

# **8th Workshop on Asian Language Resources**

**At the 23rd International Conference on Computational Linguistics 2010**

**Beijing, China  
21 – 22 August 2010**

**ISBN: 978-1-61782-324-4**

**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© (2010) by the Chinese Information Processing Society of China  
All rights reserved.

Printed by Curran Associates, Inc. (2011)

For permission requests, please contact the Chinese Information Processing Society of China  
at the address below.

Chinese Information Processing Society of China  
No. 4 Zhongguancun South 4th St.  
Haidian District Beijing  
China 100190

**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>A Thesaurus of Predicate-Argument Structure for Japanese Verbs to Deal with Granularity of Verb Meanings</i>             |     |
| Koichi Takeuchi, Kentaro Inui, Nao Takeuchi and Atsushi Fujita .....                                                        | 1   |
| <i>Collaborative Work on Indonesian WordNet through Asian WordNet (AWN)</i>                                                 |     |
| Chairil Hakim, Budiono Budiono and Hammam Riza .....                                                                        | 9   |
| <i>Considerations on Automatic Mapping Large-Scale Heterogeneous Language Resources: Sejong Semantic Classes and KorLex</i> |     |
| Heum Park, Ae sun Yoon, Woo Chul Park and Hyuk-Chul Kwon .....                                                              | 14  |
| <i>Sequential Tagging of Semantic Roles on Chinese FrameNet</i>                                                             |     |
| Jihong LI, Ruibo WANG and Yahui GAO .....                                                                                   | 22  |
| <i>Augmenting a Bilingual Lexicon with Information for Word Translation Disambiguation</i>                                  |     |
| Takashi Tsunakawa and Hiroyuki Kaji .....                                                                                   | 30  |
| <i>Construction of bilingual multimodal corpora of referring expressions in collaborative problem solving</i>               |     |
| Takenobu Tokunaga, Ryu Iida, Masaaki Yasuhara, Asuka Terai, David Morris and Anja Belz ..                                   | 38  |
| <i>Labeling Emotion in Bengali Blog Corpus A Fine Grained Tagging at Sentence Level</i>                                     |     |
| Dipankar Das and Sivaji Bandyopadhyay .....                                                                                 | 47  |
| <i>SentiWordNet for Indian Languages</i>                                                                                    |     |
| Amitava Das and Sivaji Bandyopadhyay .....                                                                                  | 56  |
| <i>Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews</i>                                             |     |
| Choochart Haruechaiyasak, Alisa Kongthon, Pornpimon Palingoon and Chatchawal Sangkeetrakarn .....                           | 64  |
| <i>The Annotation of Event Schema in Chinese</i>                                                                            |     |
| Hongjian Zou, Erhong Yang, Yan Gao and Qingqing Zeng .....                                                                  | 72  |
| <i>Query Expansion for Khmer Information Retrieval</i>                                                                      |     |
| Channa Van and Wataru Kameyama .....                                                                                        | 80  |
| <i>Word Segmentation for Urdu OCR System</i>                                                                                |     |
| Misbah Akram and Sarmad Hussain .....                                                                                       | 88  |
| <i>Dzongkha Word Segmentation</i>                                                                                           |     |
| Sithar Norbu, Pema Choejey, Tenzin Dendup, Sarmad Hussain and Ahmed Muaz .....                                              | 95  |
| <i>Building NLP resources for Dzongkha: A Tagset and A Tagged Corpus</i>                                                    |     |
| Chungku Chungku, Jurmey Rabgay and Gertrud Faaß .....                                                                       | 103 |
| <i>Unaccusative/Unergative Distinction in Turkish: A Connectionist Approach</i>                                             |     |
| Cengiz Acarturk and Deniz Zeyrek .....                                                                                      | 111 |

|                                                                                                                                  |     |
|----------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>A Preliminary Work on Hindi Causatives</i>                                                                                    |     |
| Rafiya Begum and Dipti Misra Sharma .....                                                                                        | 120 |
| <i>A Supervised Learning based Chunking in Thai using Categorical Grammar</i>                                                    |     |
| Thepchai Supnithi, Chanon Onman, Peerachet Porkaew, Taneth Ruangrajitpakorn, Kanokorn Trakultaweekoon and Asanee Kawtrakul ..... | 129 |
| <i>A hybrid approach to Urdu verb phrase chunking</i>                                                                            |     |
| Wajid Ali and Sarmad Hussain .....                                                                                               | 137 |
| <i>Development of the Korean Resource Grammar: Towards Grammar Customization</i>                                                 |     |
| Sanghoun Song, Jong-Bok Kim, Francis Bond and Jaehyung Yang .....                                                                | 144 |
| <i>An Open Source Urdu Resource Grammar</i>                                                                                      |     |
| Shafqat Mumtaz Virk, Muhammad Humayoun and Aarne Ranta .....                                                                     | 153 |
| <i>A Current Status of Thai Categorical Grammars and Their Applications</i>                                                      |     |
| Taneth Ruangrajitpakorn and Thepchai Supnithi .....                                                                              | 161 |
| <i>Chained Machine Translation Using Morphemes as Pivot Language</i>                                                             |     |
| Li Wen, Chen Lei, Han Wudabala and Li Miao .....                                                                                 | 169 |
| Author Index                                                                                                                     |     |