Current Themes in Computational Phonology and Morphology:

Seventh Meeting of the ACL Special Interest Group in Computational Phonology

Held in cooperation with ACL-2004

26 July 2004

Barcelona, Spain

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
73 Landmark Center
East Stroudsburg, PA 18301
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

SIGPHON EXECUTIVE COMMITTEE:

Jason Eisner, The Johns Hopkins University, President
Richard Wicentowski, Swarthmore College, Secretary
Paul Boersma, University of Amsterdam
Julie Carson-Berndsen, University College Dublin
John Coleman, University of Oxford
John Goldsmith, University of Chicago

ADDITIONAL PROGRAM COMMITTEE MEMBERS:

Adam Albright, University of California, Santa Cruz
Steven Bird, University of Melbourne
Bruce Hayes, University of California at Los Angeles
Lauri Karttunen, Palo Alto Research Center
Mike Maxwell, Linguistic Data Consortium
Kemal Oflazer, Sabanci University
Janet Pierrehumbert, Northwestern University
Richard Sproat, University of Illinois at Urbana-Champaign

WORKSHOP ORGANIZERS:

John Goldsmith, University of Chicago
Richard Wicentowski, Swarthmore College

WORKSHOP WEBSITE:

`http://www.cs.swarthmore.edu/~richardw/sigphon04/`

PREFACE

The ACL Special Interest Group in Computational Phonology organized a workshop on July 26, 2004, after the 24th meeting of the Association for Computational Linguistics in Barcelona, Spain. The papers in this volume were accepted for presentation at that workshop.

In the call for papers, we asked for papers addressing questions concerned either with computational morphology or phonology, and we listed the following topics as possible areas for discussion:

- Speech technologies relating to phonology
- Unsupervised learning of phonotactics, morphology, and phonology
- Stochastic or probabilistic phonology
- Computational evolutionary modeling of phonological systems
- Language-independent learning algorithms relating to computational phonology
- Finite-state formalizations of phonology
- Computational morphology and morphological modeling
- Speech generation
- Assistive learning tools for second language speakers

The papers were reviewed by a gracious committee, who agreed to read and evaluate papers in a very short period of time. This committee consisted of the executive committee of SIGPHON as well as a volunteer panel of outside experts.

We thank this committee for their work, and we want to give an extra special expression of our gratitude to Jason Eisner's efforts. Jason helped us at every step of the way, always going the extra mile to make sure things were moving the right way and at the right time.

John Goldsmith
Richard Wicentowski
June 2004

Table of Contents

Category competition drives contrast maintenance within an exemplar-based production/perception loop
Andrew Wedel 1

Non-locality all the way through: Emergent Global Constraints in the Italian Morphological Lexicon
Vito Pirrelli, Basilio Calderone, Ivan Herreros and Michele Virgilio 11

A Diachronic Approach for Schwa Deletion in Indo Aryan Languages
Monojit Choudhury, Anupam Basu and Sudeshna Sarkar 20

Automatic Acquisition of Feature-Based Phonotactic Resources
Julie Carson-Berndsen, Robert Kelly and Moritz Neugebauer 27

Priors in Bayesian Learning of Phonological Rules
Sharon Goldwater and Mark Johnson 35

Induction of a Simple Morphology for Highly-Inflecting Languages
Mathias Creutz and Krista Lugas 43

Unsupervised Induction of Natural Language Morphology Inflection Classes
Christian Monson, Alon Lavie, Jaime Carbonell and Lori Levin 52

A Comparison of Two Different Approaches to Morphological Analysis of Dutch
Guy De Pauw, Tom Laureys, Walter Daelemans and Hugo Van hamme 62

Multilingual Noise-Robust Supervised Morphological Analysis using the WordFrame Model
Richard Wicentowski 70

Segment Predictability as a Cue in Word Segmentation: Application to Modern Greek
C. Anton Rytting 78

Vi-xfst: A Visual Regular Expression Development Environment for Xerox Finite State Tool
Kemal Oflazer and Yasin Yılmaz 86

Technical Program Schedule

Monday, July 26

8:45-9:00 Welcome

9:00-9:25 *Category competition drives contrast maintenance within an exemplar-based production/perception loop*
Andrew Wedel

9:25-9:50 *Non-locality all the way through: Emergent Global Constraints in the Italian Morphological Lexicon*
Vito Pirrelli, Basilio Calderone, Ivan Herreros and Michele Virgilio

9:50-10:30 Break

10:30-10:55 *A Diachronic Approach for Schwa Deletion in Indo Aryan Languages*
Monojit Choudhury, Anupam Basu and Sudeshna Sarkar

10:55-11:20 *Automatic Acquisition of Feature-Based Phonotactic Resources*
Julie Carson-Berndsen, Robert Kelly and Moritz Neugebauer

11:20-11:45 *Priors in Bayesian Learning of Phonological Rules*
Sharon Goldwater and Mark Johnson

11:45-13:30 Lunch

13:30-13:55 *Induction of a Simple Morphology for Highly-Inflecting Languages*
Mathias Creutz and Krista Lugas

13:55-14:20 *Unsupervised Induction of Natural Language Morphology Inflection Classes*
Christian Monson, Alon Lavie, Jaime Carbonell and Lori Levin

14:20-14:45 *A Comparison of Two Different Approaches to Morphological Analysis of Dutch*
Guy De Pauw, Tom Laureys, Walter Daelemans and Hugo Van hamme

14:45-15:10 *Multilingual Noise-Robust Supervised Morphological Analysis using the WordFrame Model*
Richard Wicentowski

15:10-15:50 Break

15:50-16:15 *Segment Predictability as a Cue in Word Segmentation: Application to Modern Greek*
C. Anton Rytting

16:15-16:40 *Vi-xfst: A Visual Regular Expression Development Environment for Xerox Finite State Tool*
Kemal Oflazer and Yasin Yılmaz

16:40-17:00 General Discussions

THIS IS A BLANK PAGE PLEASE IGNORE

Author Index

Basu, Anupam 20
Calderone, Basilio 11
Carbonell, Jaime 52
Carson-Berndsen, Julie 27
Choudhury, Monojit 20
Creutz, Mathias 43
Daelemans, Walter 62
De Pauw, Guy 62
Goldwater, Sharon 35
Herreros, Ivan 11
Johnson, Mark 35
Kelly, Robert 27
Laureys, Tom 62
Lavie, Alon 52
Levin, Lori 52
Lugas, Krista 43
Monson, Christian 52
Neugebauer, Moritz 27
Oflazer, Kemal 86
Pirrelli, Vito 11
Rytting, C. Anton 78
Sarkar, Sudeshna 20
Van hamme, Hugo 62
Virgilio, Michele 11
Wedel, Andrew 1
Wicentowski, Richard 70
Yılmaz, Yasin 86