

# **9th Workshop on Asian Language Resources 2011**

**Held at the 5th International Joint Conference on Natural  
Language Processing 2011 (IJCNLP 2011)**

**Chiang Mai, Thailand  
12-13 November 2011**

**ISBN: 978-1-62276-455-6**

**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© (2011) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2012)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Participation in Language Resource Development and Sharing (Invited Talk)</i> Virach Sornlertlamvanich .....	1
<i>A Grammar Checker for Tagalog using LanguageTool</i> Nathaniel Oco and Allan Borra .....	2
<i>Bantay-Wika: towards a better understanding of the dynamics of Filipino culture and linguistic change (Short Paper)</i> Joel Ilao, Rowena Cristina Guevara, Virgilio Llenaresas, Eilene Antoinette Narvaez and Jovy Peregrino .....	10
<i>Engineering a Deep HPSG for Mandarin Chinese (Short Paper)</i> Yi Zhang, Rui Wang and Yu Chen .....	18
<i>Error Detection for Treebank Validation</i> Bharat Ram Ambati, Rahul Agarwal, Mridul Gupta, Samar Husain and Dipti Misra Sharma .....	23
<i>Experiences in Building Urdu WordNet</i> Farah Adeeba and Sarmad Hussain .....	31
<i>Feasibility of Leveraging Crowd Sourcing for the Creation of a Large Scale Annotated Resource for Hindi English Code Switched Data: A Pilot Annotation</i> Mona Diab and Ankit Kamboj .....	36
<i>Linguist's Assistant: A Resource For Linguists</i> Stephen Beale and Tod Allman .....	41
<i>Multi-stage Annotation using Pattern-based and Statistical-based Techniques for Automatic Thai Annotated Corpus Construction</i> Nattapong Tongtep and Thanaruk Theeramunkong .....	50
<i>Philippine Languages Online Corpora: Status, issues, and prospects</i> Shirley Dita and Rachel Edita Roxas .....	59
<i>Providing Ad Links to Travel Blog Entries Based on Link Types</i> Aya Ishino, Hidetsugu Nanba and Toshiyuki Takezawa .....	63
<i>Towards a Computational Semantic Analyzer for Urdu</i> Annette Hautli and Miriam Butt .....	71
<i>Word Disambiguation in Shahmukhi to Gurmukhi Transliteration</i> Tejinder Singh Saini and Gurpreet Singh Lehal .....	79