

Workshop on the Use of Computational Methods in the Study of Endangered Languages

(ComputEL 2014)

Held at ACL 2014

**Baltimore, Maryland, USA
26 June 2014**

ISBN: 978-1-63439-202-0

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2014) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Aikuma: A Mobile App for Collaborative Language Documentation</i> Steven Bird, Florian R. Hanke, Oliver Adams and Haejoong Lee	1
<i>Documenting Endangered Languages with the WordsEye Linguistics Tool</i> M. Ulinski, A. Balakrishnan, D. Bauer, B. Coyne, J. Hirschberg and O. Rambow	6
<i>Small Languages, Big Data: Multilingual Computational Tools and Techniques for the Lexicography of Endangered Languages</i> Martin Benjamin and Paula Radetzky	15
<i>LingSync & the Online Linguistic Database: New Models for the Collection and Management of Data for Language Communities, Linguists and Language Learners</i> Joel Dunham, Gina Cook and Joshua Horner	24
<i>Modeling the Noun Morphology of Plains Cree</i> C. Snoek, D. Thunder, K. Lõo, A. Arppe, J. Lachler, S. Moshagen and T. Trosterud	34
<i>Learning Grammar Specifications from IGT: A Case Study of Chintang</i> Emily M. Bender, Joshua Crowgey, Michael Wayne Goodman and Fei Xia	43
<i>Creating Lexical Resources for Endangered Languages</i> Khang Nhut Lam, Feras Al Tarouti and Jugal Kalita	54
<i>Estimating Native Vocabulary Size in an Endangered Language</i> Timofey Arkhangelskiy	63
<i>InterlinguaPlus Machine Translation Approach for Local Languages: Ekegusii & Swahili</i> Edward Ombui, Peter Wagacha and Wanjiku Ng'ang'a	68
<i>Building and Evaluating Somali Language Corpora</i> Abdillahi Nimaan	73
<i>SeedLing: Building and Using a Seed corpus for the Human Language Project</i> Guy Emerson, Liling Tan, Susanne Fertmann, Alexis Palmer and Michaela Regneri	77
<i>Short-Term Projects, Long-Term Benefits: Four Student NLP Projects for Low-Resource Languages</i> Alexis Palmer and Michaela Regneri	86
<i>Data Warehouse, Bronze, Gold, STEC, Software</i> Doug Cooper	91
<i>Time to Change the "D" in "DEL"</i> Stephen Beale	100