	2
<i>A Bayesian Model of Natural Language Phonology: Generating Alternations from Underlying Forms</i> David Ellis	12
<i>Unsupervised Word Segmentation for Sesotho Using Adaptor Grammars</i> Mark Johnson	20
<i>Invited talk: Counting Rankings</i> Jason Riggle	28
<i>Three Correlates of the Typological Frequency of Quantity-Insensitive Stress Systems</i> Max Bane and Jason Riggle	29
<i>Phonotactic Probability and the Maori Passive: A Computational Approach</i> ‘Ōiwi Parker Jones	39
<i>Evaluating an Agglutinative Segmentation Model for ParaMor</i> Christian Monson, Alon Lavie, Jaime Carbonell and Lori Levin	49

Workshop Program

Thursday, June 19, 2008

- 8:50–9:00 Opening remarks
- 9:00–10:00 *Invited talk: Phonological Models in Automatic Speech Recognition*
Karen Livescu
- 10:00–10:30 *Bayesian Learning over Conflicting Data: Predictions for Language Change*
Rebecca Morley
- 10:30–11:00 Break
- 11:00–11:30 *A Bayesian Model of Natural Language Phonology: Generating Alternations from Underlying Forms*
David Ellis
- 11:30–12:00 *Unsupervised Word Segmentation for Sesotho Using Adaptor Grammars*
Mark Johnson
- 12:00–14:00 Lunch
- 14:00–15:00 *Invited talk: Counting Rankings*
Jason Riggle
- 15:00–15:30 *Three Correlates of the Typological Frequency of Quantity-Insensitive Stress Systems*
Max Bane and Jason Riggle
- 15:30–16:00 Break
- 16:00–16:30 *Phonotactic Probability and the Maori Passive: A Computational Approach*
‘Ōiwi Parker Jones
- 16:30–17:00 *Evaluating an Agglutinative Segmentation Model for ParaMor*
Christian Monson, Alon Lavie, Jaime Carbonell and Lori Levin
- 17:00–17:30 General discussion

