

TLSX
Computational Linguistics for Less-Studied Languages

Friday, November 3, 2006 – UTC 3.124 (before lunch), UTC 3.104 (after lunch)

- 8:30 – 9:15 Registration / Coffee and refreshments
9:15 – 9:30 Opening remarks
9:30 – 10:45 KEYNOTE ADDRESS
Linguistic Data Management with the Natural Language Toolkit
Steven Bird, University of Melbourne, University of Pennsylvania
- 10:45 – 10:50 mini-break
10:50 – 11:20 **Annotating and archiving natural language paradigms online,**
Dorothee Beermann & Atle Prange
- 11:20 – 11:50 **The SIL FieldWorks Language Explorer Approach to Morphological Parsing,**
H. Andrew Black & Gary F. Simons

---LUNCH (on your own, 70 minutes)---

- 1:00 – 1:30 **A Method for Enhancing Search Using Transliteration of Mandarin Chinese**
Vijay John
- 1:30 – 2:00 **Inflectional Vocalization of Arabic Text: A MaxEnt Tagging Approach,**
Frederick M. Hoyt
- 2:05 – 2:10 mini-break
2:10 – 3:35 KEYNOTE ADDRESS
Cutting Corpus Costs: Machine Learning and Annotation
Jason Baldridge, University of Texas at Austin
- 3:35 – 3:50 break
3:50 – 5:30 **Panel Discussion**
Understanding the needs of documentary and descriptive linguistics,
gaps in the current tool set, problems of current technologies, issues involved
in doing work on less-studied languages
Tony Woodbury, moderator

Saturday, November 4, 2006 – CAL 100

- 9:00 – 10:15 KEYNOTE ADDRESS
TBA
Emily Bender, University of Washington
- 10:15 – 10:20 mini-break
10:20 – 10:50 **Two Approaches to Mayan Grammar Development in CCG**
Elias Ponvert
- 10:50 – 11:20 **A Combinatory Categorical Grammar of a Fragment of American Sign Language**
Tony Wright
- 11:20 – 11:50 **A Morphological Analyzer for Verbal Aspect in American Sign Language**
Aaron Shield & Jason Baldridge

---LUNCH (on your own, 70 minutes)---

- 1:00 – 2:15 KEYNOTE ADDRESS
**Detecting outliers: useful for word sense assignment –
and for aiding manual annotation?**
Katrin Erk, University of Texas at Austin
- 2:15 – 2:20 mini-break
- 2:20 – 2:50 **Affix Discovery based on Entropy and Economy Measurements**
Alfonso Medina-Urrea
- 2:50 – 3:20 **Enriching Language Data through Projected Structures**
William Lewis, Fei Xia, & Dan Jinguji
- 3:20 – 3:50 **Finite State Methods for Bantu Verb Morphology**
Robert Elwell
- 3:50 – 4:05 break
- 4:05 – 5:20 KEYNOTE ADDRESS
Maximizing the Utility of Small Training Sets in Machine Learning
Raymond Mooney, University of Texas at Austin

CONFERENCE PARTY

Sunday, November 5, 2006 – CAL 100

- 9:00 – 10:15 KEYNOTE ADDRESS
TBA
Mark Liberman, University of Pennsylvania
- 10:15 – 10:30 break
- 10:30 – 12:30 **Panel Discussion**
How can computational linguistics address the needs of documentary
and descriptive linguistics, and how will doing so further the state of research in
the field of computational linguistics? What are fruitful directions for future
research? Where do we go from here?

Alternate: **Using POS Tags in Word Prediction: A Statistical Language
Modeling for the Persian Language**
Masood Ghayoomi & Ehsan Daroodi
