

2014 Association for Computational Linguistics Workshop on the
Use of Computational Methods in the Study of Endangered Languages (ComputEL '14)

WORKSHOP PROGRAM (UPDATED)

Heron Room (4th floor)
Baltimore Marriott Waterfront

26 June 2014

- 9:00–9:10 Introduction
- Paper Session 1: Computational Tools for Endangered Languages Research** (Chair: Worthy Martin)
- 9:10–9:30 *Aikuma: A Mobile App for Collaborative Language Documentation*
Steven Bird, Florian R. Hanke, Oliver Adams and Haejoong Lee
- 9:30–9:50 *Small Languages, Big Data: Multilingual Computational Tools and Techniques for the Lexicography of Endangered Languages*
Martin Benjamin and Paula Radetzky
- 9:50–10:10 *LingSync & the Online Linguistic Database: New Models for the Collection and Management of Data for Language Communities, Linguists and Language Learners*
Joel Dunham, Gina Cook and Joshua Horner
- 10:10–10:40 *Modeling the Noun Morphology of Plains Cree*
Conor Snoek, Dorothy Thunder, Kaidi Lõo, Antti Arppe, Jordan Lachler, Sjur Moshagen and Trond Trosterud
- 10:40–11:00 Coffee Break
- Paper Session 2: Applying Computational Methods to Endangered Languages** (Chair: Terry Langendoen)
- 11:00–11:30 *Learning Grammar Specifications from IGT: A Case Study of Chintang*
Emily M. Bender, Joshua Crowgey, Michael Wayne Goodman and Fei Xia
- 11:30–12:00 *Creating Lexical Resources for Endangered Languages*
Khang Nhut Lam, Feras Al Tarouti and Jugal Kalita
- 12:00–12:20 *Documenting Endangered Languages with the WordsEye Linguistics Tool*
Morgan Ulinski, Anusha Balakrishnan, Daniel Bauer, Bob Coyne, Julia Hirschberg and Owen Rambow
- 12:20–2:00 Lunch
- 2:00–3:00 **Poster and Tool Demonstration Session**
- Estimating Native Vocabulary Size in an Endangered Language*
Timofey Arkhangelskiy
- InterlinguaPlus Machine Translation Approach for Local Languages: Ekegusii & Swahili*
Edward Ombui, Peter Wagacha and Wanjiku Ng'ang'a
- Building and Evaluating Somali Language Corpora*
Abdillahi Nimaan
- Additional demonstrations of tools presented in paper session 1
- Paper Session 3: Infrastructure and Community Development for Computational Research on Endangered Languages** (Chair: Steven Bird)
- 3:00–3:30 *SeedLing: Building and Using a Seed corpus for the Human Language Project*
Guy Emerson, Liling Tan, Susanne Fertmann, Alexis Palmer and Michaela Regneri
- 3:30–4:00 Coffee Break
- 4:00–4:20 *Short-Term Projects, Long-Term Benefits: Four Student NLP Projects for Low-Resource Languages*
Alexis Palmer and Michaela Regneri
- 4:20–4:50 *Data Warehouse, Bronze, Gold, STEC, Software*
Doug Cooper
- 4:50–5:20 *Time to Change the “D” in “DEL”*
Stephen Beale
- 5:20–5:30 Concluding Remarks