

"First Steps for Language Documentation of Minority Languages:
Computational Linguistic Tools for Morphology, Lexicon and Corpus Compilation"
LREC2004 (Lisbon, Portugal)
Morning of 24th of May 2004 (8:00 to 13:30)

Oral Session: First Steps for Language Documentation

08:00 Workshop Welcome and Introduction. Bojan Petek (Univ. of Ljubljana)

08:15 "Morphological Analysis and Generation: A First Step in Natural Language Processing" Kenneth R. Beesley (Xerox)

08:45 "Laying Lexical Foundations for NLP: the Case of Basque at the Ixa Research Groups" Xabier Artola (Univ. of The Basque Country))

09:15 "First steps in corpus building for linguistics and technology." Dafydd Gibbon (Univ. Bielefeld)

09:45 Break

Oral Session: HLT and the Coverage of Languages

10:15 "First steps in FP6" Bojan Petek (Univ. of Ljubljana)

10:45 Panel discussion - Invited speakers will be included as panelist members.

12:00 Poster Session

1. Atelach Alemu & Lars Asker: Building an Amharic Lexicon from Parallel Texts
2. Jens Allwood & A. P. Hendrikse: Spoken Language Corpora in South Africa
3. Akshar Bharati, Rajeev Sangal, Dipti M Sharma, Radhika Mamidi: Generic Morphological Analysis Shell
4. Emily Bender Dan Flickinger, Jeff Good, Ivan Sag: Montage: Leveraging advances in grammar engineering, linguistic ontologies, and mark-up for the documentation of underdescribed languages.
5. Bordel G., Ezeiza A., Lopez de Ipina K., Méndez M., Peñagarikano M., Rico T., Tovar C., Zulueta E.: Linguistic Resources and Tools for Automatic Speech Recognition in Basque.
6. Montserrat Civit, Núria Bufí and Ma. Pilar Valverde: Building Cat3LB: a Treebank for Catalan.
7. Mike Maxwell: From Legacy Lexicon to Archivable Resource.
8. László Németh, Viktor Trón, Péter Halácsy, András Kornai, András Rung, István Szakadát: Leveraging the open source ispell codebase for minority language analysis.
9. Ailbhe Ní Chasaide, Martha Dalton, Mika Ito, Christer Gobl: Analysing Irish prosody: A dual linguistic/quantitative approach.
10. Attila Novák: Creating a Morphological Analyzer and Generator for the Komi language.
11. Delyth Prys, Briony Williams, Bill Hicks, Dewi Jones, Ailbhe Ní Chasaide, Christer Gobl, Julie Carson-Berndsen, Fred Cummins, Máire Ní Chiosáin, John McKenna, Rónán Scaife, Elaine Uí Dhonnchadha: WISPR: Speech Processing Resources for Welsh and Irish.
12. Rajmund Piotrowski, Yuri Romanov: Imperial and Minority Languages in the Former USSR and in the Post-Soviet Area
13. Ksenia Shalnova and Roger Tucker: Issues in Porting TTS to Minority Languages.
14. Kiril Simov, Petya Osenova, Alexander Simov, Krasimira Ivanova, Ilko Grigorov, Hristo Ganev: Creation of a Tagged Corpus for Less-Processed Languages with ClaRK System.
15. Oliver Streiter, Mathias Stuflesser, Isabella Ties: CLE, an aligned Tri-lingual Ladin-Italian-German Corpus. Corpus Design and Interface.
16. Nicholas Thieberger: Building an interactive corpus of field recordings.
17. Trond Trosterud: Porting morphological analysis and disambiguation to new languages.

13:30 End